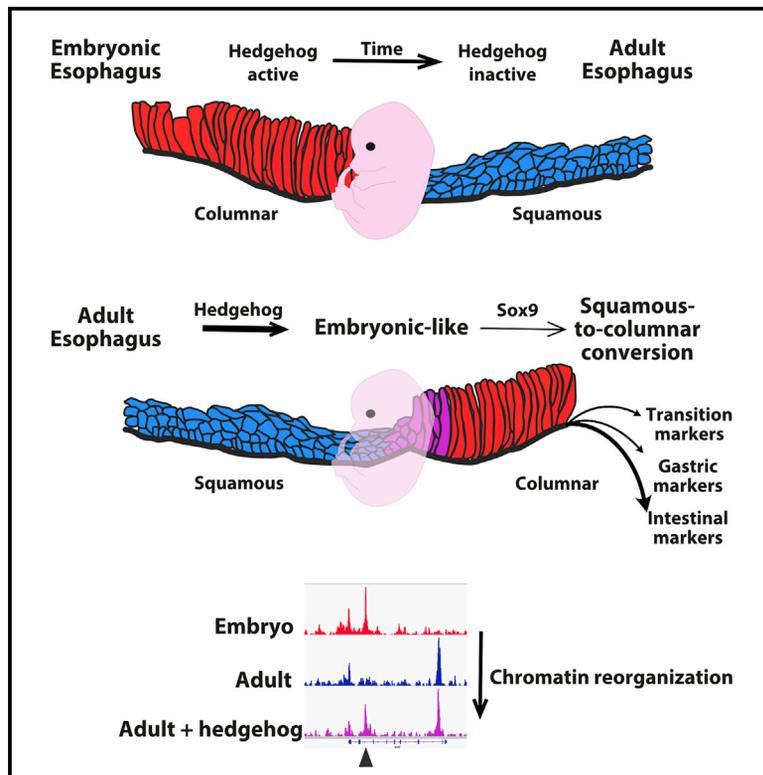


Reactivation of the Hedgehog pathway in esophageal progenitors turns on an embryonic-like program to initiate columnar metaplasia

Graphical abstract



Authors

Alizée Vercauteren Drubbel, Sheleya Pirard, Simon Kin, ..., Frédérick Libert, Sachiyo Nomura, Benjamin Beck

Correspondence

benjamin.beck@ulb.be

In brief

The Hedgehog pathway is activated in squamous progenitors from the foregut in a model of chronic reflux-induced columnar metaplasia. Activation of the Hedgehog pathway turns a subset of esophageal progenitors into columnar cells in a Sox9-dependent manner. This transcommitment involves a step of dedifferentiation into embryonic-like esophageal progenitors.

Highlights

- The Hedgehog pathway is activated in squamous foregut cells upon chronic reflux
- The Hedgehog pathway triggers the columnar conversion of a subset of esophageal cells
- Transcommitment involves a step of dedifferentiation into embryonic-like foregut cells
- Sox9 is required for squamous-to-columnar conversion but not for dedifferentiation

Article

Reactivation of the Hedgehog pathway in esophageal progenitors turns on an embryonic-like program to initiate columnar metaplasia

Alizée Vercauteren Drubbel,¹ Sheleya Pirard,¹ Simon Kin,¹ Benjamin Dassy,¹ Anne Lefort,¹ Frédéric Libert,¹ Sachiyo Nomura,² and Benjamin Beck^{1,3,4,*}

¹IRIBHM, ULB/Faculty of Medicine, 808 route de Lennik, 1070 Brussels, Belgium

²Department of Gastrointestinal Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

³WELBIO/FNRS Principal Investigator at IRIBHM, ULB/Faculty of Medicine, 808 route de Lennik, 1070 Brussels, Belgium

⁴Lead contact

*Correspondence: benjamin.beck@ulb.be

<https://doi.org/10.1016/j.stem.2021.03.019>

SUMMARY

Columnar metaplasia of the esophagus is the main risk factor for esophageal adenocarcinoma. There is a lack of evidence to demonstrate that esophageal progenitors can be the source of columnar metaplasia. In this study, using transgenic mouse models, lineage tracing, single-cell RNA sequencing, and transcriptomic and epigenetic profiling, we found that the activation of the Hedgehog pathway in esophageal cells modifies their differentiation status *in vivo*. This process involves an initial step of dedifferentiation into embryonic-like esophageal progenitors. Moreover, a subset of these cells undergoes full squamous-to-columnar conversion and expresses selected intestinal markers. These modifications of cell fate are associated with remodeling of the chromatin and the appearance of Sox9. Using a conditional knockout mouse, we show that Sox9 is required for columnar conversion but not for the step of dedifferentiation. These results provide insight into the mechanisms by which esophageal cells might initiate columnar metaplasia.

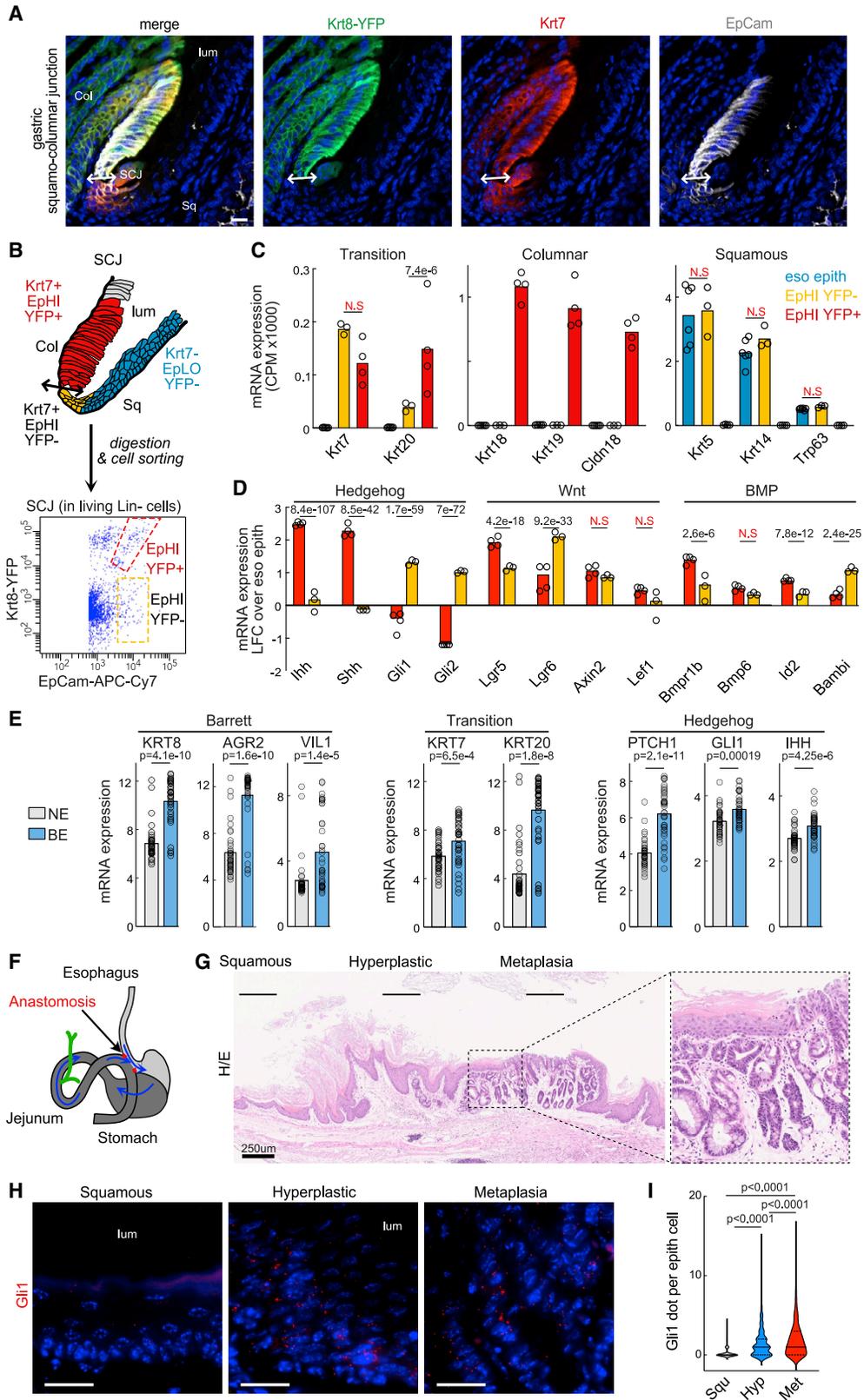
INTRODUCTION

Metaplasia is defined as the replacement of a fully differentiated cell type by another. There are several classical examples of metaplasia, such as intestinal metaplasia of the stomach, squamous metaplasia of the lung, and columnar metaplasia of the esophagus, also called Barrett's esophagus (BE) (Giroux and Rustgi, 2017). BE is considered a precancerous lesion at the origin of esophageal adenocarcinoma that increases by about 50 times the risk for this cancer (Reid et al., 2010). Nonetheless, the cellular origin of this metaplasia is complex.

Columnar metaplasia of the esophagus has been described as an adaptation to chronic gastro-esophageal reflux disease (GERD) (Souza, 2016). This chronic assault on the tissue would trigger the replacement of the distal esophageal epithelium by a columnar epithelium to protect the esophagus from the acidic content of the stomach. It has been demonstrated that Lgr5+ columnar progenitors (Quante et al., 2012) and embryonic-like Krt7+ columnar cells (Wang et al., 2011) at the squamo-columnar junction (SCJ) could contribute to BE-like metaplasia in the mouse. Recently, lineage tracing of p63+ squamous cells in a surgical model of GERD demonstrated that keratinocytes can also give rise to BE-like metaplasia (Jiang et al., 2017), suggesting that several cell populations are involved in the development of BE. Nonetheless, the role of esophageal progenitors

in the development of columnar metaplasia *in vivo* is still uncertain.

During its development, the esophagus is first lined with a simple columnar epithelium. Around embryonic day (E) 13, the transcription factor p63 triggers a wave of squamous differentiation in the mouse Krt8+ esophagus epithelium (Wang et al., 2011). This columnar-to-squamous conversion is orchestrated by many different signaling pathways, such as Hedgehog (HH) (Litingtung et al., 1998; Motoyama et al., 1998), BMP (Rodríguez et al., 2010) and Wnt/b-catenin (Woo et al., 2011). Interestingly, activation of the same pathways in esophageal keratinocytes *in vitro* alters the process of squamous differentiation, suggesting that some may promote transdifferentiation (Kong et al., 2011a; Milano et al., 2007; Wang et al., 2010). However, it is unclear to what extent reactivation of such pathways in adult esophagus *in vivo* would turn squamous cells into embryonic-like columnar cells and whether development of metaplasia actually requires a step of dedifferentiation. In this study, we show that the HH pathway is activated at the SCJ under physiological conditions and more broadly in squamous epithelial cells upon chronic acid reflux. We found that the activation of this pathway alters the squamous differentiation program in the majority of esophageal cells and induces a full squamous-to-columnar conversion in a subset of progenitors. Interestingly, an embryonic-like epigenetic and transcriptomic program



(legend on next page)

precedes the columnar conversion, suggesting that keratinocytes need to be dedifferentiated before activating another differentiation program. Conditional knockout *in vivo* demonstrates that Sox9 plays a pivotal role in the process of columnar differentiation but not in the embryonic-like epigenetic and transcriptional programs. Hence, our data highlight the mechanisms modulating cellular plasticity in a subset of progenitors that may constitute the very first step of transdifferentiation and metaplasia development from esophageal cells.

RESULTS

The HH pathway is activated in some squamous cells under physiological and pathological conditions

Co-immunostaining shows that the Krt7 transition epithelium from the SCJ is composed of two populations: Krt8+/Krt14– columnar cells and Krt8–/Krt14+ squamous cells. Both Krt7+ populations express EpCam (Figures 1A and S1). In the Krt8-YFP knockin mouse model (K8-YFP), columnar transition epithelium co-expresses YFP and EpCam, while the squamous transition epithelium expresses EpCam only (Figure 1A). We used flow cytometry to sort these two populations (EpHI/YFP+ and EpHI/YFP–) and profile them using RNA sequencing (RNA-seq) (Figure 1B). RNA-seq data confirmed that both populations express Krt7, but columnar markers (*Krt8*, *Krt18*, and *Cldn18*) are enriched in EpHI/YFP+ cells, and squamous markers (*Krt14*, *Krt5*, and *Trp63*) are enriched in EpHI/YFP– cells (Figure 1C). Several transcripts suggest that the Wnt and BMP signaling pathways are activated in both populations compared with control esophageal cells (Figure 1D). Interestingly, the HH pathway target genes *Gli1* and *Gli2* are upregulated in the squamous Krt7+ epithelium, while *Shh* and *Ihh* are among the most upregulated ligands in the neighboring columnar transition epithelium. These data show that the Krt7+ squamous cells that are prone to generate columnar metaplasia *in vivo* (Jiang et al., 2017) are localized in a HH stimulating niche constituted by the columnar transition epithelium.

Analysis of microarray data from human BE and normal esophagus samples shows that classical BE markers (*KRT8*, *AGR2*, *VIL1*, *KRT7*, and *KRT20*) are upregulated together with HH-related genes (*IHH*, *PTCH1*, and *GLI1*), suggesting that this pathway is activated in metaplasia (Figures 1E and S1). To determine whether the HH pathway may be associated with the development of columnar metaplasia *in vivo*, we investigated the expression of the HH target genes *Gli1* and *Shh* in a model of

surgically induced chronic acid reflux (Figures 1F–1I and S1). *In situ* hybridization shows the overexpression of both transcripts in the columnar tissue and the adjacent hyperplastic squamous tissue compared with the normal squamous epithelium. These data show that the HH pathway is activated in squamous cells during metaplasia development.

Activation of the HH pathway in esophageal progenitors modifies cell fate *in vivo*

To determine the consequences of HH pathway activation in esophageal keratinocytes, we used the Krt5-CreER^{T2}:R26SmoM2/EYFP mice (K5:Smo) model to induce the expression of the constitutively active form of Smoothed (SmoM2) in basal progenitors (Figure 2A). SmoM2 stimulates expression of classical HH targets *Ptch1* and *Gli1*, as measured by *in situ* hybridization (Figure S2). Immunostaining shows that SmoM2+ esophageal cells express a high level of EpCam, just like the Krt7+ transition epithelium (Figure 2B), and this upregulation (EpHI in opposition to EpLO) is confirmed by flow cytometry (Figure S2). Using qPCR, we found that *SmoM2*, *Ptch1*, and *Gli1* transcripts are enriched in EpHI cells compared with EpLO cells (Figure 2C). As the YFP fused to SmoM2 is undetectable by fluorescence-activated cell sorting (FACS), EpCam can thus be used as a surrogate of SmoM2 expression. Analysis of several microarray datasets shows that EPCAM is also upregulated in human columnar metaplasia of the esophagus (Figure S2).

