Article

Lgr5 marks stem/progenitor cells contributing to epithelial and muscle development in the mouse esophagus

Lana Kostic[1], Carly Leung[1], Katzrin Ahmad Murad[1], Snezhina Kancheva[1], Stefano Perna[2], Bernett Lee[2,3,4,5] & Nick Barker [1,6] ✉

The existence and function of Lgr5+ cells within the developing esophagus remains unknown. Here, we document multiple discrete Lgr5+ populations in the developing mouse esophagus, predominantly within nascent epithelial and external muscle layers. Lgr5 expression initially emerges in the developing proximal embryonic epithelium, but progressively extends distally and persists within the distal epithelium of neonates. Fate mapping and ablation analyses reveal a long-term contribution of epithelial Lgr5+ cells to esophageal organogenesis. Additionally, Lgr5-expressing cells are present in the developing external muscle layer, particularly during the development of the striated component. Fate mapping reveals a significant contribution of these embryonic Lgr5+ cells to the adult muscle layer. Embryonic Lgr5+ epithelial cells are also found to be important for regulating epithelial development, serving as a key source of Wnt6, among other ligands, to promote epithelial cell proliferation and formation of epithelial layers. These findings significantly enhance our understanding of esophageal development and shed light on the involvement of Lgr5+ stem/progenitor cells during organogenesis. Importantly, this study lays the foundation for investigating esophageal diseases related to the Lgr5+ stem/progenitor cell pool.

The esophagus is a part of the upper gastrointestinal (GI) tract with no known digestive, absorptive, or metabolic function, acting simply as a link for transporting food between the oral cavity and the stomach. Consistent with its basic role in the food intake process, it displays a relatively simple tissue architecture. The lining of the adult mouse esophagus comprises a uniform layer of squamous epithelial cells residing above a highly proliferative basal layer[1]. This epithelium is keratinized, providing an additional layer of protection against an abrasive diet. To facilitate swallowing and food transportation to the stomach, the epithelial layer is surrounded by a double-layered external muscle (*muscularis externa*). This external muscle comprises outer (longitudinal) and inner (circular) layers of smooth muscle cells, infiltrated by striated muscle cells during esophageal development[2,3]. Activation of radial muscle waves in this layer propels ingested food towards the stomach[4].

Esophageal development initiates with the separation of the respiratory buds from the foregut tube at embryonic day 9.5 (E9.5), with this division being completed by E11.5 in mice[5]. This process requires delicate coordination between both endodermal and mesenchymal tissues. The esophagus initially comprises a single-layered cuboidal

[1]Institute of Molecular and Cell Biology (IMCB), Agency for Science, Technology and Research (A*STAR)61 Biopolis Drive, Proteos, Singapore, Singapore. [2]Centre for Biomedical Informatics, Lee Kong Chian School of Medicine, Nanyang Technological University, 59 Nanyang Drive, Experimental Medicine, Singapore, Singapore. [3]Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A*STAR), 8A Biomedical Grove, Immunos, Singapore, Singapore. [4]Infectious Disease Labs (ID Labs), Agency for Science, Technology and Research (A*STAR), 8A Biomedical Grove, Immunos, Singapore, Singapore. [5]Khoo Teck Puat Hospital, 90 Yishun Central, Singapore, Singapore. [6]Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore. ✉e-mail: nicholas_barker@imcb.a-star.edu.sg

epithelium surrounded by a smooth muscle layer[5]. After separating from the airway, both tissues gain additional cell layers, and the epithelium undergoes a shift in marker expression as it transitions from a cytokeratin 8 (Krt8) labeled cuboidal layer to a cytokeratin 14 (Krt14) expressing squamous layer[6]. Although the existence of cellular heterogeneity within the developing mouse esophageal epithelium and muscle has been described[7,8], it is currently not clear whether any of the subsets represents a pool of dedicated stem/progenitor cells and what their role in the process of organ development and maintenance might be.

The intricate process of embryogenesis relies on tightly regulated complex interactions between different signaling cascades. Wnt signaling is a well-established player in this process, with key roles in the maintenance and repair of many tissues during embryogenesis and after birth. Lgr5 acts as a facultative component of the Wnt receptor complex and regulates the development and maintenance of a variety of tissues, particularly epithelia[9–12]. Several studies have identified Lgr5-expressing (Lgr5+) cells and Wnt-related genes as essential components for the co-ordinated development of GI tract structures[13–15]. However, knowledge about the distinct roles that putative esophageal epithelial and muscle stem/progenitor cell subpopulations have in this process is scarce. Though Lgr5+ cell pools have been extensively studied in the lower GI tract epithelia, Lgr5+ cell progenitors have not been reported in esophageal tissues during development/homoeostasis. Nevertheless, Lgr5 is robustly expressed in pathological conditions such as Barrett's Esophagus, esophageal squamous cell carcinoma, and esophageal adenocarcinoma[16–19].

Here, we employ high resolution expression profiling assays, combined with Lgr5 reporter/tracing/ablation mouse models, explant and in vitro organoid cultures to document the distribution and dynamics of Lgr5-expressing cells in the esophagus from early embryonic stages until birth, and to evaluate their roles during organ development. We show that multiple Lgr5+ cell populations reside within both the developing esophageal epithelium and external muscle layers. In the developing embryonic epithelium, Lgr5 is initially restricted to the proximal esophagus, with expression subsequently arising in the distal region as maturation of the cuboidal-squamous epithelial layers progresses. Lineage tracing analyses identify early embryonic epithelial Lgr5+ cells as a minor, fast-cycling pool that is rarely retained in the epithelium, while late embryonic epithelial Lgr5+ cells form long-lived clones within developing basal layers that persist into adulthood Lgr5 expression also arises in the embryonic external muscle, correlating with the appearance of striated muscle cells. Importantly we demonstrate that Lgr5+ cells residing in the external muscle layer function as a major pool of embryonic stem/progenitor cells contributing to the establishment and maintenance of external muscle layers in adult esophagus. scRNAseq and Spatial Transcriptomics analyses have revealed significant co-expression of Lgr5 and Wnt6 within the developing esophageal epithelium, indicating a potentially key role for nascent Lgr5+ epithelial cells as a primary source of this crucial ligand during esophageal development. Functional assays combining Lgr5+ cell-ablation with Wnt6 supplementation and Wnt6 knockout (KO) in epithelial organoids document Wnt6 as being indispensable for esophageal epithelial growth in vitro. In adults, a minor pool of Lgr5+ cells is present within the esophageal epithelium, predominantly within the basal layer of more distal regions, where it makes limited contributions to epithelial maintenance throughout life. In contrast, Lgr5 expression in the external muscle layer of adults was almost undetectable, with highly infrequent, very low-level expression visible under homoeostatic conditions.

Results

Lgr5 is expressed in multiple tissue compartments during esophageal development

To accurately define the lineages harboring Lgr5-expressing (Lgr5+) cells during esophageal organogenesis and the key signaling cascades

involved in this process, we performed single-cell RNA sequencing (scRNA-seq) analysis of mouse esophagi at two embryonic timepoints, E13.5 and E16.5, and one postnatal timepoint, P0 (Fig. 1a). We microdissected the embryonic/neonatal esophagi as previously described[20] and dissociated cells using the cold protease digestion method[21], enabling effective capture of in vivo gene expression patterns whilst minimizing expression artefacts. We identified 25 distinct clusters, which corresponded to 14 major cell types. This classification was based on widely recognized marker genes, extensively validated for each cell type. The markers used for validating the major esophageal cell types are detailed in Supplementary Fig. 1a. Lgr5+ cells were present across all identified lineages at all analyzed developmental time points, but were most frequently associated with epithelial, muscle, and neuronal populations, where they constituted approximately 30% and 25% of the proliferating and non-proliferating squamous epithelial clusters respectively, together with approximately 10% of the striated muscle and neuronal clusters (Fig. 1b–d, Supplementary Fig. 1b). The epithelial Lgr5+ subpopulations include early (cuboidal) epithelium, distinguished by high Krt8 expression, late (squamous) epithelium with high Krt14 expression, and mixed populations of cycling epithelial cells expressing known proliferation markers Mki67, Diaph3 and Cdca8, in addition to established epithelial markers Sox2, Trp63 and Krt14[22]. The epithelial cell pool size expands markedly as development proceeds and additional suprabasal layers are formed, resulting in the highest total number of epithelial cells captured at P0. Concurring with this, most Lgr5+ epithelial cells were detected in clusters originating from the P0 time point and were predominantly located in more mature Krt14+ or proliferating subsets (Fig. 1d, e, Supplementary Fig. 1b). By integrating the timepoint origin information and expression profile for each cluster, and performing pseudotime prediction rooted to Lgr5+ cells, we validated the positioning of Lgr5+ cells within cell trajectories and their correlation with distinct developmental stages (Fig. 1f). Despite being a relatively minor cell pool in early stages, epithelial Lgr5+ were found at roots of early-stage clusters and shown to have direct trajectories towards multiple differentiated (late) subsets (Fig. 1d–f). UMAP analysis based on the timepoint origin revealed intriguing insights into the origin of the captured striated muscle cells. As suggested by Gopalakrishnan et al.[3], striated muscle cells of cranial origin have an active role in external muscle development and were observed during E13.5 and E16.5 but subsequently lost by P0. The Lgr5+ cell trajectory mirrors this pattern and shows similar changes during E13.5-16.5 and P0 stages. In contrast, non-cranial striated muscle cells emerged as a prominent muscle cell population after birth, still encompassing Lgr5+ cells (Fig. 1d). Leveraging pseudotime trajectories of Lgr5-expressing cells throughout embryogenesis, we were able to delineate their dynamic transformations within the muscle clusters, transitioning from an initially restricted pool of muscle cells and fibroblasts to fully differentiated populations of striated and smooth muscle cells detected in postnatal tissues. This detailed analysis further underscores the intricate dynamics of Lgr5+ cells within both the epithelial and muscle layers of the developing esophagus.

Further analysis of cellular identity using established lineage markers revealed that as development proceeds, captured epithelial Lgr5+ cells display elevated expression of stem/progenitor markers (Trp63, Sox2, Sox4, Sox6) whilst downregulating differentiation markers (Krt13, Ivl) (Fig. 1g) suggesting a potential stem/progenitor cell identity. Timepoint analysis also indicated significant changes in the expression profile of the Lgr5+ epithelial cell pool during development. Comparison of the top 100 upregulated genes in Lgr5+ epithelial cells at analyzed timepoints revealed only five genes were consistently expressed during development (Fig. 1h). These shared genes include Lgr5, together with well-established factors essential for organ development or cell functioning (e.g. Plcb1, necessary for organ development; the Hist1 family for chromatin structure regulation). However, the remarkably diverse composition of the Lgr5+ cell populations
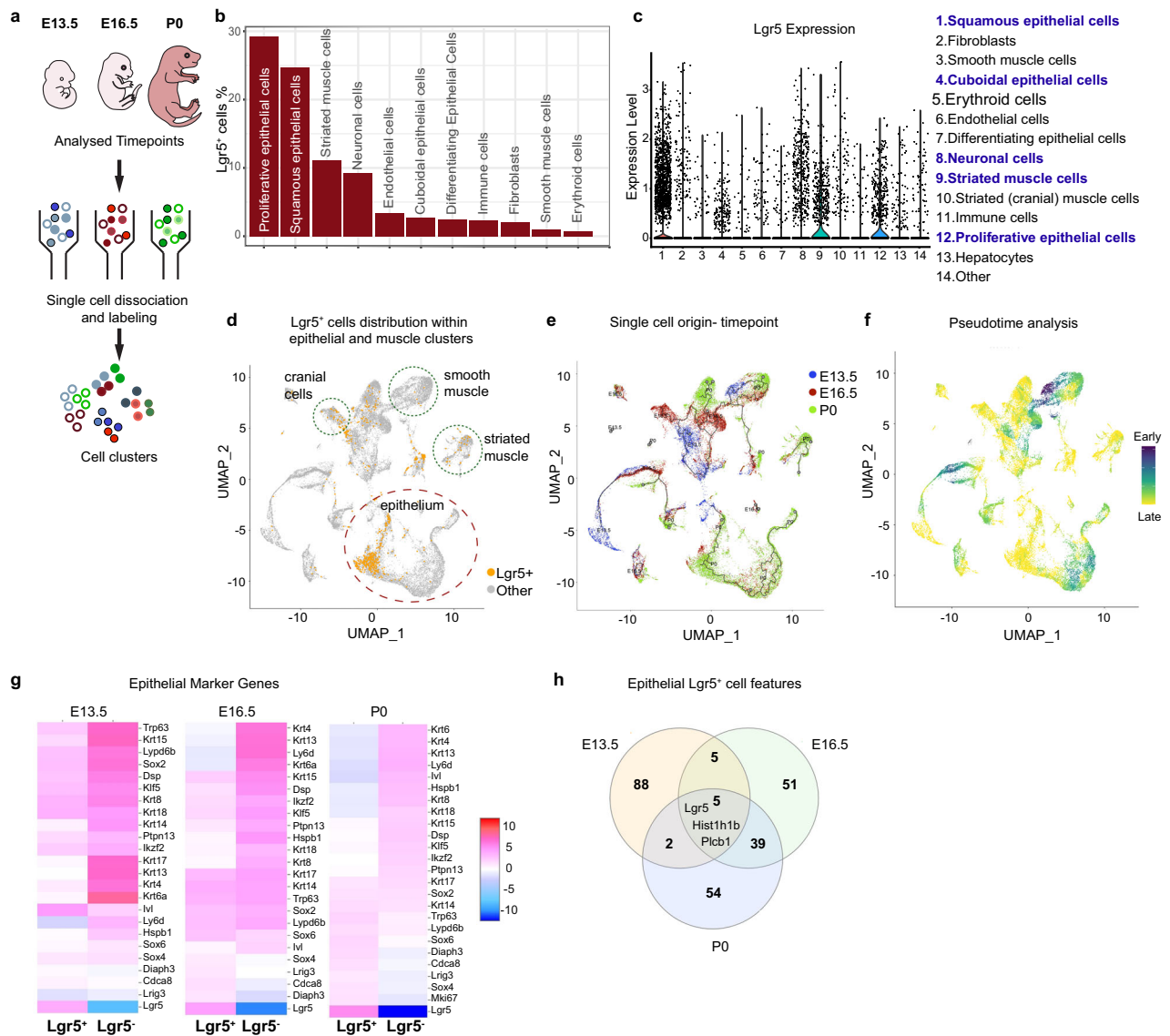
Fig. 1 | Single-cell profiling reveals the dynamic changes in the expression profile of Lgr5+ cells during esophageal development. a Experimental design for isolating esophageal single-cell suspension from E13.5, E16.5, and P0 embryos/neonates. **b** Percentage of Lrg5+ cells detected in indicated clusters. **c** Violin plots representing the distribution and expression levels of Lgr5 in the identified clusters. **d** UMAP representation of identified clusters showing Lgr5-expressing cells distribution at all analyzed timepoints. **e** UMAP clustering depicting cell timepoint origin for each cluster. **f** Pseudotime trajectory rooted in Lgr5+ cells. **g** Heatmaps showing expression of selected epithelial marker genes in Lgr5+ and Lgr5- epithelial single cells detected at E13.5, E16.5, and P0. **h** Venn diagram of the differentially expressed genes in Lgr5+ cells at analyzed timepoints.