Comparison of EpHI cells with EpLO and control esophageal cells using RNA-seq 3, 8, and 12 weeks after tamoxifen (TAM) administration shows a progressive modification of gene expression (Figures 2D and S2). Gene set enrichment analysis (GSEA) of significantly modified transcripts at these time points suggests a repression of the squamous differentiation program that starts after 3 weeks and an enrichment of genes related to tissue morphogenesis after 12 weeks (Figure S2). Notably, squamous markers (*Krt14*, *Krt4*, and *Krt13*) are downregulated, and *Krt20* appears in esophageal cells 3 weeks after SmoM2 expression (Figure 2E). Several Wnt and BMP target genes are strongly upregulated in EpHI cells compared with control cells, suggesting that both pathways are activated (Figure S2). After 12 weeks, columnar markers (*Krt7*, *Krt8*), several mucins (*Muc5ac*, *Muc2*, *Muc6*), and some BE markers (*Agr2*, *Vil1*) are significantly upregulated at the mRNA level (Figures 2E and S2). At this time point, 9% of the transcripts stimulated by the HH pathway in esophageal cells are also enriched in squamous SCJ cells. Surprisingly, even more transcripts (24%) are shared with columnar SCJ cells,

Figure 1. The Hedgehog pathway is activated in some squamous cells under physiological and pathological conditions

- (A) Co-immunostaining for Krt8-YFP, Krt7, and EpCam at the gastric SCJ from K8-YFP mouse.
(B) Scheme of the SCJ and representative FACS plot for Krt8-YFP and EpCam in SCJ cells from K8-YFP mice.
(C) Expression of transition, columnar, and squamous markers in EpHI/YFP+, EpHI/YFP–, and control esophageal cells measured by RNA-seq.
(D) Expression of markers of the HH, Wnt, and BMP pathways in EpHI/YFP+ and EpHI/YFP– measured by RNA-seq.
(E) Expression of BE markers (*KRT8*, *AGR2*, and *VIL1*), transition epithelium (*KRT7* and *KRT20*), and HH pathway (*PTCH1*, *GLI1* and *IHH*) in human BE (BE; n = 40) and esophagus samples (NE; n = 40) (GEO: GSE39491).
(F) Scheme of esophago-gastro-jejunostomy (EGJ).
(G) H&E staining in EGJ samples.
(H) Fluorescence *in situ* hybridization (FISH) for *Gli1* in EGJ samples.
(I) Quantification of *Gli1* expression measured using FISH in normal squamous epithelium (n = 100 cells/sample), hyperplastic epithelium (n = 200), and columnar metaplastic epithelium (n = 200). Data are represented as the sum of eight samples.
Scale bars, 20 μ m. Col, columnar; Sq, squamous; SCJ, squamo-columnar junction; lum, lumen; LP, lamina propria. p values were calculated using the Mann-Whitney test.

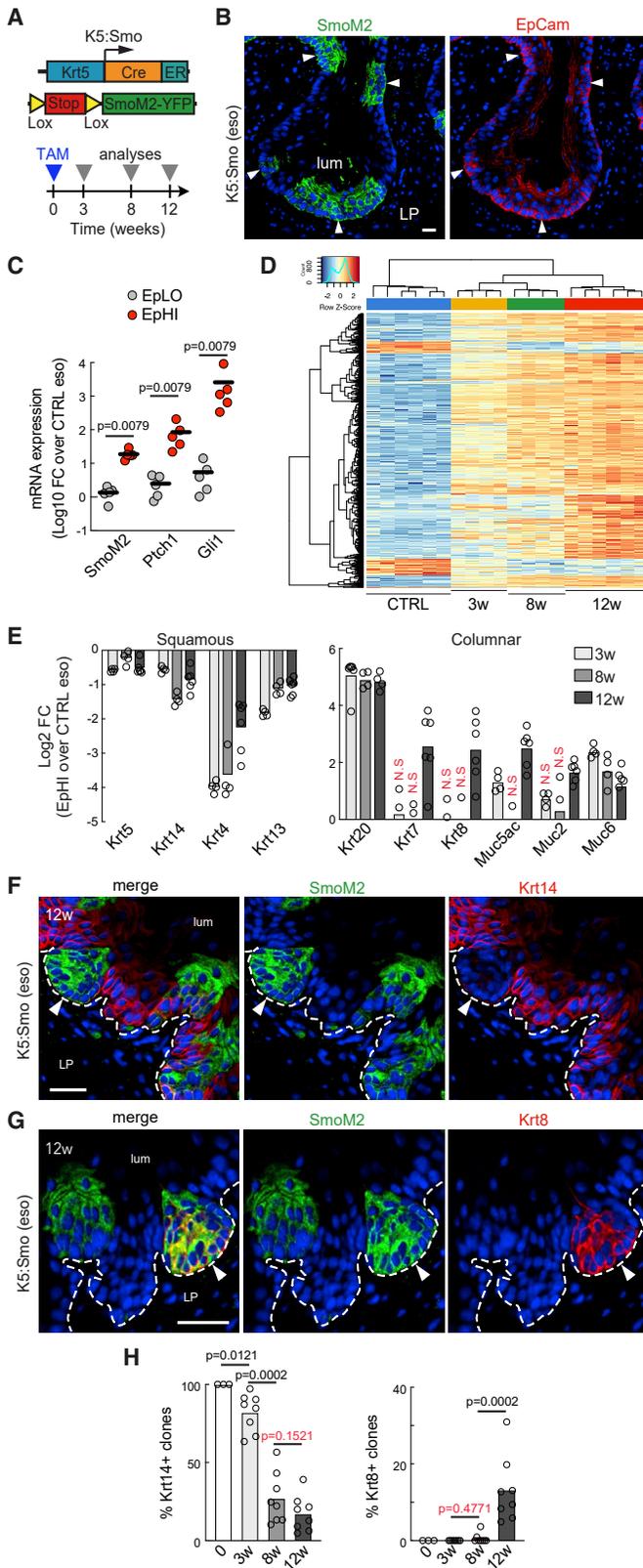


Figure 2. Activation of the Hedgehog pathway in esophageal progenitors modifies cell fate

(A) Genetic strategy.

(B) Co-immunostaining for SmoM2 and EpCam in K5:Smo esophagus 6 weeks after TAM.

(C) *SmoM2*, *Ptch1*, and *Gli1* expression measured using qPCR in EpHI and paired EpLO sorted cells (n = 5). Data are represented as the log₁₀ fold change over control esophagus epithelium.

(D) Heatmap of the 500 most variable genes between CTRL and K5:Smo EpHI cells 12 weeks after TAM, for CTRL and K5:Smo EpHI cells (3, 8, and 12 weeks after TAM) samples.

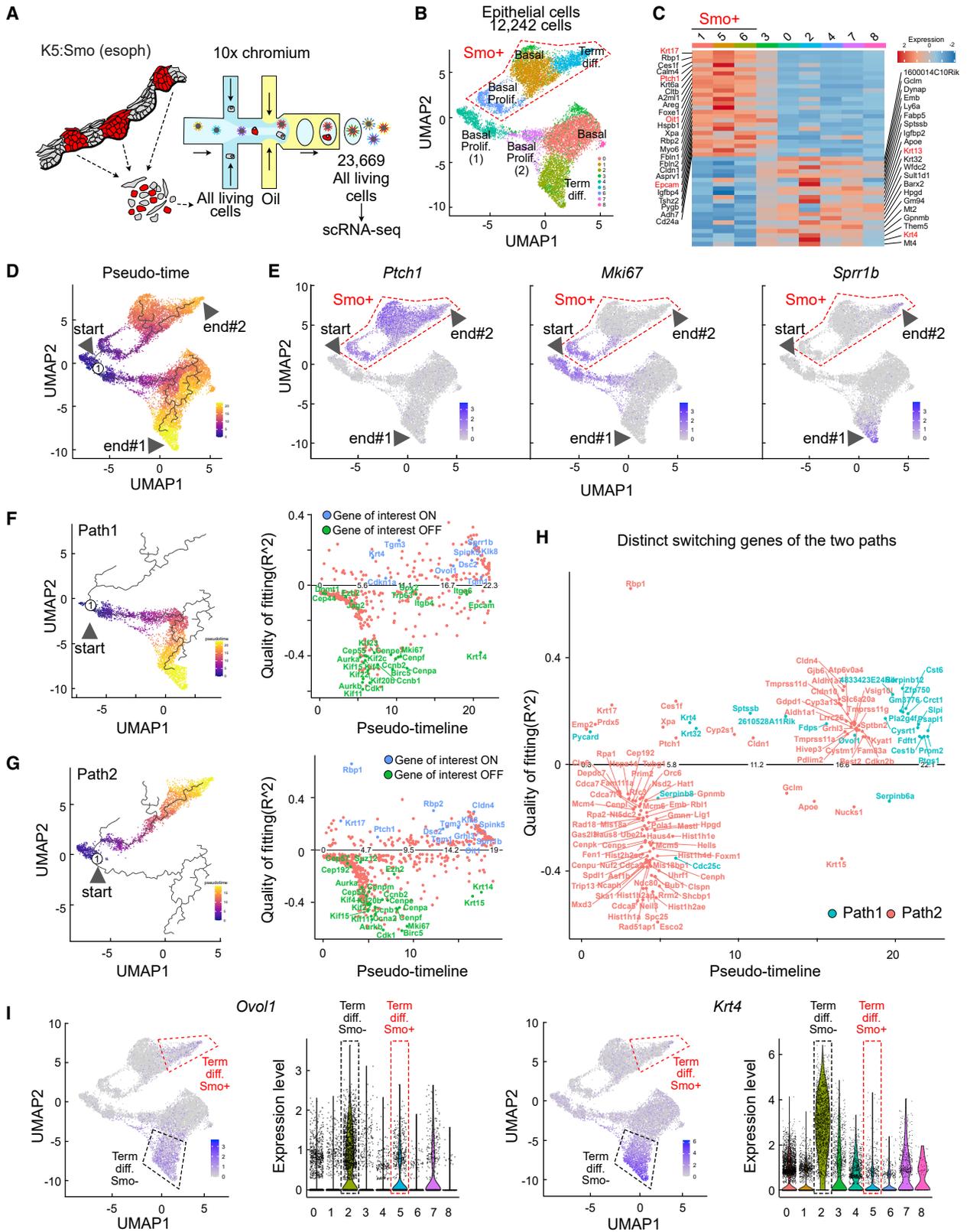
(E) Expression of squamous and columnar markers measured using RNA-seq in FACS-sorted EpHI cells 3, 8, and 12 weeks after TAM.

(F) Co-immunostaining for SmoM2 and Krt14 in K5:Smo esophagus 12 weeks after TAM.

(G) Same as in (F) with Krt8.

(H) Quantification of the number of Krt14+ clones and the number of clones containing at least one Krt8+ cell at 0, 3, 8, and 12 weeks after TAM (n = 8/condition, n = 3 for CTRL).

Scale bars, 20 μm. lum, lumen; LP, lamina propria; TAM, tamoxifen. p values were calculated using the Mann-Whitney test.



(legend on next page)

including transcription regulators (*Runx2*, *Hmga2*, *Foxa2*, and *Ascl1*) (Figure S2).

We followed the fate of SmoM2⁺ individual clones over time. All SmoM2⁺ clones are Krt20⁺, while some still express Krt14 (Figure S2). Between 3 and 12 weeks after TAM administration, the squamous marker Krt14 is progressively lost (Figures 2F and 2H), while the columnar marker Krt8 appears in some clones only after 12 weeks (Figures 2G and 2H). These data suggest that the fate of SmoM2⁺ cells is heterogeneous and that a subset of esophageal progenitors might be competent to undergo a squamous-to-columnar conversion.

The HH pathway modifies the squamous differentiation program

To identify the different cell phenotypes induced by the HH pathway, we used single-cell RNA-seq (scRNA-seq) on FACS-sorted esophageal cells from K5:Smo mice 12 weeks after induction (Figure 3A). Clustering on 12,242 epithelial cells highlighted nine different cell groups. Three clusters (1, 5, and 6) are enriched for *Ptch1* and *EpCam*, suggesting an activation of the HH pathway (Figures 3B, 3C, 3E, and S3). We investigated the expression of markers of proliferation (*Mki67*, *Cenpa*), basal epithelial cells (*Krt14*), suprabasal differentiated cells (*Krt13*), and terminal differentiation (*Sprr1b*) to identify and annotate cell clusters (Figures 3B, 3E, and S3).

Pseudo-time analysis highlights two hypothetical paths both starting from *Ptch1*-negative basal proliferating cells and finishing in *Sprr1b*⁺ terminally differentiated cells (Figures 3D and 3E). One path links cell clusters with low *Ptch1* expression (Smo⁻/path 1), and the second one connects clusters in which the HH pathway is activated, as shown by *Ptch1* expression (Smo⁺/path 2). We analyzed the genes that are significantly modified along pseudo-time for these two paths (Figures 3F–3I and S3). In path 1, Smo⁻ cells seem to undergo a classical process of squamous differentiation. They first exit the cell cycle (downregulation of *Mki67*, upregulation of *Cdkn1a*), repress their basal fate (downregulation of *Itgb4*), and start expressing markers of early (*Krt4*) and terminal (*Sprr1b*) differentiation. Although the process is quite similar in Smo⁺ cells along path 2 (Figures 3F and 3G), some markers of differentiation, such as *Krt4* and *Ovol1*, do not appear, and others, such as *Sprr1b*, seem to be expressed in a smaller proportion of cells (Figures 3E, 3H, 3I, and S3), suggesting that the HH pathway impairs the process of squamous differentiation.

Activation of the HH pathway in esophageal progenitors leads to two distinct fates

To increase the resolution of our analysis, we then analyzed 14,763 FACS-sorted EpHI cells from K5:Smo mice and a subset

of 2,706 cells on the basis of YFP mRNA expression. Consistent with our clonal analysis data, scRNA-seq highlights a cluster of Krt8⁺ cells (cluster 7) (Figures 4A, 4B, and S4). These cells are enriched in *Krt8*, *Krt18*, and *Krt19* and characterized by a downregulation of the squamous differentiation master regulator *Trp63* (Figures 4B, 4C, and S4). To determine the connection between the cluster 7 and the others, we aggregated total epithelial cells and FACS-sorted EpHI cells from K5:Smo mice. We annotated the clusters from these 26,495 epithelial cells and performed pseudo-time analysis. Using two different methods, we could identify two paths starting from basal cells and finishing in either Smo⁺ differentiated cells (path 1/curve 1) or in Smo⁺/Krt8⁺ columnar-like cells (path 2/curve 2) (Figures 4E and 4F). Characterization of the differentially expressed transcripts along these two paths suggests that in path 1, cells seem to undergo a program of squamous differentiation (repression of *Itga6* and *Itgb4* and upregulation of *Abca12* and *Sprr1b*); in path 2, esophageal cells turn on a different program that leads to Krt8 expression (Figures 4G–4J and S4) and that is associated with the appearance of Sox9 (Figures 4I–4K). However, the phenotype of these *de novo* Krt8⁺ cells is unclear.