across the different timepoints highlights the dynamic changes they are undergoing during embryogenesis. The expression profile of Lgr5+ cells at the earliest timepoint analyzed, E13.5, was highly distinct from that of Lgr5+ cells/populations at later stages. Indeed, 88 out of 100 analyzed genes were exclusively expressed in Lgr5+ epithelial cells at E13.5, showcasing the distinct molecular landscape of this subset at early developmental stages (Fig. 1h). Epithelial Lgr5+ cells at E13.5 display expression of selected epithelial stem/progenitor markers (Krt8, Krt18, Trp63) or genes known to be critical for gut development, such as, Prox1, Pdx1, Gata6 and Foxp2. Many of the detected genes are associated with cell adhesion functions (Nrp1, Magi2, Tinag, Cobll1), potentially implicating early embryonic Lgr5+ cells in the formation of epithelial layers and proper structuring of the developing esophagus (Supplementary Table 1. b). In contrast, E16.5 and P0 Lgr5+ cell expression profiles demonstrated a significant overlap, sharing nearly 50% of upregulated genes. Common upregulated genes in late developmental/early neonatal Lgr5+ cells were linked to cell adhesion

(Cdh13, Lamb3, Col17a1), cell cycle regulation (Top2a, Prc1, Mki67, Diaph3), together with DNA replication and chromatin organization (Atad2, Smc2, Dut). Uniquely expressed genes at E16.5 were associated with different signaling cascades known to play a significant role in organ development (Sox11, Cdh3, Sparc, Agrn) and cytoskeletal organization and adhesion (Cdh3, Col11a1, Birc5) (Supplementary Table 1b). At P0, the uniquely expressed genes remain associated with tissue development (Prickle2, Robo1, Fras1, Ltbp1, Lrig3), regulation of different signaling cascades (Wnt5b, Gli3, Sox4, Sema5a) and cytoskeletal adhesion (Dtnb, Itgb4, Ank3). This comprehensive analysis of Lgr5+ cells across developmental timepoints hints at their evolving role during esophageal development, predominantly in the formation of epithelial layers and tissue structuring, potentially facilitated by their engagement in cell adhesion functions and participation in critical signaling pathways.

Within the developing esophagus, the muscle layer harbors significant populations of Lgr5+ cells (Fig. 1b–d, Supplementary Fig. 1b).

These muscle cells are initially categorized into distinct clusters: striated muscle cells expressing Pax7, MyoD1, and Myf5; cranial striated muscle cells expressing Isl1, Tbx1, and Pitx2; and smooth muscle cells expressing αSMA, Cald1, and Cnn1. The majority of Lgr5+ muscle cells primarily reside within the striated muscle clusters, with a smaller population found within the smooth muscle subset. Notably, at E16.5, cranial muscle cells emerge as the dominant pool of striated cells, aligning with their active role in forming the striated external layer. By P0, as external muscle formation nears completion, the cranial cluster diminishes, giving way to the detection of striated muscle cell clusters (Supplementary Fig. 1a, b) uniformly harboring Lgr5+ cells.

To document transcriptional differences between the Lgr5+ and Lgr5- muscle populations, we conducted an analysis of the top upregulated genes in each subset. At E13.5, Lgr5+ cells represent a limited pool within the striated and smooth muscle clusters (Supplementary Fig. 1c). Comparative expression profiling of Lgr5+ and Lgr5- smooth muscle cells reveals that the former have predominantly upregulated genes associated with development and signaling (Hoxc8, Tbx3os1, Hs3st3b1). Conversely, Lgr5+ cells within the striated pool display a significant upregulation of various genes directly linked to muscle development, including Pax3, Pax7, Msc, and Col19a1. By E16.5, smooth muscle Lgr5+ cells largely maintain an expression profile not exclusively associated with muscle development, while striated Lgr5+ cells maintain robust expression of striated stem cell markers (Supplementary Fig. 1d). Postnatally, smooth muscle Lgr5+ cells express genes known to play roles in adhesion and cell signaling, such as Fbln2, Fndc1 and Pcdh9. At P0, the divergence of the Lgr5+ and Lgr5- striated muscle expression profiles markedly diminishes, likely correlating with completion of external muscle layer formation. At this stage, Lgr5+ cells are primarily enriched for transcripts involved in muscle contraction (Myh3, Myh8, Mylpf). Nevertheless, the differences are less pronounced, signifying a potential shift in the contributions of Lgr5+ cells at this stage of postnatal development (Supplementary Fig. 1e).

In summary, the combined findings from the scRNA Seq analyses at three developmental time points suggest that early Lgr5+ populations are enriched for epithelial and muscle stem cell markers and critical signaling cascades that are potentially associated with development of the esophageal epithelium and the external muscle.