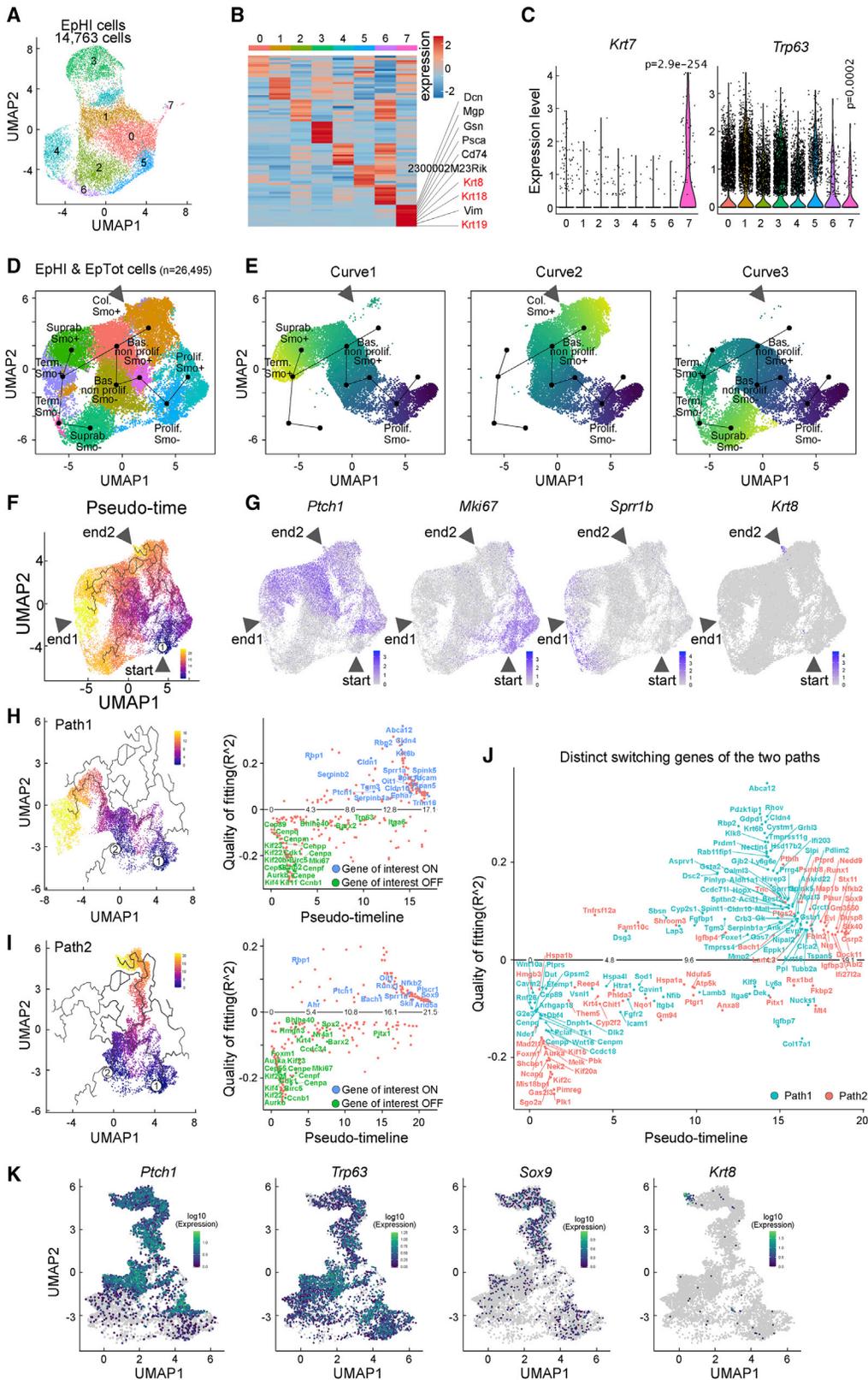
The HH pathway activates an embryonic-like program and the appearance of intestinal transcripts in esophageal progenitors

We used the K8-YFP knockin mouse model to isolate and profile the *de novo* Krt8⁺ cells in the K5:Smo esophagus by RNA-seq (Figures 5A and 5B). About 60% of the upregulated genes found in Krt8^{neg} (EpHI/K8^{neg}) cells are also upregulated in Krt8⁺ cells (EpHI/K8⁺) (Figure 5C). GSEA shows that genes related to “brush border” and “digestion” are enriched in EpHI/K8⁺ cells but not in EpHI/K8^{neg} (Figure 5D). Comparison of HH-induced EpHI/K8⁺ with EpHI/K8^{neg} cells shows no modifications of HH, BMP, or Wnt target genes but highlights the upregulation of several markers of columnar cells (*Krt7*, *Muc5ac*, *Tff1*) and the downregulation of squamous markers (*Krt5*, *Krt14*, *Trp63*) in EpHI/K8⁺ cells (Figures 5E and S5). These features suggest a squamous-to-columnar conversion, but the type of columnar epithelium generated is unclear.

To determine whether *de novo* Krt8⁺ cells would be similar to the fetal esophagus columnar epithelium, we first compared FACS-sorted embryonic esophageal cells at E13.5 with adult esophageal epithelial cells by RNA-seq (Figure S5). Embryonic cells express high levels of *Krt8*, *Krt7*, and *Krt20*, and of HH, Wnt, and BMP target genes, but low levels of the differentiation marker *Krt13* (Figure S5). Embryonic esophageal cells thus share similarities with HH-induced cells. However, *de novo* Krt8⁺ cells may be more similar to differentiated columnar

Figure 3. The Hedgehog pathway modifies the squamous differentiation program

- Experimental design.
- Annotated UMAP of epithelial cells isolated from K5:Smo mice showing the segregation between SmoM2⁻ and SmoM2⁺ cells.
- Heatmap of the 50 most differentially regulated genes between SmoM2⁺ and SmoM2⁻ clusters (the mixed cluster 3 has been removed from the differential analysis) (false discovery rate [FDR] < 0.05).
- Pseudo-time analysis on all epithelial cells from K5:Smo mice.
- Feature plots of transcripts differentially enriched in the UMAP.
- Plot of the genes whose expression changes along path 1. Switched-on genes are plotted above the line, while switched-off genes are below.
- Same as in (F) along path 2.
- Plot of the distinct switching genes of the two paths (quality fitting threshold = 0.1).
- Feature plots and violin plots for two transcripts differently expressed between the two paths.



(legend on next page)

epithelia. Indeed, three types of columnar metaplasia have been described in the esophagus: cardia-like (transition epithelium), oxyntic (gastric), and intestinal (Spechler, 2012). We therefore profiled EpCam+/YFP+ sorted epithelial cells from these tissues in K8-YFP mice and listed all the significantly up-regulated genes compared with adult esophagus. We defined four signatures using RNA-seq: embryonic, transitional, gastric, and intestinal (Figures 5F and S5). We then listed the transcripts that are specific of each signature (embryo only, junction only, stomach only, and intestine only), as well as the transcripts that are common to these four columnar epithelia (columnar) (Figure 5F). Upregulated genes in EpHI/K8neg are preferentially found in the “embryo only” signature, while upregulated genes in EpHI/K8+ cells are also found in “columnar” and “intestine only” signatures (Figures 5G and S5). GSEA confirms a significant enrichment of the “embryonic only” signature in EpHI/K8neg and EpHI/K8+ cells, while the “intestine only” signature is enriched in EpHI/K8+ cells only (Figure 5H). Comparison of EpHI RNA-seq data at 3, 8, and 12 weeks with the same signatures shows that the embryonic markers decrease between 3 and 12 weeks after induction, while the intestinal markers increase, suggesting that the acquisition of embryonic features precedes the appearance of intestinal transcripts (Figure S5). Hence, upon activation of the HH pathway, adult esophageal cells turn on a transcriptional program that resembles the one from columnar embryonic esophagus. Then, some progenitors start expressing columnar and intestinal markers, although there is no sign of specialized glandular epithelium.

Activation of the HH pathway induces opening of embryonic esophagus-specific chromatin regions to facilitate transcommitment

It has been suggested that epigenetic regulators may fuel chromatin remodeling to allow the binding of transcription factors, thus creating a framework for metaplasia (Kaz et al., 2011, 2015). We have therefore used assay for transposase-accessible chromatin using sequencing (ATAC-seq) to profile the open chromatin regions of esophageal cells in control condition and following activation of the HH pathway (Figure 6A). We found 89,407 peaks in CTRL cells and 98,160 peaks in EpHI cells, 12 weeks after induction (Figure 6B). Following activation of the HH pathway, about 20% of open chromatin regions get closed, and about 25,000 new regions are opened. One-fifth (22%) of the transcripts upregulated by the HH pathway are associated with significant chromatin opening compared with

control, including genes coding for markers such as *Krt7* and *Lgr5* (Figures 6C, 6D, and S6).

The *de novo* motif discovery in the chromatin regions significantly more opened in EpHI cells than in control condition revealed that the two most frequent motifs match the best with sex-determining region Y (SRY)-related high-mobility group (HMG) box (Sox) and Forkhead box (Fox) transcription factor families (Figure 6E). The known motifs enrichment analysis highlighted putative binding sequences for several transcription factors involved in esophageal cell physiology (AP-1, Sox2, Klf4) but also some that have been reported to be expressed in columnar metaplasia, such as Sox9 and Foxa2 (Figure S6).

To determine whether the embryonic-like transcription program found in EpHI cells could be due to the reopening of chromatin regions that are specifically opened in fetal esophagus epithelium, we profiled FACS-sorted embryonic esophageal cells at E13.5 using ATAC-seq (Figure 6F). These data highlight 26,763 chromatin regions that are significantly more opened in embryonic epithelial cells than in adult epithelium. Strikingly, 54% of the chromatin regions that are more opened following activation of the HH pathway belong to this embryonic epigenetic signature (Figure 6G). The *de novo* motif discovery in the common open chromatin regions indicates that Sox15, FoxP2, and Gli2 are the three most significant motifs (Figure 6H). These data suggest that the HH pathway induces epigenetic modifications that may facilitate activation of new transcriptional programs.

As the Sox transcription factor family-related motif is the most represented in HH-induced open chromatin regions, we investigated the expression of these transcription factors in EpHI cells using RNA-seq. We found that Sox4, Sox21, Sox6, and Sox9 are the four most abundant Sox family members (Figure 6I). Pseudo-time analysis of our scRNA-seq data shows that Sox9 is upregulated in esophageal cells along squamous-to-columnar conversion (Figures 4I–4K). Moreover, Sox9 is the most expressed in epithelial cells from the gastrointestinal (GI) tract but not in embryonic esophagus (Figure 6J). Immunostaining shows that under control conditions, Sox9 is restricted to some epithelial cells in the GI tract but is virtually absent from the squamous epithelium in the foregut and from embryonic esophageal cells at E13.5 (Figures S5 and S6). Our ATAC-seq data show that almost 15% of the chromatin regions that are opened following activation of the HH pathway contain binding motifs for Sox9, while CTRL-specific chromatin regions do not (Figure S6). Interestingly, the majority of the genes from the columnar signature

Figure 4. Activation of the Hedgehog pathway in esophageal progenitors leads to two distinct fates

- (A) UMAP of EpHI cells from K5:Smo mice.
- (B) Heatmap showing the ten most expressed transcripts in each of the eight groups compared with the others. The transcripts enriched in group 7 are listed.
- (C) Violin plots of transcripts upregulated and downregulated in group 7 compared with remaining ones. p value is calculated using the Wilcoxon test.
- (D) Annotated UMAP of cells coming from the fusion of all epithelial cells (Figure 3B) and EpHI cells from K5:Smo mice (A).
- (E) Pseudo-time analysis of this UMAP using Slingshot. Starting point is in the proliferative basal SmoM2– cells. Curves 1, 2, and 3 end respectively in the terminal differentiated Smo+ cluster, the columnar-like SmoM2+ cluster, and the suprabasal Smo– cluster.
- (F) Pseudo-time analysis of the UMAP using monocle. Starting point is in the proliferative basal SmoM2– cells. End 1 and end 2 are respectively in terminal differentiated and columnar SmoM2+ cells.
- (G) Feature plots of transcripts differentially enriched in the UMAP.
- (H) Plot of the genes whose expression changes along path 1. Switched-on genes are plotted above the line, and those switching off are below.
- (I) Same as in (F) but along path 2.
- (J) Plot of the distinct switching genes of the two paths (quality fitting threshold = 0.05).
- (K) Plots depicting transcripts expression along path 2.

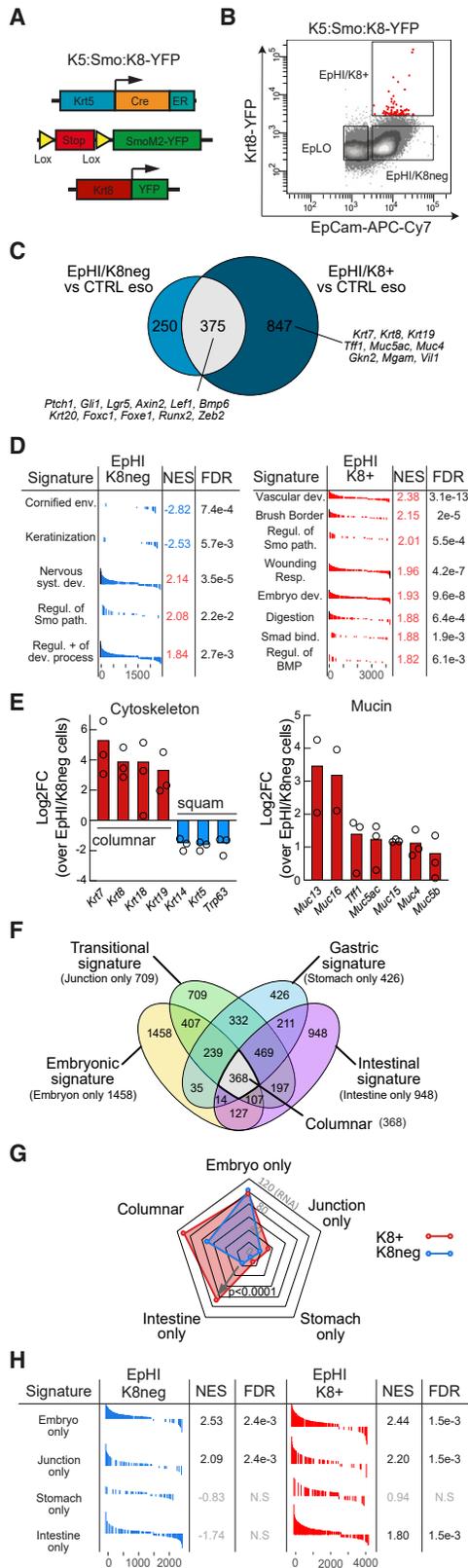


Figure 5. The Hedgehog pathway activates an embryonic-like program and the appearance of intestinal transcripts in esophageal progenitors

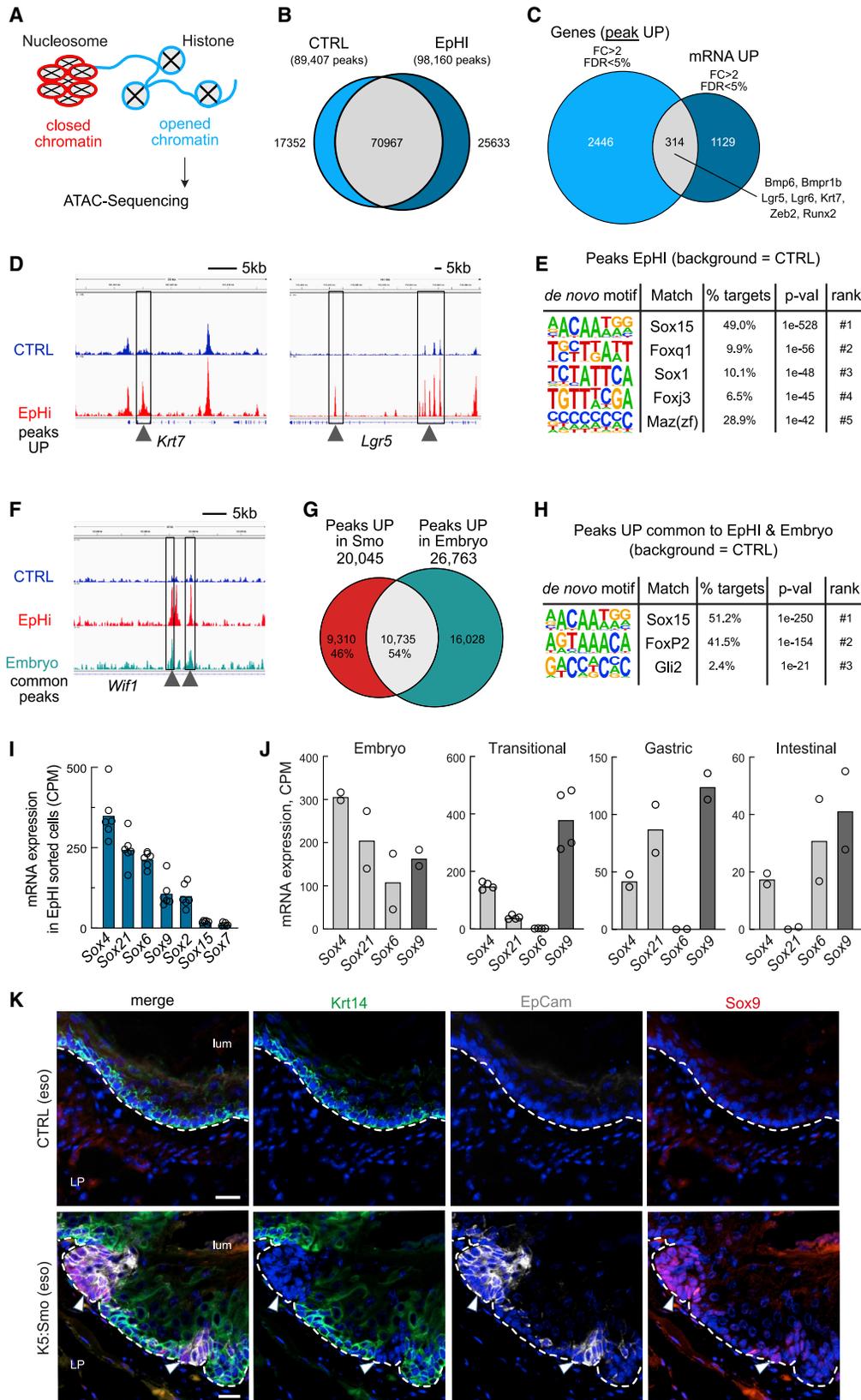
(A) Genetic strategy. (B) Representative FACS plot showing the EpCam and Krt8-YFP co-staining in K5:Smo:K8-YFP mouse esophageal cells 12 weeks after TAM. (C) Venn diagrams between the transcripts that are significantly upregulated in EpHI/K8neg and EpHI/K8+ compared with control esophageal cells. (D) Gene set enrichment analysis (GSEA) of the significantly modified transcripts in EpHI/K8neg and EpHI/K8+ compared with control esophageal cells. (E) Expression of columnar and squamous keratins as well as mucins measured using RNA-seq in EpHI/K8+ compared with paired EpHI/K8neg (n = 3) 12 weeks after TAM. (F) Venn diagram between embryonic, transitional, gastric, and intestinal transcriptional signatures. The transcripts that belong to only one of these signatures constitute the “embryo only,” “junction only,” “stomach only,” and “intestine only” signatures. The intersection among all signatures is the “columnar” signature. (G) Radar plot showing the number of transcripts from the “embryo only,” “junction only,” “stomach only,” “intestine only,” or “columnar” signatures upregulated in EpHI/K8+ or EpHI/K8neg compared with control adult esophagus. (H) Measure of the enrichment of the four different signatures “only” in EpHI/K8neg and EpHI/K8+ cells using GSEA.

(Figure 5F) contain binding motifs for Sox9 (Figure S6). Immunostaining show that Sox9 protein appears specifically in EpCam-positive cells from K5:Smo mouse esophagus (Figure 6K). Sox9 locus does not contain Gli motifs but contain a Smad1/5 motif (Figure S6), suggesting that it may be a BMP target gene indirectly regulated by HH. HH-induced chromatin modifications may thus facilitate the binding of newly expressed transcription factors such as Sox9, which would in turn drive squamous-to-columnar conversion. In line with this, SOX9 is significantly upregulated in human BE samples compared with control esophagus (Figure S6).

Sox9 expression is required for the squamous-to-columnar conversion of esophageal progenitors *in vivo*

Although Sox9 chromatin regions are open in both CTRL and EpHI cells, Sox9 is upregulated in EpHI cells compared with control esophageal cells. Its expression is similar in EpHI/K8+ and EpHI/K8neg cells (Figure S7). We found Sox9 binding motifs in open chromatin regions of several columnar markers, such as Krt7, Krt8, Krt18, and Krt20, in EpHI cells 12 weeks after TAM administration (Figure S7). To determine whether Sox9 drives HH-induced squamous-to-columnar conversion, we combined the inducible activation of the HH pathway with the deletion of Sox9 specifically in esophageal progenitors (K5:Smo:Sox9cKO) (Figure 7A). First, we confirmed that Sox9 is absent from K5:Smo:Sox9cKO esophageal cells following TAM administration (Figure 7B). Then, we followed the fate of SmoM2+ cells in both models. Interestingly, while K5:Smo esophagi are characterized by the appearance of Krt20+ and Krt8+ cells, these cells are absent from K5:Smo:Sox9cKO. This is not the consequence of the loss of SmoM2+ cells, as the percentage of EpHI cells is increased in K5:Smo:Sox9cKO (Figures 7C and S7).

We compared EpHI cells from K5:Smo:Sox9cKO and K5:Smo mice using RNA-seq 12 weeks after TAM induction and found that Sox9 deletion represses 20% of HH-induced genes, preferentially the ones that are related to columnar differentiation



(legend on next page)

(*Krt8*, *Krt7*, *Foxa2*, *GATA6*, *Lef1*) (Figures 7D–7F). Representation of the transcriptome from K5:Smc:Sox9cKO EpHI cells, control esophageal cells, embryonic esophageal cells, EpHI/*Krt8*⁺, and EpHI/*Krt8*^{neg} cells on a multidimensional plot shows that the Sox9cKO cells are very similar to the EpHI/*Krt8*^{neg} cells (Figure 7G). These data suggest that Sox9 regulates the acquisition of the columnar markers and not the dedifferentiation program. Indeed, the upregulated RNA in Sox9cKO cells belong preferentially to the embryonic signature (Figures 7H and S7). In the same line, ATAC-seq shows that 51% of the chromatin regions that are more opened in K5:Smc:Sox9cKO EpHI cells than in control esophageal cells are common with fetal esophagus epithelium (Figures 7I and 7J). These results suggest that Sox9 is not fully required for the acquisition of an embryonic epigenetic program. However, Sox9 is required for the process of squamous-to-columnar conversion that follows the step of dedifferentiation in HH-stimulated esophageal cells *in vivo* (Figure 7K).

DISCUSSION

A growing body of evidence has established that cell identity is not completely fixed (Smith et al., 2016). First, transcription factors capable of generating specific cell types such as cardiomyocytes or neurons have been identified (Ieda et al., 2010; Takahashi and Yamanaka, 2006; Vierbuchen et al., 2010). Also, a combination of transcriptomic and epigenetic studies have elucidated the mechanisms driving cell reprogramming (Buganim et al., 2012; Liu et al., 2017; Treutlein et al., 2016). Together, these studies paved the way for understanding basis of lineage conversion and cell plasticity.

Multiple cell populations at the origin of columnar metaplasia

Metaplasia seems to occur in tissues exposed to environmental agents, such as air in the lungs and trachea, and food in the GI tract, which can be injurious. As a consequence, the epithelial structure of the resident tissue adapts through metaplasia and therefore constitutes an interesting pathological model of cell plasticity. In the esophagus, a mixture of acid and bile would induce mucosal injury, inflammation, and oxidative stress, thereby creating an environment permissive for columnar metaplasia (Chen et al., 2000; Feng et al., 2017; Jenkins et al., 2007; Song et al., 2007). These columnar metaplasia can have three

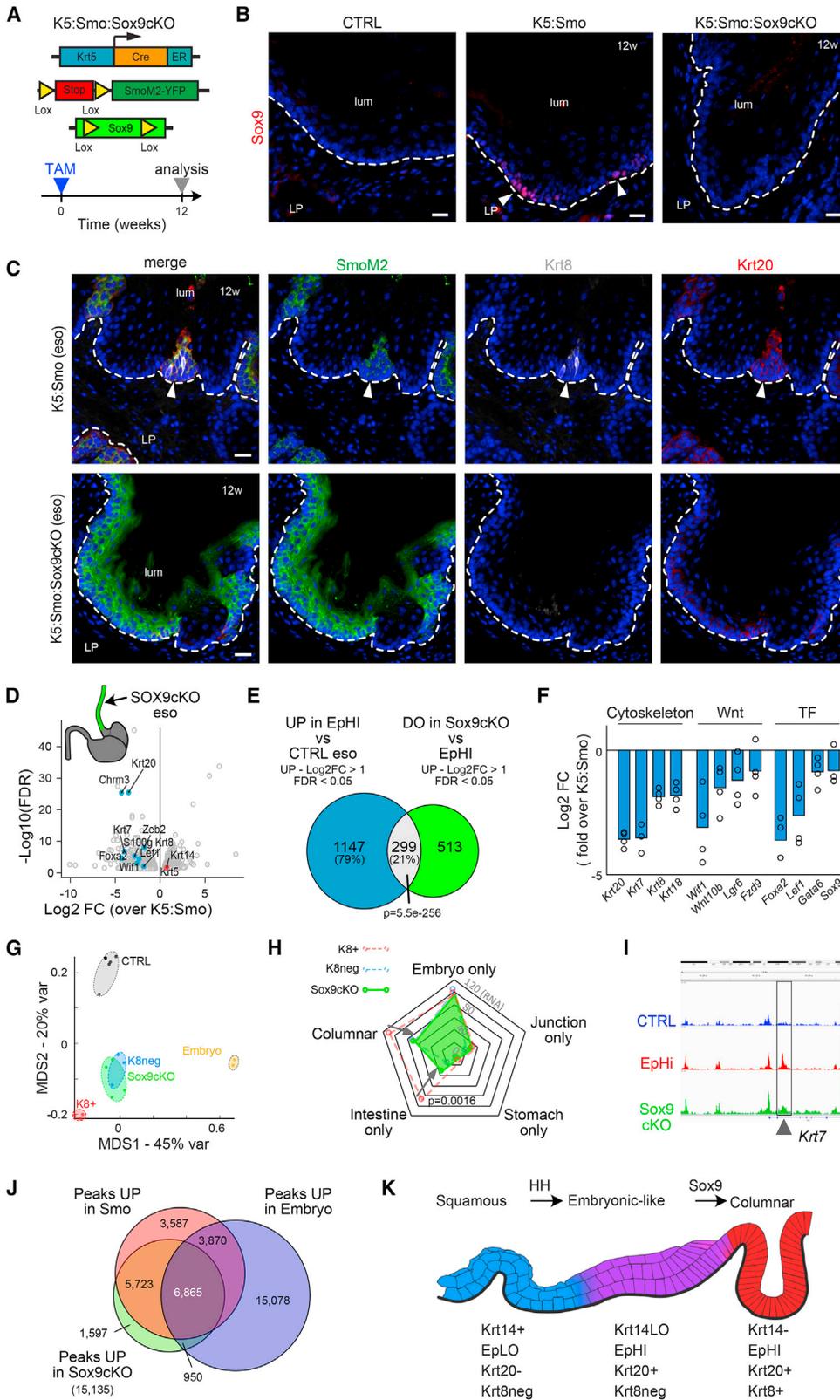
forms: cardiac, oxyntic, and intestinal. Although all types of columnar metaplasia are considered BE in some countries, in the United States at least, only intestinal metaplasia with goblet cells is called BE because it has the highest probability to progress toward dysplasia and adenocarcinoma (Spechler, 2012). There is compelling evidence that columnar cells at the SCJ are competent to initiate BE-like metaplasia *in vivo* (Quante et al., 2012). Because BE samples share similarities with embryonic esophageal cells, it has been reported that BE may arise from embryonic-like cells localized at the SCJ (Wang et al., 2011). A recent study even demonstrated that a subset of *Krt7*⁺ squamous cells at the SCJ are capable of transdifferentiation to generate BE-like metaplasia upon overexpression of *Cdx2* (Jiang et al., 2017). Nonetheless, there is a lack of evidence, at least in genetic mouse models, to demonstrate that esophageal progenitors can also be the source of columnar metaplasia (Giroux and Rustgi, 2017). Our data show that some esophageal progenitors keep the plasticity allowing a conversion into a columnar epithelium *in vivo*. This plasticity may at least partially explain how columnar metaplasia can develop in the mid-esophagus, far from the SCJ (Roliim et al., 2017) or even when the cardia has been surgically removed (Franchimont et al., 2003).

Dedifferentiation and cell reprogramming

Our results show that the HH pathway is activated in the epithelium of the embryonic esophagus and in a subpopulation of *Krt7*⁺ squamous cells at the SCJ under physiological conditions. Upon chronic gastro-esophageal reflux, the HH pathway is activated in a broad area of hyperplastic squamous epithelium. Sonic Hedgehog (Shh) expression has been reported to be stimulated by acidic pH in esophageal cells *in vitro* (Wang et al., 2010). Our data show that the sole activation of the HH pathway in the esophagus triggers transdifferentiation of a subset of keratinocytes *in vivo*. Expression of SmoM2 in some progenitors indeed inhibits squamous differentiation and stimulates expression of several classical columnar metaplasia markers. Our data show that the majority of SmoM2⁺ cells keep expressing squamous markers and express only few markers of metaplasia, such as *Krt20*. Interestingly, an intermediate phenotype called multilayered epithelium has also been reported within foci in Barrett's metaplasia, where cells co-express squamous markers (Boch et al., 1997). Such an intermediate phenotype suggests

Figure 6. Activation of the Hedgehog pathway induces opening of embryonic esophagus-specific chromatin regions to facilitate transcommitment

- (A) Scheme representing ATAC-seq.
 - (B) Venn diagram between open chromatin regions in normal adult esophageal cells (CTRL) and EpHI esophageal cells from K5:Smc mice.
 - (C) Venn diagram between the genes in which the chromatin is significantly more open and the genes upregulated in EpHI cells compared with CTRL.
 - (D) Illustration of open chromatin regions in *Krt7* and *Lgr5* loci from CTRL and EpHI cells.
 - (E) *De novo* motif enriched in the chromatin regions significantly more open in EpHI cells compared with CTRL.
 - (F) Illustration of open chromatin regions in CTRL, EpHI, and embryonic esophageal cells at E13.5.
 - (G) Venn diagram between the chromatin regions significantly more open in EpHI cells from K5:Smc compared with CTRL and the chromatin regions significantly more open in embryonic esophageal cells compared with CTRL (\log_2 fold change [LFC] > 1; FDR < 0.05).
 - (H) *De novo* motif enriched in the 10,375 common peaks between EpHI cells and embryonic esophageal cells illustrated in (G).
 - (I) Expression of the most frequent Sox transcription factors in EpHI cells measured using RNA-seq.
 - (J) Expression of four Sox transcription factors measured using RNA-seq in embryonic esophageal cells, columnar transitional, gastric, and intestinal epithelial cells.
 - (K) Co-immunostaining for *Krt14*, EpCam, and Sox9 in CTRL and K5:Smc esophagus.
- Scale bars, 20 μ m. lum, lumen; LP, lamina propria. All data from K5:Smc esophagus were measured 12 weeks after TAM administration.