Dynamic expression of Lgr5 throughout embryonic development of the esophagus

To validate scRNAseq findings and accurately document the spatial distribution and expression of Lgr5 in the developing esophagus, we employed Lgr5-2A-enhanced green fluorescent protein (eGFP) reporter mice, which facilitate the accurate visualization of endogenous Lgr5+ populations whilst retaining physiological Lgr5 expression levels[23] (Fig. 2a). To generate a more detailed expression profile for Lgr5 during embryonic development, we expanded our analysis to include five embryonic time-points: E11.5, 12.5, 13.5, 14.5, and 16.5. By incorporating a broader range of embryonic stages, we aimed to capture a more detailed picture of the evolving landscape and unravel the intricate Lgr5+ cell dynamics throughout esophageal organogenesis. Lgr5 expression in the developing esophagi was visualized via IHC for endogenous GFP (Fig. 2b). The anatomical positioning of the embryonic esophagus at each developmental stage was determined using H&E-stained sagittal embryo sections (Supplementary Fig. 2a). At E11.5, which corresponds to the final phase of gut/airway separation, we observed low-level Lgr5-GFP expression within the developing epithelium. The highest levels were consistently detected in the proximal region, indicating a potential Lgr5 expression gradient at this early time point (Fig. 2b; Supplementary Fig. 2b). Robust, localized GFP expression on well-documented Lgr5+ stem cell populations at crypt bases of small intestine from Lgr5-2A-eGFP reporter mice confirmed the accuracy of the GFP staining protocol for identifying endogenous

Lgr5 expression (Supplementary Fig. 2c). Lgr5 expression was also detected at this stage within nascent muscle layers wrapped around the developing esophagus (Fig. 2b). During the dynamic phase of epithelial development and layering after separation from the trachea[24] and until E13.5, Lgr5 expression was enriched in the most proximal regions of both the epithelial and muscle layers (Fig. 2b, Supplementary Fig. 2b).

After E13.5, Lgr5 expression in the proximal epithelium gradually declines, and the observed proximal-distal epithelial expression gradient fades as development progresses (Fig. 2b; Supplementary Fig. 2b), resulting in almost uniformly low-level expression throughout the epithelium at E16.5, when GFP reporter signals become almost undetectable. In contrast, Lgr5-GFP expression within the emerging external muscle layer becomes increasingly apparent as development advances from E11.5 to E16.5 (Fig. 2b; Supplementary Fig. 2b). Indeed, the level and distribution of GFP expression increases during late embryonic stages, indicating a potential role for Lgr5+ cells in the development and maturation of the esophageal muscle layer.

To independently validate the endogenous expression pattern of Lgr5 in the embryonic esophagus, we performed RNA in situ hybridization (ISH) using the RNAScope assay. Consistent with the observations made in reporter mouse models, we detected Lgr5 transcripts as early as E11.5 within the proximal esophageal epithelium, as well as in the mesenchymal cells adjacent to the basal layer (Fig. 2c; Supplementary Fig. 2d). Subsequently, Lgr5 signals increased within the proximal epithelium and external layers, including the submucosa and developing muscle, reaching a peak around E13.5 (Fig. 2c). After E13.5, proximal epithelial Lgr5 expression gradually declined, mirroring observations from Lgr5 reporter mice. In the surrounding submucosal region, a low level of Lgr5 expression persisted throughout the developmental period from E11.5 to 16.5. This independent confirmation further reinforces the dynamic expression profile of Lgr5 in the esophageal epithelial and muscle layers during embryonic development.

To further characterize the Lgr5-GFP+ cells in the nascent epithelium and muscle layers, we conducted immunofluorescent co-labeling (Co-IF) for GFP along with established markers for early epithelial (Krt8), early external muscle (αSMA), and proliferating (Ki67) cells. Starting from E11.5, we observed co-localization of GFP and Krt8, an epithelial marker expressed throughout the early epithelium and retained in suprabasal layers during late pregnancy/early neonatal stages (Supplementary Fig. 3a, b). In the developing muscle layer, starting from E12.5 we observed clear co-localization of GFP and smooth muscle actin (αSMA), an early marker of the external muscle layer. Additionally, the overlap between Lgr5-GFP+ cells and Ki67 in both the nascent epithelial and muscle layers highlighted the active proliferation of Lgr5+ cells during esophageal development (Supplementary Fig. 3a, b).

To further validate the scRNAseq findings, we examined the expression pattern of Lgr5 in the neonatal esophagus immediately after birth. The esophageal epithelium reaches maturation by P5, while muscle development is completed by P14[24–27]. In neonates derived from Lgr5-2A-eGFP reporter mice, GFP expression was now more prominent in the distal region of the developing esophageal epithelium (Fig. 2d). In contrast, external muscle GFP expression was most prominent in the proximal and middle regions. To validate the accuracy of the detected GFP signal in neonatal esophagus, we analyzed neonatal SI sections from Lgr5-2A-GFP and wild-type (WT) mice, as well as WT P0 esophagi (Supplementary Fig. 3c). RNAScope analyses further highlighted a new proximal-to-distal gradient in the neonatal epithelium, with highest Lgr5 levels in the distal esophagus, contrary to the pattern observed during embryonic stages, and scattered Lgr5 signal in neonatal muscle cells (Fig. 2e). We confirmed these results through qPCR analysis of
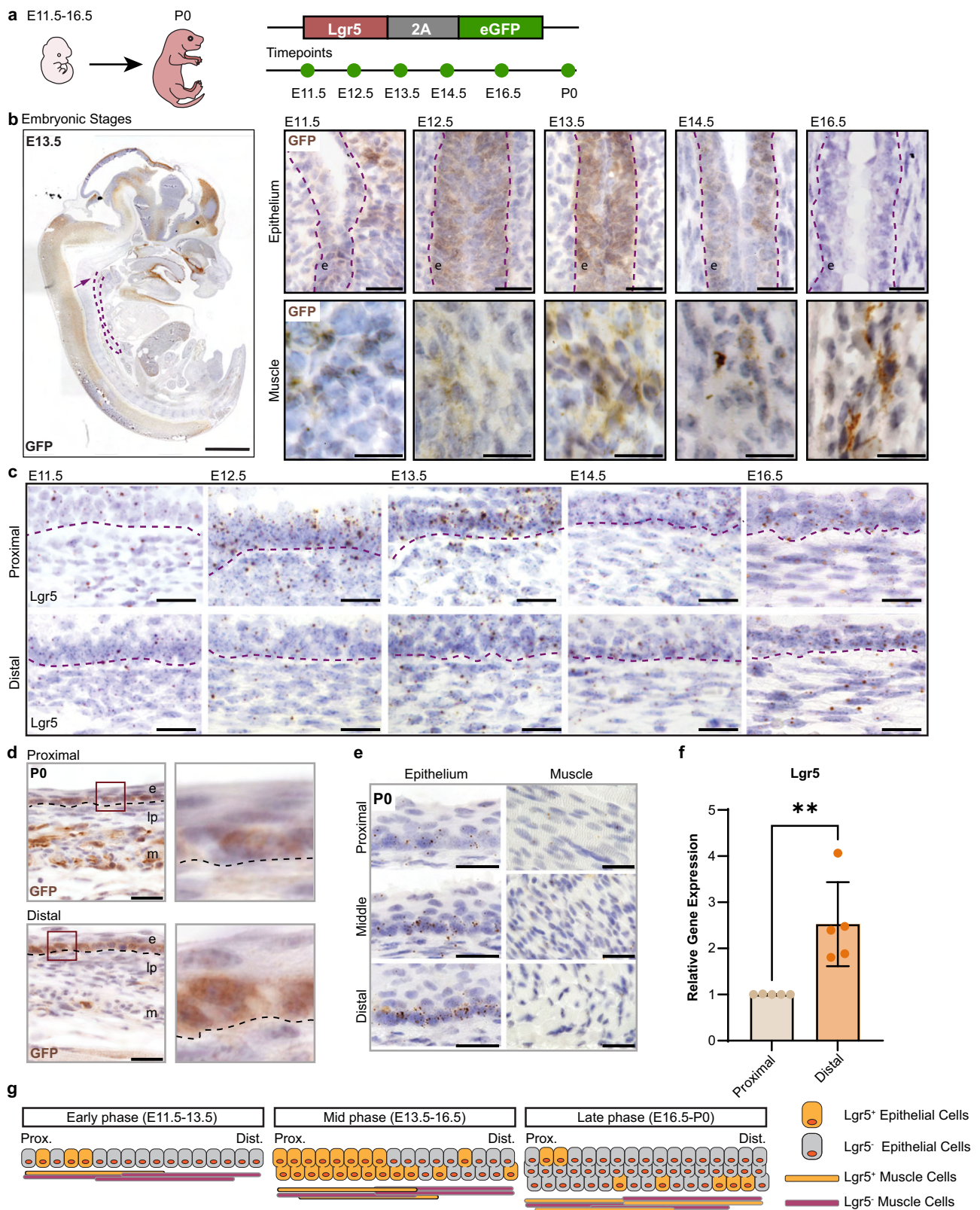
Lgr5 in neonatal esophageal epithelium from WT mice, which revealed a similar gradient of Lgr5 expression along the proximal-distal axis of the tissue (Fig. 2f).