(legend on next page)

that most of the keratinocytes are unable to turn off squamous differentiation, which is orchestrated by p63 (Wang et al., 2011). Our data thus suggest that Krt5+ basal esophageal progenitors are functionally heterogeneous, as only a fraction of them are able to achieve a complete squamous-to-columnar conversion. Along the same line, there is compelling evidence that esophageal basal progenitors are heterogeneous (DeWard et al., 2014), although it is probably not due to a cellular hierarchy (Doupe et al., 2012). Tackling the question of this functional heterogeneity will require the development of new mouse models to characterize esophageal cells subpopulations and their competence to transdifferentiate. This will be important to determine whether resistance to reprogramming is an intrinsic mechanism, whether it requires a specific microenvironment, or both.

Cell plasticity, dedifferentiation, and transdifferentiation *in vivo* have been a matter of intense debate for the past few years (Gupta et al., 2019; Merrell and Stanger, 2016). Our work provides an illustration of a physiopathological condition in which a pathway induces dedifferentiation to then allow the priming of another differentiation program. It has been reported that columnar-lined esophagus develops through wound repair in a surgical model of reflux esophagitis (Agoston et al., 2018). This process is associated with the upregulation of transcription factors involved in the epithelial-to-mesenchymal transition (EMT). The EMT is a well-known process involved in stemness and cell plasticity (Gupta et al., 2019). Our data show that key EMT-related transcription regulators, such as Twist1 and Zeb2, are upregulated in the embryonic esophagus and upon activation of the HH pathway in adult esophagus. Hence, it will be important to determine whether some of these transcription factors are involved in the process of dedifferentiation and/or transdifferentiation.

Our data show similarities at transcriptional and epigenetic levels between esophageal cells in which the HH pathway is activated and embryonic cells. A process of dedifferentiation has been reported in case of injury and regeneration in other tissues, such as heart in zebrafish (Jopling et al., 2010; Kikuchi et al., 2010), limbs in urodele amphibians (Kragl et al., 2009), Schwann cells in mammalian nerves (Mirsky et al., 2008; Woodhoo et al., 2009), and intestine in mouse (Nusse et al., 2018; Yui et al., 2018). Along the same line, activation of the HH pathway in epidermal keratinocytes induces reprogramming into embryonic hair follicle placode (Youssef et al., 2012). More-

over, BE shares similarities with the transition epithelium at the SCJ and embryonic esophageal cells (Wang et al., 2011). Interestingly, an initial step of dedifferentiation into embryonic-like progenitors could be the explanation for the three types of columnar metaplasia found in the esophagus: cardiac, oxyntic, and intestinal. As metaplasia is considered an adaptation to chronic stress, cues from the environment, such as inflammation or the content of the gastro-esophageal reflux, could stimulate differentiation of embryonic-like cells into one of these fates.

Synergy of signaling pathways to drive transdifferentiation

A previous study already reported that overexpression of Gli1, which is a major transcription factor downstream of the HH pathway, in esophageal cells for 2 weeks impairs squamous differentiation (van Dop et al., 2013). In addition, Shh signaling has also been reported to induce expression of columnar markers (Krt8 and Sox9) in a three-dimensional esophageal tissue reconstitution model (Wang et al., 2010). Our data show that SmoM2 expression in the esophagus *in vivo* leads to the activation of the HH pathway and subsequently to other pathways, such as Wnt and BMP pathways, that may synergize to modulate cell fate. Interestingly, activation of the b-catenin pathway in esophageal cells has been reported to promote expression of CyclinD1, Sox9, and Krt8 *in vivo* (Moyes et al., 2012). Our data also highlight an upregulation of *Bmp6* and *Bmpr1b* in esophageal cells following activation of the HH pathway. Knockout of the BMP inhibitor Noggin during development leads to the formation of mucin-producing columnar cells in the esophagus. In addition, overexpression of a constitutive form of the BMP receptor BMPR1a in the embryonic esophageal epithelium inhibits squamous differentiation (Rodriguez et al., 2010). Reactivation of this pathway in esophageal cells might therefore participate in the HH-induced dedifferentiation process and the acquisition of columnar features *in vivo*. Consistent with this notion, GSEA of *de novo* Krt8+ cell RNA-seq data highlights significant enrichment for genes related to “Smad binding” and “BMP regulation.” Furthermore, Smad motifs are found upstream of Sox9. However, it is unclear whether all the esophageal progenitors are equally sensitive to the HH, Wnt, or BMP pathways and what is the impact of each of these pathways on cell fate determination *in vivo*.

Figure 7. Sox9 expression is required for the squamous-to-columnar conversion of esophageal progenitors *in vivo*

- (A) Genetic strategy.
(B) Immunostaining for Sox9 in CTRL, K5:Smo, and K5:Smo:Sox9cKO esophagus.
(C) Co-immunostaining for SmoM2, Krt8, and Krt20 in K5:Smo and K5:Smo:Sox9cKO esophagus.
(D) Volcano plots representing results of RNA-seq in K5:Smo:Sox9cKO compared with K5:Smo esophagus.
(E) Venn diagrams between the transcripts that are upregulated in K5:Smo and downregulated in K5:Smo:Sox9cKO.
(F) Expression of columnar differentiation-related genes in K5:Smo:Sox9cKO compared with K5:Smo.
(G) MDS plot representing mRNA sequencing data from CTRL adult esophageal cells, embryonic esophageal cells, EpHI/K8neg, and paired EpHI/K8+ from K5:Smo:K8-YFP mice, as well as EpHI from K5:Smo:Sox9cKO esophagus.
(H) Radar plot showing the number of transcripts from the “embryo only,” “junction only,” “stomach only,” “intestine only,” or “columnar” signatures upregulated in EpHI from K5:Smo:Sox9cKO compared with control adult esophagus. p value was calculated using Fisher’s exact test.
(I) Illustrations of chromatin regions in normal adult esophageal cells (CTRL) and EpHI cells from K5:Smo and K5:Smo:Sox9cKO.
(J) Venn diagram between the chromatin regions significantly more open in EpHI cells from K5:Smo, in embryonic esophageal cells and in EpHI cells from K5:Smo:Sox9cKO compared with control esophagus (LFC > 1, FDR < 0.05).
(K) Scheme depicting our model.
- Scale bars, 20 μ m. lum, lumen; LP, lamina propria. All data were measured 12 weeks after TAM administration.

Intestinalization of columnar metaplasia

Our data confirm that the vast majority of esophageal cells are unable to undergo a process of squamous-to-columnar conversion *in vivo*, as previously suggested by the lack of transdifferentiation in K14-Cdx2 and p63CreER:Rosa-rtTA:TetO-Cdx2 models (Jiang et al., 2017; Kong et al., 2011b). In our model, only a subset of progenitor can undergo squamous-to-columnar conversion. These columnar cells express Sox9 but not Cdx2. Moreover, Sox9 is required for squamous-to-columnar conversion and the expression of GI tract differentiation markers. Sox9 expression is stimulated by the activation of the HH pathway in a BMP4-dependent manner in esophageal keratinocytes (Wang et al., 2010). In a model of tissue reconstruction *in vivo*, Sox9 overexpression in esophageal progenitors is sufficient to induce K8 and A33 columnar/intestinal marker expression (Clemons et al., 2012). In the same study, the authors reported that Cdx2 overexpression alone fails to induce columnar differentiation in esophageal cells and does not promote intestinalization of Sox9-induced columnar epithelium when grown on rat devitalized trachea. Here, we show that the activation of the HH pathway induces chromatin remodeling that uncovers enhancers containing putative binding motifs for Sox9 and Foxa2. Sox9 deletion prevents the upregulation of Foxa2 in epithelial cells, showing that Foxa2 is downstream of Sox9 in our model. Foxa2 has been reported to be a HH pathway target gene, which is upregulated in human BE samples and regulates Agr2 and Muc2 expression (Wang et al., 2014). As Sox9 protein upregulation in esophageal progenitors does not require a modification of its chromatin landscape or a strong increase in its mRNA synthesis, this is compatible with an early appearance of Sox9 along the transdifferentiation process. Sox9 may therefore act as a pioneer factor that appears quickly following activation of the HH pathway to start modifying the chromatin landscape and turn on a specific transcriptional program. Consistent with this notion, activation of the HH pathway in esophageal cells leads to the opening of thousands of chromatin regions. Almost a third of these regions are no longer significantly more opened in the absence of Sox9. Still, chromatin immunoprecipitation would be required to demonstrate whether Sox9 binds specific chromatin regions and directly regulates gene expression. The HH pathway on its own does not lead to the appearance of specialized intestinal cells such as goblet cells. As Cdx2 is the master regulator of goblet cell differentiation, it might be the switch that induces a complete intestinal differentiation program in the *de novo* Krt8+ columnar cells in the esophagus. It will be important to determine what may induce Cdx2 expression and what would be the consequences of its expression in transcommitted cells.

In conclusion, our work demonstrates that some esophageal progenitors can undergo a full squamous-to-columnar conversion *in vivo*. These results provide mechanistic insights into the understanding of the multistep process involved in squamous-to-columnar conversion and the mechanisms by which keratinocyte may participate to columnar metaplasia.

Limitations of study

This work was performed in transgenic mice to investigate the plasticity of esophageal cells *in vivo*. The specific conditions of chronic acid reflux are not recapitulated in this model

(low pH, bile acids, inflammation). The priming of intestinal transcripts described in our study does not lead to the development of intestinal metaplasia. Finally, although the markers we used, such as EpCam, Krt8, and Sox9, are expressed in human columnar metaplasia of the esophagus, the relevance of our observations for human pathology should be tested.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Histology and immunostaining
 - *In situ* hybridization
 - Antibodies
 - Clonal analysis
 - Image acquisition
 - Tissue digestion
 - FACS isolation
 - RNA extraction and quantitative real-time PCR
 - RNA-seq and analysis of bulk samples
 - GSEA analysis
 - ATAC-sequencing and analysis
 - Human data analysis
 - Single cell-RNA sequencing and analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.stem.2021.03.019>.

ACKNOWLEDGMENTS

We thank the FACS core facility and the animal house facility of Université libre de Bruxelles (ULB) for their help. Sequencing was performed at the Brussels Interuniversity Genomics High Throughput core (www.brightcore.be). The Krt8-YFP knockin mice were kindly provided by Prof. Rudolf Leube. B.B. is an investigator of Walloon Excellence in Lifesciences & BIOTEchnology (WEL-BIO) and Fonds de la Recherche Scientifique (FNRS) at ULB. A.V.D. is supported by a fellowship from the TELEVIE and Fondation contre le Cancer. B.D. is supported by Fonds pour la formation à la recherche dans l'industrie et dans l'agriculture (FRIA). This work was supported by Worldwide Cancer Research (1611-92), the region Wallonie-Bruxelles (ARC consolidator), WEL-BIO (FRFS-WELBIO-CR-2017S-01), FNRS (F4538.16 MIS), and Fondation contre le Cancer (FCC/ULB 2018-067).

AUTHOR CONTRIBUTIONS

A.V.D. and B.B. designed the experiments and performed data analysis. A.V.D. performed all the experiments. S.K. participated in the analysis of scRNA-seq data. B.D. participated in GSEA analysis. S.P. provided technical support. A.L. and F.L. are members of the genomic facility. S.N. provided esophago-gastro-jejunostomy (EGJ) samples. A.V.D. and B.B. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 18, 2020

Revised: December 18, 2020

Accepted: March 24, 2021

Published: April 20, 2021

REFERENCES

- Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* *44* (W1), W3–W10.
- Agoston, A.T., Pham, T.H., Odze, R.D., Wang, D.H., Das, K.M., Spechler, S.J., and Souza, R.F. (2018). Columnar-lined esophagus develops via wound repair in a surgical model of reflux esophagitis. *Cell. Mol. Gastroenterol. Hepatol.* *6*, 389–404.
- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* *31*, 166–169.
- Blighe, K., Rana, S., and Lewis, M. (2020). EnhancedVolcano: publication-ready volcano plots with enhanced colouring and labeling. R package version 1.8.0. <https://github.com/kevinblighe/EnhancedVolcano>.
- Boch, J.A., Shields, H.M., Antonioli, D.A., Zwas, F., Sawhney, R.A., and Trier, J.S. (1997). Distribution of cytokeratin markers in Barrett's specialized columnar epithelium. *Gastroenterology* *112*, 760–765.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* *30*, 2114–2120, <https://doi.org/10.1093/bioinformatics/btu170>.
- Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* *10*, 1213–1218.
- Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., van Oudenaarden, A., and Jaenisch, R. (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* *150*, 1209–1222.
- Cao, E.Y., Ouyang, J.F., and Rackham, O.J.L. (2020). GeneSwitches: ordering gene expression and functional events in single-cell experiments. *Bioinformatics* *36*, 3273–3275.
- Chen, X., Ding, Y.W., Yang, G.y., Bondoc, F., Lee, M.J., and Yang, C.S. (2000). Oxidative damage in an esophageal adenocarcinoma model with rats. *Carcinogenesis* *21*, 257–263.
- Clemons, N.J., Wang, D.H., Croagh, D., Tikoo, A., Fennell, C.M., Murone, C., Scott, A.M., Watkins, D.N., and Phillips, W.A. (2012). Sox9 drives columnar differentiation of esophageal squamous epithelium: a possible role in the pathogenesis of Barrett's esophagus. *Am. J. Physiol. Gastrointest. Liver Physiol.* *303*, G1335–G1346.
- DeWard, A.D., Cramer, J., and Lagasse, E. (2014). Cellular Heterogeneity in the Mouse Esophagus Implicates the Presence of a Nonquiescent Epithelial Stem Cell Population. *Cell Rep.* *9*, 701–711, <https://doi.org/10.1016/j.celrep.2014.09.027>.
- di Pietro, M., Lao-Sirieix, P., Boyle, S., Cassidy, A., Castillo, D., Saadi, A., Eskeland, R., and Fitzgerald, R.C. (2012). Evidence for a functional role of epigenetically regulated midcluster HOXB genes in the development of Barrett esophagus. *Proc. Natl. Acad. Sci. U S A* *109*, 9077–9082.
- Doupe, D.P., Alcolea, M.P., Roshan, A., Zhang, G., Klein, A.M., Simons, B.D., and Jones, P.H. (2012). A Single Progenitor Population Switches Behavior to Maintain and Repair Esophageal Epithelium. *Science* *337*, 1091–1093, <https://doi.org/10.1126/science.1218835>.
- Feng, C., Luo, Y., Nian, Y., Liu, D., Yin, X., Wu, J., Di, J., Zhang, R., and Zhang, J. (2017). Diallyl disulfide suppresses the inflammation and apoptosis resistance induced by DCA through ROS and the NF- κ B signaling pathway in human Barrett's epithelial cells. *Inflammation* *40*, 818–831.
- Franchimont, D., Covas, A., Brasseur, C., Laethem, J.L., El-Nakadi, I., and Devière, J. (2003). Newly developed Barrett's esophagus after subtotal esophagectomy. *Endoscopy* *35*, 850–853.
- Giroux, V., and Rustgi, A.K. (2017). Metaplasia: tissue injury adaptation and a precursor to the dysplasia-cancer sequence. *Nat. Rev. Cancer* *17*, 594–604.
- Gupta, P.B., Pastushenko, I., Skibinski, A., Blanpain, C., and Kuperwasser, C. (2019). Phenotypic plasticity: driver of cancer initiation, progression, and therapy resistance. *Cell Stem Cell* *24*, 65–78.
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* *38*, 576–589.
- Huang, W., Loganathanaraj, R., Schroeder, B., Fargo, D., and Li, L. (2013). PAVIS: a tool for Peak annotation and visualization. *Bioinformatics* *29*, 3097–3099.
- Hyland, P.L., Hu, N., Rotunno, M., Su, H., Wang, C., Wang, L., Pfeiffer, R.M., Gherman, B., Giffen, C., Dykes, C., et al. (2014). Global changes in gene expression of Barrett's esophagus compared to normal squamous esophagus and gastric cardia tissues. *PLoS ONE* *9*, e93219.
- Ieda, M., Fu, J.-D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* *142*, 375–386.
- Jenkins, G.J.S., D'Souza, F.R., Suzen, S.H., Eltahir, Z.S., James, S.A., Parry, J.M., Griffiths, P.A., and Baxter, J.N. (2007). Deoxycholic acid at neutral and acid pH, is genotoxic to oesophageal cells through the induction of ROS: The potential role of anti-oxidants in Barrett's oesophagus. *Carcinogenesis* *28*, 136–142.
- Jiang, M., Li, H., Zhang, Y., Yang, Y., Lu, R., Liu, K., Lin, S., Lan, X., Wang, H., Wu, H., et al. (2017). Transitional basal cells at the squamous-columnar junction generate Barrett's oesophagus. *Nature* *550*, 529–533.
- Jopling, C., Sleep, E., Raya, M., Martí, M., Raya, A., and Izpisua Belmonte, J.C. (2010). Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* *464*, 606–609.
- Kaz, A.M., Wong, C.-J., Luo, Y., Virgin, J.B., Washington, M.K., Willis, J.E., Leidner, R.S., Chak, A., and Grady, W.M. (2011). DNA methylation profiling in Barrett's esophagus and esophageal adenocarcinoma reveals unique methylation signatures and molecular subclasses. *Epigenetics* *6*, 1403–1412.
- Kaz, A.M., Grady, W.M., Stachler, M.D., and Bass, A.J. (2015). Genetic and epigenetic alterations in Barrett's esophagus and esophageal adenocarcinoma. *Gastroenterol. Clin. North Am.* *44*, 473–489.
- Kikuchi, K., Holdway, J.E., Werdich, A.A., Anderson, R.M., Fang, Y., Egnaczyk, G.F., Evans, T., Macrae, C.A., Stainier, D.Y.R., and Poss, K.D. (2010). Primary contribution to zebrafish heart regeneration by gata4(+) cardiomyocytes. *Nature* *464*, 601–605.
- Kong, J., Crissey, M.A.S., Stairs, D.B., Sepulveda, A.R., and Lynch, J.P. (2011a). Cox2 and β -catenin/T-cell factor signaling intestinalize human esophageal keratinocytes when cultured under organotypic conditions. *Neoplasia* *13*, 792–805.
- Kong, J., Crissey, M.A., Funakoshi, S., Kreindler, J.L., and Lynch, J.P. (2011b). Ectopic Cdx2 expression in murine esophagus models an intermediate stage in the emergence of Barrett's esophagus. *PLoS ONE* *6*, e18280.
- Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2016). Fast gene set enrichment analysis. *bioRxiv*. <https://doi.org/10.1101/060012>.
- Kragl, M., Knapp, D., Nacu, E., Khattak, S., Maden, M., Epperlein, H.H., and Tanaka, E.M. (2009). Cells keep a memory of their tissue origin during axolotl limb regeneration. *Nature* *460*, 60–65.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* *9*, 357–359, <https://doi.org/10.1038/nmeth.1923>.

- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Litingtung, Y., Lei, L., Westphal, H., and Chiang, C. (1998). Sonic Hedgehog is essential to foregut development. *Nat. Genet.* 20, 58–61.
- Liu, Z., Wang, L., Welch, J.D., Ma, H., Zhou, Y., Vaseghi, H.R., Yu, S., Wall, J.B., Alimohamadi, S., Zheng, M., et al. (2017). Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte. *Nature* 551, 100–104.
- Merrell, A.J., and Stanger, B.Z. (2016). Adult cell plasticity in vivo: de-differentiation and transdifferentiation are back in style. *Nat. Rev. Mol. Cell Biol.* 17, 413–425.
- Milano, F., van Baal, J.W.P.M., Buttar, N.S., Rygiel, A.M., de Kort, F., DeMars, C.J., Rosmolen, W.D., Bergman, J.J.G.H.M., Van Marle, J., Wang, K.K., et al. (2007). Bone morphogenetic protein 4 expressed in esophagitis induces a columnar phenotype in esophageal squamous cells. *Gastroenterology* 132, 2412–2421.
- Mirsky, R., Woodhoo, A., Parkinson, D.B., Arthur-Farraj, P., Bhaskaran, A., and Jessen, K.R. (2008). Novel signals controlling embryonic Schwann cell development, myelination and dedifferentiation. *J. Peripher. Nerv. Syst.* 13, 122–135.
- Morikawa, M., Koinuma, D., Tsutsumi, S., Vasilaki, E., Kanki, Y., Heldin, C.-H., Aburatani, H., and Miyazono, K. (2011). ChIP-seq reveals cell type-specific binding patterns of BMP-specific Smads and a novel binding motif. *Nucleic Acids Res.* 39, 8712–8727.
- Motoyama, J., Liu, J., Mo, R., Ding, Q., Post, M., and Hui, C.C. (1998). Essential function of Gli2 and Gli3 in the formation of lung, trachea and oesophagus. *Nat. Genet.* 20, 54–57.
- Moyes, L.H., McEwan, H., Radulescu, S., Pawlikowski, J., Lamm, C.G., Nixon, C., Sansom, O.J., Goings, J.J., Fullarton, G.M., and Adams, P.D. (2012). Activation of Wnt signalling promotes development of dysplasia in Barrett's oesophagus. *J. Pathol.* 228, 99–112.
- Nusse, Y.M., Savage, A.K., Marangoni, P., Rosendahl-Huber, A.K.M., Landman, T.A., de Sauvage, F.J., Locksley, R.M., and Klein, O.D. (2018). Parasitic helminths induce fetal-like reversion in the intestinal stem cell niche. *Nature* 559, 109–113.
- Ostrowski, J., Mikula, M., Karczmarski, J., Rubel, T., Wyrwicz, L.S., Bragoszewski, P., Gaj, P., Dadlez, M., Butruk, E., and Regula, J. (2007). Molecular defense mechanisms of Barrett's metaplasia estimated by an integrative genomics. *J. Mol. Med. (Berl.)* 85, 733–743.
- Powell, D. (2019). *drpowell/degust* 4.1.1. <https://zenodo.org/record/3501067#.YGsmbOhKhhE>.
- Quante, M., Bhagat, G., Abrams, J.A., Marache, F., Good, P., Lee, M.D., Lee, Y., Friedman, R., Asfaha, S., Dubeykovskaya, Z., et al. (2012). Bile acid and inflammation activate gastric cardia stem cells in a mouse model of Barrett-like metaplasia. *Cancer Cell* 21, 36–51.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. In *Bioinformatics*, 26 (Doi), pp. 841–842, <https://doi.org/10.1093/bioinformatics/btq033>.
- Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 42, W187–W191.
- Reid, B.J., Li, X., Galipeau, P.C., and Vaughan, T.L. (2010). Barrett's oesophagus and oesophageal adenocarcinoma: time for a new synthesis. *Nat. Rev. Cancer* 10, 87–101.
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26.
- Rodríguez, P., Da Silva, S., Oxburgh, L., Wang, F., Hogan, B.L.M., and Que, J. (2010). BMP signaling in the development of the mouse esophagus and forestomach. *Development* 137, 4171–4176.
- Rolim, I., Rodrigues, R.V., Bettencourt, A., Barros, R., Camilo, V., Dias Pereira, A., Almeida, R., and Chaves, P. (2017). Mid-esophagus columnar metaplasia: what is the biopathogenic pathway? *Int. J. Surg. Pathol.* 25, 262–265.
- Schwarz, N., Windoffer, R., Magin, T.M., and Leube, R.E. (2015). Dissection of keratin network formation, turnover and reorganization in living murine embryos. *Sci. Rep.* 5, 9007.
- Smith, Z.D., Sindhu, C., and Meissner, A. (2016). Molecular features of cellular reprogramming and development. *Nat. Rev. Mol. Cell Biol.* 17, 139–154.
- Song, S., Guha, S., Liu, K., Buttar, N.S., and Bresalier, R.S. (2007). COX-2 induction by unconjugated bile acids involves reactive oxygen species-mediated signalling pathways in Barrett's oesophagus and oesophageal adenocarcinoma. *Gut* 56, 1512–1521.
- Souza, R.F. (2016). From reflux esophagitis to esophageal adenocarcinoma. *Dig. Dis.* 34, 483–490.
- Spechler, S.J. (2012). Barrett's esophagus: is the goblet half empty? *Clin. Gastroenterol. Hepatol.* 10, 1237–1238.
- Stairs, D.B., Nakagawa, H., Klein-Szanto, A., Mitchell, S.D., Silberg, D.G., Tobias, J.W., Lynch, J.P., and Rustgi, A.K. (2008). Cdx1 and c-Myc foster the initiation of transdifferentiation of the normal esophageal squamous epithelium toward Barrett's esophagus. *PLoS ONE* 3, e3534.
- Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* 19, 477.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. *Cell* 177, 1888–1902.e21.
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.
- Terabe, F., Aikou, S., Aida, J., Yamamichi, N., Kaminishi, M., Takubo, K., Seto, Y., and Nomura, S. (2017). Columnar metaplasia in three types of surgical mouse models of esophageal reflux. *Cell. Mol. Gastroenterol. Hepatol.* 4, 115–123.
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, L. (2014). Pseudo-temporal ordering of individual cells reveals dynamics and regulators of cell fate decisions. *Nat. Biotechnol.* 32, 381–386.
- Treutlein, B., Lee, Q.Y., Camp, J.G., Mall, M., Koh, W., Shariati, S.A.M., Sim, S., Neff, N.F., Skotheim, J.M., Wernig, M., and Quake, S.R. (2016). Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* 534, 391–395.
- Vallejos, C.A., Risso, D., Scialdone, A., Dudoit, S., and Marioni, J.C. (2017). Normalizing single-cell RNA sequencing data: challenges and opportunities. *Nat. Methods* 14, 565–571.
- van Dop, W.A., Rosekrans, S.L., Uhmman, A., Jaks, V., Offerhaus, G.J.A., van den Bergh Weerman, M.A., Kasper, M., Heijmans, J., Hardwick, J.C.H., Verspaget, H.W., et al. (2013). Hedgehog signalling stimulates precursor cell accumulation and impairs epithelial maturation in the murine oesophagus. *Gut* 62, 348–357.
- Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463, 1035–1041.
- Wang, D.H., Clemons, N.J., Miyashita, T., Dupuy, A.J., Zhang, W., Szczepny, A., Corcoran-Schwartz, I.M., Wilburn, D.L., Montgomery, E.A., Wang, J.S., et al. (2010). Aberrant epithelial-mesenchymal Hedgehog signaling characterizes Barrett's metaplasia. *Gastroenterology* 138, 1810–1822.
- Wang, X., Ouyang, H., Yamamoto, Y., Kumar, P.A., Wei, T.S., Dagher, R., Vincent, M., Lu, X., Bellizzi, A.M., Ho, K.Y., et al. (2011). Residual embryonic cells as precursors of a Barrett's-like metaplasia. *Cell* 145, 1023–1035.
- Wang, D.H., Tiwari, A., Kim, M.E., Clemons, N.J., Regmi, N.L., Hodges, W.A., Berman, D.M., Montgomery, E.A., Watkins, D.N., Zhang, X., et al. (2014).

Hedgehog signaling regulates FOXA2 in esophageal embryogenesis and Barrett's metaplasia. *J. Clin. Invest.* *124*, 3767–3780.

Woo, J., Miletich, I., Kim, B.-M., Sharpe, P.T., and Shivdasani, R.A. (2011). Barx1-mediated inhibition of Wnt signaling in the mouse thoracic foregut controls tracheo-esophageal septation and epithelial differentiation. *PLoS ONE* *6*, e22493.

Woodhoo, A., Alonso, M.B.D., Droggiti, A., Turmaine, M., D'Antonio, M., Parkinson, D.B., Wilton, D.K., Al-Shawi, R., Simons, P., Shen, J., et al. (2009). Notch controls embryonic Schwann cell differentiation, postnatal myelination and adult plasticity. *Nat. Neurosci.* *12*, 839–847.

Youssef, K.K., Lapouge, G., Bouvrée, K., Rorive, S., Brohée, S., Appelstein, O., Larsimont, J.-C., Sukumaran, V., Van de Sande, B., Pucci, D., et al. (2012). Adult interfollicular tumour-initiating cells are reprogrammed into an embryonic hair follicle progenitor-like fate during basal cell carcinoma initiation. *Nat. Cell Biol.* *14*, 1282–1294.