In contrast, assessment of Lgr5 expression in the adult esophagus (8-week-old mice) using the Lgr5 reporter model showed no detectable GFP signal (Supplementary Fig. 3d), likely due to the low endogenous expression levels of Lgr5. As a positive control, adult small intestine (SI) from the same Lgr5 reporter line was analyzed, revealing strong GFP expression. However, ISH analysis revealed restricted expression of Lgr5 in the epithelium, predominantly within the basal layer of the distal epithelium, and very low, almost undetectable levels in the external muscle (Supplementary Fig. 3e). This finding supports largely developmental-specific roles for Lgr5$^+$ cells in the mouse esophagus, at least during homoeostasis.

Fig. 2 | Spatial and temporal expression of Lgr5 during esophageal development. a Schematic representation of the Lgr5-2A-eGFP mouse model and the experimental setup for Lgr5 expression analysis. **b** Immunohistochemistry (IHC) showing endogenous eGFP expression of a whole sagittal E13.5 embryo section; the dashed line outlines the esophagus; the arrow indicates the proximal esophagus. Scale bar: 1000 μm. IHC for GFP expression in proximal epithelium and muscle from E11.5 to E16.5. The dashed line indicates the separation between the epithelium and subepithelial layers. Scale bar: epithelium 25 μm, muscle 15 μm. **c** RNA in situ hybridization (ISH) demonstrating Lgr5 expression in the embryonic esophagus (proximal and distal). The dashed line indicates the separation between the epithelium and subepithelial layers. Scale bar: 25 μm. **d** Immunohistochemistry

(IHC) showing endogenous GFP expression in proximal and distal epithelium and muscle at P0. The dashed line indicates the separation between the epithelium and subepithelial structures; *e*- epithelium, *lp*- lamina propria, *m*- muscle. Scale bar: 25 μm. **e** RNA in situ hybridization (ISH) showing Lgr5 expression gradient in the P0 neonatal esophagus (proximal, middle and distal); epithelium and muscle. Scale bar 25 μm. **f** Quantitative PCR (qPCR) analysis of relative gene expression for Lgr5 at the P0 stage, normalized to the proximal region. n = 5, biological replicates. Graph represent mean ± SD with two-tailed unpaired t-test. Statistical significance: p = 0.0057. **g** Graphic illustration of Lgr5+ cell distribution in the muscle and epithelium during esophageal development. Source data are provided as a Source Data file.

Collectively, these results confirm the presence of discrete Lgr5+ populations within the epithelial and muscle compartments of the developing esophagus. The discovery of dynamic Lgr5 expression gradients along the esophagus indicates potential stage-specific contributions of Lgr5+ populations to the developing epithelial and muscle compartments during embryogenesis. During the initial stages (E11.5–13.5), a notable abundance of Lgr5+ cells is evident in the proximal epithelium, concomitant with the onset of Lgr5 expression in the developing external muscle. At intermediate stages of development (E14.5–16.5), a gradual redistribution of epithelial Lgr5 expression towards the middle and distal regions occurs, whereas Lgr5+ expression persists in the proximal muscle layer and expands towards the middle regions. At later developmental stages (E16.5-P0), characterized by transition of the immature cuboidal epithelium to a more differentiated squamous-type epithelium, Lgr5 expression shifts towards more distal regions. In parallel, robust Lgr5 expression is retained in proximal and middle regions of the developing muscle layer, whilst also expanding towards more distal areas coincident with the appearance of striated components (Fig. 2g, Supplementary Fig. 3f).
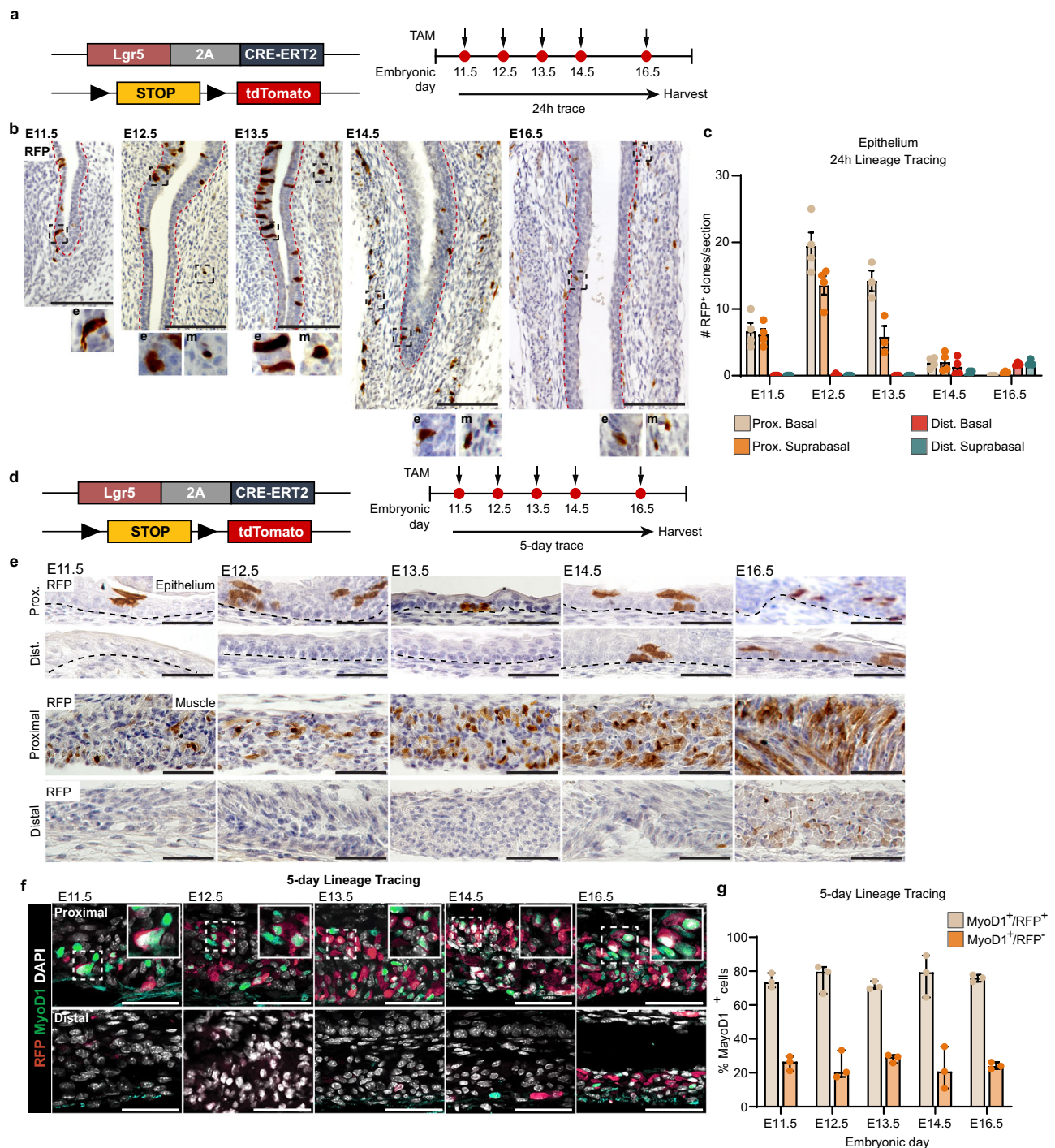
Stage-specific contributions of embryonic Lgr5+ cells to esophageal development

To functionally evaluate the stem/progenitor potential of the newly-identified embryonic esophageal Lgr5+ cell populations, we employed the Lgr5-2A-CreERT2;R26-tdTomato in vivo lineage tracing mouse model, previously utilized to document the stem/progenitor potential of Lgr5+ cells in numerous tissues[9–12,28–30]. Lgr5-driven lineage tracing was initiated at the same embryonic timepoints detailed above by injecting pregnant females with Tamoxifen (TAM), and the fate of labeled Lgr5+ cells and their progeny within the developing epithelial and muscle layers was subsequently traced over time (Fig. 3a). 24 h post-tamoxifen injection, tdTomato (RFP) reporter expression was detected in single cells and their presumptive immediate progeny within the developing proximal epithelium and external muscle layers, consistent with the localization and distribution of endogenous Lgr5-GFP+ cells at the same embryonic stage (Fig. 3b). The induction efficiency of reporter gene expression and the specificity of the IHC stains were confirmed by analyzing RFP expression in SI from the same embryo, together with a non-induced control (Supplementary Fig. 4a). The frequency of tracing initiation at each developmental stage correlated with the observed endogenous Lgr5 expression pattern, with highest levels present in the proximal nascent epithelium at initial stages E11.5–13.5 (Fig. 3b, c). Labeled epithelial cells were found in the presumptive progenitor compartment within the basal layer, and suprabasal layers (Fig. 3c). To evaluate a potential stem/progenitor cell identity for the labeled Lgr5+ epithelial cells, we examined the co-localization of RFP expression and Sox2, a reported marker of both embryonic and adult stem/progenitor cells in the esophagus[31], together with co-localization of RFP and Ki67 (Supplementary Fig. 4b). Co-expression of Sox2 was consistently observed in RFP+ cells within the basal and suprabasal layers of the proximal epithelium across the different developmental stages examined (Supplementary Fig. 4c). Co-

staining with Ki67 revealed the majority of epithelial RFP+ cells to be actively proliferating from E11.5 until E14.5. In contrast, many of RFP+ epithelial cells at E16.5 were non-proliferative, as evidenced by the decrease in RFP/Ki67 co-labeling (Supplementary Fig. 4c), potentially reflecting a change in Lgr5+ cell identity/function during epithelial maturation. Induction at later stages of pregnancy resulted in fewer labeled proximal epithelial cells, with RFP+ cells visible in middle and distal regions for the first time. While RFP signals were more apparent in the proximal epithelium initially, they were also observed in individual cells within the external muscle layer, especially in the proximal region where striated muscle cells infiltrate the developing smooth muscle layer (Fig. 3b). The number of labeled muscle cells increased with the progression of pregnancy, consistent with observations in reporter mice (Supplementary Fig. 4d).

To evaluate the contribution of labeled Lgr5+ cells and their progeny to the development of the embryonic esophagus, we initiated lineage tracing at the same developmental stages described earlier and examined expression of RFP in the esophagi collected after 5 days (Fig. 3d). We analyzed the distribution of RFP+ cells along the proximal-distal axis and their position within the epithelium, specifically within the basal-suprabasal layers. During the initial stages of esophageal development (E11.5–E13.5), epithelial cells are mainly highly proliferative cuboidal lineages. Nevertheless, our analysis reveals that a subset of these early Lgr5+ cells possess a unique ability to remain in the basal layer for extended periods and expand over time (Fig. 3e, Supplementary Fig. 4e). Furthermore, by comparing the number of RFP+ clones detected after 24 h and 5-days of lineage tracing we observed an increase in basal cell clones positive for RFP after 5 days (Fig. 3c, Supplementary Fig. 4e). Although the induction rate was relatively low in both the basal and suprabasal layers of the epithelium at later developmental stages, the 5-day time point exhibited the highest retention of labeled RFP+ clones (Fig. 3e, Supplementary Fig. 4e). The initially restricted pool of RFP+ clones, detected 24 h after initiation of lineage tracing, underwent significant expansion over the 5-day period. This resulted in the presence of more than 30 RFP+ clones per analyzed section in the proximal basal and suprabasal layers, together with ~20 RFP+ clones in the distal region per section (Supplementary Fig. 4e). Despite the relatively low labeling frequency documented in the developing epithelium at later stages, we consistently observed more robust Lgr5+ cell-driven clonal expansion at later time points post-induction, highlighting the likely stem/progenitor identity of Lgr5+ cells at later embryonic stages (Fig. 3c, Supplementary Fig. 4e).