Yui, S., Azzolin, L., Maimets, M., Pedersen, M.T., Fordham, R.P., Hansen, S.L., Larsen, H.L., Guiu, J., Alves, M.R.P., Rundsten, C.F., et al. (2018). YAP/TAZ-dependent reprogramming of colonic epithelium links ECM remodeling to tissue regeneration. *Cell Stem Cell* *22*, 35–49.e7.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Chicken polyclonal anti-Krt14 1/10000	BioLegend	906004
Rabbit monoclonal anti-Cldn18 1/2000	abcam	ab203563
Rabbit polyclonal anti-EpCam 1/500	abcam	ab71916; RRID: AB_1603782
Rat monoclonal anti-EpCam 1/500	BioLegend	118 202; RRID:AB_1089027
Goat polyclonal anti-GFP 1/1000	abcam	ab6673; RRID AB_305643
Rabbit polyclonal anti-GFP 1/2000	abcam	ab6556; RRID:AB_305564
Chicken polyclonal anti-GFP 1/500	Invitrogen	A10262; RRID:AB_2534023
Rabbit monoclonal anti-Krt7 1/1000	abcam	ab181598
Rat monoclonal anti-Krt8 1/250	DSHB	TROMA-I; RRID:AB_531826
Rabbit polyclonal anti-Krt20 1/1000	LabNed	LN2020180
Rabbit monoclonal anti-p63 1/1000	abcam	ab124762; RRID:AB_10971840
Rabbit polyclonal anti-Sox9 1/10000	Merck	AB5535; RRID:AB_2239761
Rabbit polyclonal anti-Krt5 1/4000	LabNed	LN0316197
Alexa Fluor 488 donkey anti-chicken	Jackson ImmunoResearch	703-545-155; RRID:AB_2340375
Alexa Fluor 488 donkey anti-rabbit	Jackson ImmunoResearch	711-545-152; RRID:AB_2313584
Alexa Fluor 488 donkey anti-goat	Jackson ImmunoResearch	705-545-147; RRID:AB_2336933
Rhodamine Red X donkey anti-chicken	Jackson ImmunoResearch	703-295-155; RRID:AB_2340371
Rhodamine Red X donkey anti-rabbit	Jackson ImmunoResearch	711-295-152; RRID:AB_2340613 711-297-003; RRID:AB_2340615
Alexa Fluor 647 donkey anti-chicken	Jackson ImmunoResearch	703-605-155; RRID:AB_2340379
Alexa Fluor 647 donkey anti-rat	Jackson ImmunoResearch	712-605-153; RRID:AB_2340694
Rat anti-CD45 PE	BioLegend	103106; RRID:AB_312971
Rat anti-CD140a PE	BioLegend	135906; RRID:AB_1953269
Rat anti-CD31 PE	BioLegend	102508; RRID:AB_312915
Rat anti-CD326(EpCam) APC-Cy7	BioLegend	118218; RRID:AB_2098648
Chemicals, peptides, and recombinant proteins		
Bovine Serum Albumin	Capricorn scientific	BSA-IT Lot CP14-1028
Triton X-100	Sigma-Aldrich	T8787
Glycergel	Dako	C0563
1,4-Diazabicyclo[2,2,8]octane (Dabco)	Sigma-Aldrich	D27802
Hoechst 33258 (10mg/mL solution)	Molecular probes	Cat # H3569
Tamoxifen	A&E Scientific	10540-29-1
Sunflower	Sigma-Aldrich	S5007
Formaldehyde 4%	VWR	116 994 55
Collagenase I	A.G. Scientific	C-2823
O.C.T.	Tissue-Tek	Cat#4583
Horse Serum	Capricorn scientific	HOS-1b
Foetal Bovine Serum	GIBCO	Cat# 10270-106
AccuMax Cell Detachment Solution	Capricorn Scientific	ACM-1G
Y-27632 dihydrochloride (Rock inhibitor)	Sigma-Aldrich	Y0503
Trypsin solution 2.5%	A&E Scientific	TRY-2B10
Dispase I	Sigma-Aldrich	D4818
Igepal 0.1%	Sigma-Aldrich	I8896
TDE1 transposase	Illumina	15027865

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TDE buffer	Illumina	15027866
Critical commercial assays		
E.Z.N.A Total RNA Kit (2100 Bioanalyzer)	Omega BIO-TEK Agilent	SKU: R6834-01 G293BA)
MinElute purification kit	QIAGEN	Cat # 28004
Ovation SoLo RNA-Seq System	NuGEN	Part # 0501-32
NEBNext High-Fidelity 2x PCR Master Mix	New England BioLabs	M0541S
AmpureXP magnetic beads	Beckman Coulter	A63881
RNAScope H2O2 & Protease plus reagents	Advanced Cell Diagnostics	322 330
RNAScope Retrieval reagents	Advanced Cell Diagnostics	322 000
RNAScope Protease III & IV	Advanced Cell Diagnostics	322340
RNAScope Fluorescent multiplex detection reagents	Advanced Cell Diagnostics	320851
Deposited data		
SuperSeries composed of three SubSeries	This manuscript	GSE148876
Single cell mRNA sequencing	This manuscript	GSE148875
mRNA sequencing	This manuscript	GSE148874
ATAC sequencing	This manuscript	GSE148872
Experimental models: organisms/strains		
K5-CreER ^{T2} knock-in B6N.129S6(Cg)-Krt5 < tm1.1(cre/ERT2) Blh > /J	The Jackson Laboratory	Stock# 029155; RRID:IMSR_JAX:029155
R26SmoM2 Gt(ROSA)26Sor < tm1(Smo/EYFP) Amc > /J	The Jackson Laboratory	Stock# 005130; RRID:IMSR_JAX:005130
Krt8-YFP knock-in	Rudolf E. Leube lab	N/A
Sox9 ^{fl} B6.129S7-Sox9 ^{tm2Crm} /J	The Jackson Laboratory	Stock# 013106; RRID:IMSR_JAX:013106
Oligonucleotides		
Primers for qPCR see Table S1	This manuscript	N/A
Software and algorithms		
Zen Blue	Zeiss	N/A
FACSDiva	BD Biosciences	N/A
STAR version 2.5.3a	Alexander Dobin (2009-2019)	https://github.com/alexdobin/STAR ; PMID:23104886; https://doi.org/10.1093/bioinformatics/bts635
HTseq	Anders et al., 2015	https://htseq.readthedocs.io/en/release_0.11.1/install.html#install
Degust 4.1.1: <i>interactive RNA-seq analysis</i>	David R. Powell	https://degust.erc.monash.edu/
R version 3.6.3	GNU project	https://cran.r-project.org/
Bioconductor		https://www.bioconductor.org/install/
HOMER	Heinz et al., 2010	http://homer.ucsd.edu/homer/download.html ; RRID:SCR_010881
GraphPad Prism 8	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Trimmomatic	Bolger et al., 2014	http://www.usadellab.org/cms/?page=trimmomatic
Bowtie2 (version 2.2.6)	Langmead and Salzberg, 2012	https://sourceforge.net/projects/bowtie-bio/files/bowtie2/2.4.1/
Picard tools		http://broadinstitute.github.io/picard/
Macs2 (version 2.1.1.20160309)		
Galaxy server	Afgan et al., 2016	https://usegalaxy.org/

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BED tool	Quinlan and Hall, 2010	https://bedtools.readthedocs.io/en/latest/content/installation.html
PAVIS2	Huang et al., 2013	https://manticore.niehs.nih.gov/pavis2/
Integrative Genomics Viewer (IGV)	Robinson et al., 2011	http://software.broadinstitute.org/software/igv/
Cell Ranger (version 3.1.0)		https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/installation

Other

Mouse reference genome GRCm38/mm10	Genome Reference Consortium	https://www.ncbi.nlm.nih.gov/grc/mouse
C5 collection MSigDB for mouse	Walter+Eliza Hall Bioinformatic Resources	http://bioinf.wehi.edu.au/software/MSigDB/
Human microarray data on Barrett samples and normal squamous samples	GEO: Ostrowski et al.	GSE36223
Human microarray data on Barrett samples and normal squamous samples	GEO: di Pietro et al.	GSE34619
Human microarray data on Barrett samples and normal squamous samples	GEO: Hyland et al.	GSE39491
Human microarray data on Barrett samples and normal squamous samples	GEO: Stairs et al.	GSE13083

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Benjamin Beck (benjamin.beck@ulb.be).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The datasets generated during this study are available under GEO accession SuperSeries GSE148876 composed of the following SubSeries GSE148872, GSE148875, GSE148875.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The Krt5-CreER^{T2} knock-in mice (The Jackson Laboratory, Stock#029155, RRID:IMSR_JAX:029155) were combined with the R26SmoM2 mice (The Jackson Laboratory, Stock#005130, RRID:IMSR_JAX:005130) in order to generate K5-CreER^{T2}:R26SmoM2, a tamoxifen-inducible mouse model. The Krt8-YFP knock-in mice (Schwarz et al., 2015) were crossed with this model to generate the K5-CreER^{T2}:R26SmoM2:Krt8-YFP. The Sox9^{fl^{ox}} (The Jackson Laboratory, Stock#013106, RRID:IMSR_JAX:013106) mice were crossed with the K5-CreER^{T2}:R26SmoM2 in order to generate K5-CreER^{T2}:R26SmoM2:Sox9^{fl^{ox}/fl^{ox}} mice. All mice used in this study were composed of males and females with mixed genetic background. Tamoxifen (TAM) was diluted at 25 mg/mL in sunflower oil (Sigma). 2.5 mg TAM was administered intraperitoneally (IP) to K5-CreER^{T2}:R26SmoM2 mice and K5-CreER^{T2}:R26SmoM2:Krt8-YFP at P28 and followed over up to 15 weeks. K5-CreER^{T2}:R26SmoM2:Sox9^{fl^{ox}/fl^{ox}} mice and their control (K5-CreER^{T2}:R26SmoM2) were injected with 10 mg IP. Littermates of the same sex were randomly assigned to experimental groups. Mouse colonies were maintained in a certified animal facility in accordance with the European guidelines. All the experiments were approved by the ethical committee from the university and conform with regulatory standards (LA1230406 – project 666N).

METHOD DETAILS

For all experiments presented in this study, sample size was large enough to measure the effect size. No randomization and no blinding were performed in this study.

Histology and immunostaining

For the staining on frozen sections, tissues were harvested, directly embedded in O.C.T. (Tissue Tek) and flash frozen for cryopreservation. For the following staining: EpCam (Rt), Krt7, Sox9, Krt8-YFP, tissues were pre-fixed in 4% formaldehyde during 2 h at RT, washed in PBS, incubated overnight in 30% sucrose in PBS at 4°C, embedded in O.C.T. (Tissue Tek) and flash frozen for cryopreservation. Samples were sectioned at 6 µm sections using a M1860 cryostat (Leica Microsystems GmbH). Nonspecific antibody binding was blocked with 5% horse serum (HS), 1% Bovine Serum Albumin (BSA), and 0.2% Triton X-100 during 1 h at room temperature. Primary antibodies were incubated overnight at 4°C in blocking buffer. Sections were rinsed three times in PBS and incubated with secondary antibodies during 1 h at room temperature. Nuclei were stained with Hoechst (4 mM). Slides were mounted using Glycergel (Dako) supplemented with 2.5% DABCO (Sigma-Aldrich).

In situ hybridization

In situ hybridization has been realized on OCT embedded, flash frozen samples using RNAscope kits (RNAscope Protease III & IV and RNAscope Fluorescent multiplex detection reagents, Advanced Cell Diagnostics) following manufacturer's recommendations. The following probes were used: Gli1 (cat 311001) and Ptch1 (cat 402811).

Paraffin slides of esophagojejunostomy have been kindly provided by the lab of Sachiyo Nomura (Terabe et al., 2017) and *in situ* hybridization has been realized in the laboratory using RNAscope kits (RNAscope H2O2 & Protease plus reagents, RNAscope Retrieval reagents and RNAscope Fluorescent multiplex detection reagents). The following probes were used: Gli1 (cat 311001) and Shh (cat 314361).

Antibodies

The following primary antibodies were used: anti-Krt14 (polyclonal chicken, 1:10000, Biolegend), anti-GFP (polyclonal rabbit, 1:1000, Abcam, RRID:AB_305564), anti-Krt5 (polyclonal rabbit, 1:4000, LabNed), anti-GFP (polyclonal goat, 1:1000, Abcam, RRID:AB_305643), anti-GFP (polyclonal chicken, 1:500, Invitrogen, RRID:AB_2534023) anti-p63 (rabbit monoclonal, 1:1000, Abcam, RRID:AB_10971840), anti-EpCam (rabbit polyclonal, 1:500, Abcam, RRID:AB_1603782), anti-EpCam (rat polyclonal, 1:500, Biolegend, RRID:AB_1089027), anti-Krt7 (rabbit monoclonal, 1:1000, Abcam), anti-Krt20 (rabbit polyclonal, 1:1000, LabNed), anti-Krt8 (rat monoclonal, 1:250, DSHB, RRID:AB_531826), anti-Sox9 (rabbit polyclonal, 1:10000, Merck, RRID:AB_2239761), anti-Cldn18 (rabbit monoclonal, 1:2000, abcam).

The following secondary antibodies were used: anti-rabbit, anti-rat, anti-chicken, conjugated to AlexaFluor488 (1:500, Jackson ImmunoResearch), to rhodamine Red-X (1:500, Jackson ImmunoResearch) or to Cy5 (1:1000, Jackson ImmunoResearch).

Clonal analysis

For clonal analysis, co-immunostaining for YFP and Krt14 or Krt8 were used. A minimum of 42 clones from 8 different animals for each time point were counted (3, 8 and 12 weeks following TAM administration). Baseline has been measured in 3 control animals. Krt14 downregulation was assessed by comparison to the cells surrounding the clones. Clones with at least 1 positive Krt8 cell were considered as positive.

Image acquisition

Imaging was performed on a Zeiss Axio Imager M2 fluorescence microscope with a Zeiss Axiocam 503 mono camera and a Zeiss macroscope axiozoom V16 with axiocam MRN camera for immunofluorescence microscopy using Zen Blue (Zeiss) software. Brightness, contrast, and picture size were adjusted using Zen Blue (Zeiss). Anastomosis was imaged using a NanoZoomer-SQ Digital slide scanner (Hamamatsu).

Tissue digestion

Esophagi and squamous part of the squamocolumnar junction were dissected, minced and digested in collagenase I 2 mg/mL (A&E scientific) during 1 h 30. Collagenase I activity was blocked by the addition of EDTA (5 mM), incubated for 30 min. Trypsin (0.125%) was then added for 15 min. All incubations have been done on a rocking plate at 37°C.

Embryonic esophagi were dissected and digested in collagenase I 2mg/mL (A&E scientific) during 45 min. Collagenase I activity was blocked by the addition of EDTA (5 mM), incubated for 5 min. Trypsin (0.125%) was then added for 5 min and then the cells were rinsed in PBS supplemented with 2% FBS. All incubations have been done on a rocking plate at 37°C.

Columnar part of the squamocolumnar junctions were dissected, minced and washed 2 times in PBS supplemented with 10 mM HEPES by centrifugation 5 min at 50 × g. Tissues were then digested for 1 h 30 in Accumax (Capricorn Scientific) supplemented with 20 mM HEPES, 10 nM ROCK inhibitor (Y-27632 dihydrochloride, Sigma-Aldrich), 0.2% BSA at room temperature on a rocking plate. To improve digestion, EDTA (5mM) was added for 15 min at 37°C on a rocking plate.

The corpus was separated from the rest of the stomach, and the muscular layer was eliminated with micro scissors. Tissues were then minced and washed 2 times in PBS by centrifugation 2 min at 100 × g at 4°C. Tissues were then digested in PBS supplemented with EDTA 10mM, HEPES 10mM during 1h at 37°C on a rocking plate. Cells were then washed in PBS twice and filtered through a 70 µm cell strainers (BD). The glands were then digested by adding PBS supplemented with Trypsin (0.125%) and HEPES 10 mM for 15 min at 37°C on a rocking plate.

Guts were dissected and digested 1h30 in HBSS supplemented with 0.6 unit (30 $\mu\text{g}/\text{mL}$) dispase I (sigma D4818). Guts were then minced and digested in collagenase I (A&E scientific) 2 mg/mL for 15 min. Collagenase I activity was blocked by the addition of EDTA (5 mM) for 15 min. Finally, trypsin (0.125%) was added for 5 min. Each incubation was done on a rocking plate at 37°C.

For all the tissues, cells were then rinsed in PBS supplemented with 2% FBS and filtered through a 70 μm cell strainers (BD) to ensure the elimination of cell debris and clumps of cells.

FACS isolation

Immunostaining was performed on single cell suspension using PE-conjugated anti-CD45 (1:500, BioLegend), PE-conjugated anti-CD31 (1:500, BioLegend), PE-conjugated anti-CD140a (1:500, BioLegend) and APC-Cy7-conjugated anti-EpCam (clone G8.8, 1:250, Biolegend), during 45 min at 4°C on a rocking plate. Living epithelial cells were selected by forward scatter, side scatter, doublets discrimination and by Hoechst dye exclusion. EpCam+/Lin- cells were selected based on the expression of EpCam and the exclusion of CD45, CD31, CD140a (Lin-). Fluorescence-activated cell sorting analysis was performed using FACSaria III and FACS-Diva software (BD Biosciences).