In external muscle, the frequency of tracing induction positively correlated with progression of development, with highest levels observed at later embryonic time-points for both the proximal and distal esophagus. Tracing induction was markedly higher in the proximal versus distal esophagus at initial stages (E11.5-E13.5) but was comparable after E13.5 correlating with the relative endogenous expression levels of Lgr5 in these different regions (Supplementary Fig. 4d). After 5 days, a marked expansion of RFP+ clones was selectively observed within the proximal esophagus in mice induced at all

Fig. 3 | Stage-specific contribution of embryonic Lgr5⁺ cells to development of the esophagus. a Schematic representation of the Lgr5-2A-CreERT2; R26-tdTomato mouse model used in the experiment and the experimental setup for Tamoxifen (TAM) induction and 24-hlineage tracing. **b** Immunohistochemistry (IHC) showing 24-h lineage tracing of tdTomato⁺ (RFP⁺) cells in the developing esophagus, induced at E11.5-E16.5. Magnified regions within the epithelium (*e*) and the muscle (*m*) are shown. The dashed line separates the epithelium and sub-epithelial tissues. Scale bar: 100 μm. **c** Quantification of RFP⁺ cell number per section detected in Lgr5-2A-CreERT2; R26-tdTomato proximal and distal epithelium 24 h post-TAM induction. n = 4 (E11.5-E14.5), n = 3 (E16.5) biological replicates. **d** Illustration of the experimental setup for 5-day lineage tracing. **e** IHC assay showing 5-day lineage tracing of RFP⁺ cells in the developing esophagus, induced at E11.5-E16.5. Scale bar: 50 μm. **f** Co-IF labeling for RFP/MyoD1 at E11.5-E16.5 after 5 days of lineage tracing. Scale bar: 50 μm. **g** Percentage of MyoD1⁺/RFP⁺ and MyoD1⁺/RFP⁻ cells detected per field of view in developing muscle 5 days post-induction at the indicated developmental timepoints. n = 3, biological replicates. Data are presented as mean values ± SD. Source data are provided as a Source Data file.

time-points. In contrast, expansion of RFP⁺ tracing units was observed in distal esophagus at 5 days post- induction from E13.5 embryonic time-point (Fig. 3d, e, Supplementary Fig 4f). Due to significant expansion of RFP⁺ clones and technical challenges in accurately distinguishing between clones originating from distinct Lgr5⁺ cells, we

calculated the percentage of RFP coverage per section using the Semi-Quantitative[32] approach. (Supplementary Fig. 4f). Analysis of RFP⁺ area along the developing esophagus 5 days post-induction from E11.5–E16.5 stages revealed a proximal-distal expansion of RFP contributions to the muscle compartment during development

(Supplementary Fig. 4f). The robust expansion of Lgr5+ cell-driven tracing units within the developing proximal muscle layer correlates well with the known maturation pattern and regional expansion of striated muscle components at this stage[33]. Co-IF for RFP and striated muscle marker (MyoD1) at different embryonic stages, revealed frequent co-expression of RFP and MyoD1+ in the proximal esophagus (Fig. 3f, g), further highlighting the contribution of Lgr5+ muscle progenitors to striated muscle formation in the developing esophagus. In contrast, only a minority of Lgr5+ cell-derived tracing units colocalized with the smooth muscle marker αSMA after 5 days of tracing (Supplementary Fig. 4g). Thus, we confirm that Lgr5+ cells and their progeny in the muscle layer predominantly participate in the development of the striated layer component, with only a minor contribution to the smooth muscle layer.

To assess the long-term contributions of detected Lgr5+ embryonic cells, we initiated lineage tracing at the previously described developmental stages and examined RFP expression in adults (8-week-old mice) (Supplementary Fig. 5a). We found that only a limited number of epithelial RFP+ clones persisted following induction. Notably, retention was most prominent in mice induced at E16.5, with approximately 7 RFP+ clones per section retained in adult tissues (Supplementary Fig. 5b,c). This finding indicates that embryonic Lgr5+ cells are not restricted to short-term contributions to epithelial development during embryogenesis, but also have a role in forming the adult epithelial stem/progenitor compartments. Conversely, in the external muscle, after 8 weeks, there was a major expansion of RFP+ clones starting from the E13.5 time point. The expansion was observed throughout the muscle layer of the proximal and distal esophagus (Supplementary Fig. 5b), as evidenced by the RFP coverage area (Supplementary Fig. 5d). The detected RFP+ clones predominantly correspond to striated muscle lineages, as supported by their co-localization with Pax7 (striated muscle marker) and limited co-localization with αSMA (Supplementary Fig. 5e). This finding highlights an important contribution of embryonic Lgr5+ muscle cell progenitors to the development and maintenance of the muscle layer even in the adult esophagus.

Lineage tracing experiments in adult Lgr5-2A-CreERT2;R26-tdTomato mice confirmed the retention of a very limited pool of epithelial Lgr5+ cells with stem/progenitor cell activity, predominantly within the distal basal layer (Supplementary Fig. 5f). 8-week-old Lgr5-2A-CreERT2;R26-tdTomato mice were administered 2 doses of Tamoxifen and the esophagi were harvested 2 and 60 days later (Supplementary Fig. 5f). After 2 days, infrequent, single RFP-labeled cells were largely restricted to the basal layer of the epithelium, where they co-expressed the Krt14 marker. At 60 days, rare RFP+ clones which had expanded to encompass additional basal and differentiated epithelial cell layers were present in the distal esophagus, consistent with a stem cell/progenitor identity of a subset of adult Lgr5+ cells (Supplementary Fig. 5f). No lineage tracing was observed in the adult muscle layer at any timepoint following induction.

Technical challenges in isolating embryonic epithelium and the low expression of Lgr5 in adult tissues poses major challenges in functionally evaluating Lgr5+ populations and understanding their role in epithelial development and tissue maintenance in vivo. We therefore generated near-physiological esophageal epithelial organoids derived from adult Lgr5 reporter mice to directly evaluate Lgr5+ cell contributions to epithelial organoid development and maintenance via lineage tracing. The organoid system is a powerful tool that faithfully mimics various processes detected in organ injury repair response, development, or disease in vitro[34–37]. Organoid cultures were routinely initiated by embedding isolated epithelial cells from adult esophagus into Matrigel and cultured for up to 12 days, as previously described[38].