RNA extraction and quantitative real-time PCR

Sorted cells were collected into TRK lysis buffer (Omega bio-tek) and RNA was extracted using E.Z.N.A Total RNA Kit (Omega bio-tek) according to the manufacturer's recommendations with DNase I digestion protocol on column (Omega bio-tek). After nanodrop RNA quantification, the first strand cDNA was synthesized, using Superscript II (Invitrogen) and random hexamers (Roche) in 50 μl final volume. Control of genomic contaminations was measured for each sample by performing the same procedure with or without reverse transcriptase. Quantitative PCR assays were performed using 1 ng of cDNA as template, PowerUP SYBRGreen master mix (Life Technologies Limited) and a Quantstudio 3 real-time PCR system (Applied Biosystems). Actin beta housekeeping gene was used for normalization. Primers were designed using the NCBI primer designing tool – Primer-blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) and are presented in Table S1. Quantitative PCR Analysis was performed using QuantStudio 3 software and the DDCT method with Actin beta as a reference. The entire procedure was repeated in at least three biologically independent samples and always with technical replicates.

RNA-seq and analysis of bulk samples

RNA quality was checked using a Bioanalyzer 2100 (Agilent technologies). Indexed cDNA libraries were obtained using the Ovation Solo RNA-Seq System (NuGen) following manufacturer's recommendations. The multiplexed libraries were loaded on a NovaSeq 6000 (Illumina) using a S2 flow cell and sequences were produced using a 200 Cycle Kit. Paired-end reads were mapped against the mouse reference genome GRCm38 using STAR software (version 2.5.3a) to generate read alignments for each sample. Annotations Mus_musculus.GRCm38.90.gtf were obtained from ftp.Ensembl.org. After transcripts assembling, gene level counts were obtained using HTSeq (Anders et al., 2015). Total raw counts were loaded on degust 4.1.1 (Powell, 2019). All analyses were performed using EdgeR, TMM normalization and “Min gene read count” set at 10. Control condition (CTRL) is defined by 6 control FACS sorted adult esophagus epithelium samples, these 6 biological samples were used as reference all along the paper to calculate the fold change of gene expressions. Through all the analysis we tried to keep the same order of magnitude for the number of differentially regulated genes in the different tissues. Therefore we defined 3 different thresholds: $\text{abs(LFC)} > 2$ and $\text{FDR} < 0.05$ for K5:SmO 3 weeks ($n = 4$), 8 weeks ($n = 4$) and 12 weeks ($n = 6$) after TAM induction, for EpHI/YFP- from the transition epithelium of K8-YFP mice ($n = 3$), for Ep+/YFP+ cells from gastric and intestinal epithelium from K8-YFP mice ($n = 2$); $\text{abs(LFC)} > 2$ and $\text{FDR} < 0.01$ for EpHI/K8+ and EpHI/K8- from K5:SmO:K8-YFP mice ($n = 3$); $\text{abs(LFC)} > 2.5$ and $\text{FDR} < 1\text{e-}4$ for EpHI/YFP+ from the transition epithelium of K8-YFP mice ($n = 4$) and for Ep+/YFP+ cells from embryonic esophagus ($n = 2$ pools of 6 esophagus). For the K5:SmO:Sox9cKO EpHI ($n = 4$), cells were compared to K5:SmO EpHI cells ($n = 6$) and a threshold of $\text{abs(LFC)} > 0.5$ and $\text{FDR} < 0.05$ was chosen. All volcano plots represent results of RNA-seq as the statistical significance versus the magnitude of fold of change (FC) and were generated using the package “EnhancedVolcano” (Blighe et al., 2020) from Bioconductor in R version 3.6.3. Data are available under GEO accession number GSE148874. MDS plots have been generated using degust 4.1.1 (Powell, 2019). Heatmap was generated using “heatmap.2” function and represent values in logCPM scaled by row for the 500 most variable genes between CTRL and EpHI cells 12 weeks after TAM induction.

GSEA analysis

GSEA analysis was performed using preranked gene set enrichment analysis from the fgsea package (Korotkevich et al., 2016) in R version 3.6.3, with “nperm=1000” and “maxSize=500.” The values of LFC were used as the ranking metric. For this purpose, thresholds were enlarged to include more genes in the analysis and therefore increase the strength of the analysis. For K5:SmO 3, 8 and 12 weeks after TAM, ranked fold change values correspond to EpHI cells from K5:SmO over CTRL cells ($\text{abs(LFC)} > 1$, $\text{FDR} < 0.05$). For K5:SmO:K8-YFP analysis, ranked fold change values correspond to EpHI/K8+ and EpHI/K8neg over CTRL cells ($\text{abs(LFC)} > 1$, $\text{FDR} < 0.05$). Gene sets were generated by taking the most upregulated genes compared to CTRL, only present in each ensemble (see Figure 5F “Junction only,” “Stomach only,” “Embryo only,” “Intestine only”). Finally, to perform unbiased gene ontology analyzes, the C5 collection adapted for mouse which contains gene sets annotated by GO terms have been downloaded (<http://>

bioinf.wehi.edu.au/software/MSigDB/) and the function “fgseaMultilevel” has been used. To remove redundant pathways, the function “reduce_pathways” from “GeneSwitches” package (Cao et al., 2020) was used with rate fixed at 0.8. Pathways with adjusted p value < 0.05 were considered as significant.

ATAC-sequencing and analysis

Assay for transposase accessible chromatin (ATAC) followed by sequencing was performed as following: between 50,000 and 100,000 sorted cells were collected in 1 mL of PBS + 3% FBS at 4°C. Cells were centrifuged, then cell pellets were resuspended in 100 μ L of lysis buffer (Tris HCl 10 mM, NaCl 10 mM, MgCl₂ 3 mM, Igepal 0.1%) and centrifuged (500 \times g) for 25 min at 4°C with break set on 4. Supernatant was discarded and nuclei were resuspended in 50 μ L of reaction buffer (TDE1 transposase 2.5 μ L (Illumina), TDE buffer 25 μ L (Illumina)). The reaction was performed for 30 min at 37°C and then blocked by addition of 5 μ L of stop buffer (NaCl 900 mM, EDTA 300 mM). DNA was purified using the MinElute purification kit (QIAGEN) according to the manufacturer’s recommendations. For embryonic esophagi E13.5, 60 embryos have been dissected to sort 10,000 cells in duplicate by flow cytometry. To compensate the low number of cells harvested, the tagmentation reaction was carried out in 4 times less volume. The reaction was blocked by freezing and no DNA purification was done to avoid loss of material. DNA libraries were PCR amplified (NEB-Next High-Fidelity 2x PCR Master Mix, New England Biolabs), indexed using the primers described previously (Buenrostro et al., 2013), and double size selected from 150 to 1200 base pairs (bp) using the AmpureXP magnetic beads (Beckman) following the manufacturer’s recommendations. The multiplexed libraries were loaded on a NovaSeq 6000 (Illumina) using a S2 flow cell and paired-end sequences were produced using a 200 Cycle Kit.

Three samples of CTRL, K5:Smo, K5:Smo:Sox9cKO esophagi and 2 samples of fetal esophagi were sequenced. ATAC-seq paired-end reads of 50 bp were trimmed for adaptor sequences using Trimmomatic. ATAC-seq paired-end reads were then aligned to the mouse GRCm38 genome using Bowtie2 (version 2.2.6) using options “-X 2000-fr-very-sensitive-no-discordant-no-unal-no-mixed-non-deterministic..” Approximately 34 million reads paired-end were mapped to mouse genomic DNA in each condition. Mitochondrial reads and reads aligned to scaffolds and undefined chromosomes were excluded from downstream analysis. Duplicated reads were removed by Picard tools (<http://broadinstitute.github.io/picard/>). Read start sites were adjusted to represent the center of the transposon binding event as described in Li et al. (2009). Peak calling was performed on each individual sample using Macs2 (version 2.1.1.20160309) with parameters set on “callpeak -f BAMPE -g mm -q 0.05 -nomodel -call-summits -B -SPMR.” The data were loaded to the public Galaxy server at usegalaxy.org (Afgan et al., 2016). Coordinates from all ATAC-sequencing peaks were merged using BED tool (Quinlan and Hall, 2010) with default options. The merged coordinates were reattributed to each sample using HTseq count (Anders et al., 2015) with the options: “feature type= exon” and “mode= union.” Identification of differentially opened chromatin regions between K5:Smo samples and CTRL samples were performed using EdgeR on degust 4.1.1 (Powell, 2019). Peaks coordinates were annotated using PAVIS2 (Huang et al., 2013) with default options (upstream region length of Transcript Start Site (TSS): 5000 bp and downstream regions length of Transcription Termination Site (TTS): 1000 bp). Integrative Genomics Viewer (IGV) (Robinson et al., 2011) allowed the visualization of differentially opened chromatin regions.

The peaks obtained following peak calling and the peak coordinates considered as differentially regulated compared to control following degust analysis, were both analyzed. To identify *de novo* and known motif enrichment in the coordinates of K5:Smo EpHI cells and embryonic esophageal cells analyzed with degust, we used findMotifsGenome.pl from HOMER software (Heinz et al., 2010) with background set to Control cells and options set to “-size 200 -len 8.” To look for known motif enrichments in common and differential opened peaks in K5:Smo and CTRL we used the peak files obtained following peak calling and findMotifsGenome.pl from HOMER software with options “-size 200 and -len 8.” Files with common and differential opened peaks in K5:Smo and CTRL cells were generated using the function mergePeaks from HOMER software with option “-d given -prefix -venn” after we have fused the peaks coordinates by condition. The function annotatepeaks.pl from HOMER software was used to find and annotate the opened chromatin regions in K5:Smo EpHI cells where Sox9 motif was present with option “-size 500.” Smad1/5 motif was not in Homer database and was directly taken from a publication (Morikawa et al., 2011) and used in IGV software. Heatmaps were generated using bedTools (Ramírez et al., 2014). First, bigwig files were generated using “bamCoverage” with option “-normalizeUsing RPGC.” Then matrices were generated using “computeMatrix scale-regions -m 10000 -b 3000 -a 3000” and UCSC mouse ncbiRefSeq mm10 (<https://genome.ucsc.edu/cgi-bin/hgTables>) as regions to plot. Finally, plotHeatmap was used with default parameters. Data are available under GEO accession number GSE148872.

Human data analysis

Human gene expression datasets from normal and Barrett’s esophagus of 4 different studies available on the Database Gene Expression Omnibus (GEO) were downloaded and analyzed using GEO2R and R software. Dataset #1 (Database: GEO: GSE39491; n = 80; Hyland et al., 2014), Dataset #2 (Database: GEO: GSE13083; n = 14; Stairs et al., 2008), #3 (Database: GEO: GSE34619; n = 18 (di Pietro et al., 2012)) and #4 (Database: GEO: GSE36223; n = 46; Ostrowski et al., 2007).

Single cell-RNA sequencing and analysis

Single cell sequencing preparation was carried out as following: for K5:Smo epithelial cells analysis, 50,000 and 100,000 living cells were sorted from 4 mice and charged on two different chips. For K5:Smo EpHI cells analysis 34,000 and 100,000 EpHI cells were sorted also from 4 mice and charged on two different chips. Cells were collected in 1mL of PBS + 0.04% BSA at 4°C. Samples were loaded on a chromium chip targeting a recovery between 2500 and 7000 cells. Samples were then processed using the Chro-

mium Single Cell 3' Reagent kits v3 (10x Genomics) following manufacturer's recommendations. The multiplexed libraries were loaded on a NovaSeq 6000 (Illumina) using a S1 flow cell and paired-end sequences were produced using a 100 Cycle Kit (Read1 28 cycles, i7 Index 8 cycles, i5 Index 0 cycles, Read2 87 cycles). Cell Ranger version 3.1.0 pipeline with the RNA-seq aligner STAR were used to generate output files aligned on mm10 reference genome with EYFP custom reference added (Clontech (TaKaRa) https://www.snapgene.com/resources/plasmid-files/?set=fluorescent_protein_genes_and_plasmids&plasmid=EYFP). Package Seurat v3 from Bioconductor (Stuart et al., 2019) was used in R to perform all the analyzes. The analyzes followed recommendations from Satijalab (<https://satijalab.org>). The datasets were converted into Seurat objects. Potential empty or low-quality droplets and multiplets were removed. The data were normalized using "SCTransform" that outperforms traditional global scaling normalization methods derived from bulk-RNA-seq (Vallejos et al., 2017). To avoid the influence of mitochondrial reads in the analyses, we filtered them using "vars.to.regress=percent.mt" with "percent.mt" calculated using "PercentageFeaturesSet." The features for datasets were selected for downstream integration with "SelectIntegrationFeatures" set to 1500. To work with datasets in the same space, and minimize batch effect from the 2 different sequencings, anchors were defined with "FindIntegrationAnchors" prior to integration of the data with "IntegrateData." The data were then linear transformed prior to PCA (npcs set to 20) and UMAP analyzes (dims set to 1:20). Distances between cells were defined using "FindNeighbors" with dims set to 1:20. The cells were grouped using "FindClusters" with a resolution set to 0.45. Clusters enriched in mitochondrial and/or ribosomal genes have been removed to clean the analysis. For all living cells analysis, epithelial cells have been subset based on clustering to keep only clusters devoid of immune, fibroblastic and endothelial markers, prior to analysis and integration with EpHI cells. After integration of all the datasets, one cluster very far from all other clusters in the clustering space and very enriched in Vimentin was removed. To visualize data and proceed to differential expression analysis, the unnormalized RNA assay was normalized and scaled using "NormalizeData" and "ScaleData." Positive markers for each group comparing to remaining ones were then identified with Wilcoxon test using "FindMarkers" and "p-value < 0.05." Average expression for the 10 most upregulated markers were illustrated with "DoHeatmap" for each group. Violin plots were done with "VlnPlot" function. To perform pseudo-time analysis, Seurat objects were manually converted to be compatible with *Monocle3* and/or *Slingshot* (Street et al., 2018) with UMAP projection used for trajectory inference (Trapnell et al., 2014). Starting points were chosen based on biology. Slingshot allowed us to plot the pseudotime values for each lineage. Following *Monocle3* analysis, pseudo-times were subsetted using `choose_graph_segments()`, expression plots were generated using `plot_cells()`. Analysis of the genes switching on and off along the pseudo-time was performed using GeneSwitches package (Cao et al., 2020). Data are available under GEO accession number GSE148875.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical and graphical data analyses were performed using Prism 8 (Graphpad) and R softwares. All experiments shown were replicated at least twice. All data in histograms represent mean. Statistical significance was calculated by Mann-Whitney or Fisher exact test using the Graphpad Prism and R softwares, considering $p < 0.05$ as statistically significant. All tests are two-sided. One sided Fisher exact test was used to calculate the significance of genesets overlap, setting the total number of transcripts as 52,637 and the parameter "alternative" as greater. For human data, normality of samples distribution was evaluated using Shapiro-Wilk test, then the appropriated test was performed (t test for normal distribution, Mann-Whitney for non-normal distribution).