Isolated epithelial cells from adult Lgr5-2A-CreERT2;R26-tdTomato mice were plated in Matrigel and treated 1 day after seeding with 4-OHT (the active metabolite of Tamoxifen[39]) for 18 h to selectively activate tdTomato (RFP) expression in Lgr5+ cells (Supplementary Fig. 6a). No RFP accumulation was visible on 2-days old organoids (24 h post-treatment) using brightfield microscopy. By day 4 post-seeding, scattered RFP signals were observed in approximately 20% of organoids, documenting the presence of limited numbers of Lgr5+ cells and any early progeny within developing organoids. By day 12 post-seeding, RFP+ clones had expanded significantly in more than 70% of the organoids to cover 20-70% of the total organoid surface area (Supplementary Fig. 6b–d). Immunofluorescent co-labeling confirmed that the day-12 RFP+ clones consisted of Krt14-positive (basal layer) and Krt13-positive (suprabasal layer) lineages (Supplementary Fig. 6e). These lineage tracing analyses highlight the major contribution of adult Lgr5+ esophageal cells to epithelial organoid development in vitro.

To directly evaluate the stem cell potential of adult esophageal Lgr5+ cells via their ability to initiate organoid formation in vitro, we administered adult Lgr5-2A-CreERT2;R26-tdTomato mice with high dose Tamoxifen and FACsorted CD104+/RFP+ (enriched for Lgr5+ basal cells) and CD104+/RFP- (enriched for Lgr5- basal cells) populations after 24 h (Supplementary Fig. 6f). In agreement with the low frequency of endogenous Lgr5 expression in adult esophagus, numbers of CD104+/RFP+ cells isolated were low. Strikingly, RFP+ cells consistently initiated long-term organoid cultures, even from as few as 150 cells, that could be passaged long-term, highlighting their remarkable stem cell potential in vitro. In contrast, CD104+/RFP- cells seeded at the same low density typically failed to generate organoids or did so at a greatly reduced frequency with accompanying aberrant morphology (Supplementary Fig. 6g).

Collectively, these lineage-tracing analyses document stage-specific contributions of embryonic Lgr5+ cells to the different compartments of the mouse esophagus. In the developing epithelium, proliferating Lgr5+ cells transiently contribute to epithelial development at early stages, but later transition into a minor, yet significant pool of dedicated progenitors contributing to long-term epithelial maintenance in neonates and adults. In vitro lineage tracing and organoid initiation assays further highlight the stem cell potential of adult epithelial Lgr5+ cells. A separate Lgr5+ embryonic pool functions as dedicated progenitors providing major contributions to external muscle development and maintenance.

Epithelial Lgr5+ cells are a source of Wnt6 during development and organoid formation

Having established the dynamic, stage-dependent distribution of Lgr5+ cells and their diverse expression profiles during esophageal development, we set out to accurately map the transcriptional profiles of Lgr5+ and Lgr5- populations in situ using spatial transcriptomics. To facilitate this, we employed the 10X Visium Spatial Gene Expression platform using frozen sections of the esophagus at E13.5, E16.5, and P0. The distribution and relative expression levels of Lgr5 were analyzed for each time point, which revealed similar temporal changes as observed in the reporter and lineage tracing models (Supplementary Fig. 7a). It is important to note that the 10X Visium platform lacks single cell resolution, instead providing a merged expression profile of adjacent cells (spot diameter- 55 μm). Nevertheless, through a comparative analysis of expression profiles between esophageal Lgr5+ and Lgr5- spots at each time point, we have documented inherent transcriptional heterogeneity. In addition, we have identified potential ligands and pathways specific to Lgr5+ cells, shedding light on their potential contributions to development. We consistently observed robust enrichment of Wnt ligand expression within Lgr5+ spots, suggesting a potentially crucial role for embryonic Lgr5+ cells as signaling regulators during development (Supplementary Fig. 7b). While some ligands were found to be expressed in both Lgr5- and Lgr5+ spots (Wnt4, Wnt5a, Wnt10a), Wnt6 exhibited a strong association with Lgr5+ spots. Similar expression dynamics were observed with Wnt5b and

Wnt7b. However, for Wnt5b, although levels were higher in Lgr5+ spots, a notable portion of Lgr5− spots also showed expression, indicating that its presence was not exclusive to Lgr5+ spots. Wnt7b displayed a pattern very similar to that of Wnt6. Nevertheless, when validated using scRNAseq profiling results, the expression in the epithelium was not confirmed, suggesting that the detected signal from Spatial Transcriptomics results likely originates from non-epithelial cells (Supplementary Fig. 7c).

Integrating both our scRNAseq and spatial transcriptomics analyses, we confirmed a substantial enrichment of Wnt6 expression in Lgr5+ epithelial cells compared to their Lgr5− counterparts within the same cluster or spots (Fig. 4a). The distribution of Wnt6+ spots closely resemble the pattern observed for Lgr5+ spots (Supplementary Fig. 7d), and Wnt6+ spots displayed enrichment for Lgr5 over Wnt6− spots (Supplementary Fig. 7e), supporting a strong correlation between Wnt6 and Lgr5. The observed discrepancy between the Wnt6 enrichment levels detected by scRNAseq and spatial transcriptomics at E13.5 can likely be attributed to technical challenges arising from profiling the highly limited Lgr5+ epithelial population present at this early embryonic stage, together with the lack of single-cell resolution using the 10X Visium platform. To validate our findings, we performed qPCR on additional esophageal samples. Due to organ size limitations during embryonic development (E13.5 and E16.5) and the inability to mechanically separate the embryonic epithelium from the surrounding external muscle, we analyzed proximal (Lgr5-low region) and distal (Lgr5-high region) esophageal epithelium from P0 neonates. Our qPCR results confirmed the positive correlation between Lgr5 and Wnt6 expression in these regions (Fig. 4b). To further confirm the colocalization of Lgr5 and Wnt6 in the developing epithelium, we performed duplex ISH RNAScope (Fig. 4c), which showed robust co-expression of Lgr5 and Wnt6, particularly within the proximal embryonic epithelium. These findings were consistent with Lgr5+ cell distribution and the observed expression patterns in reporter and lineage tracing models at previously analyzed embryonic timepoints. Co-immunofluorescence analysis of embryos from Lgr5-2A-eGFP reporter mice confirmed the co-localization of GFP+ cells with Wnt6-expressing cells in the epithelium (Fig. 4d). To document whether the limited Lgr5+ epithelial pool present in adult esophagus also expresses Wnt6, we performed duplex ISH RNAScope for Lgr5 and Wnt6 on 8-week adult esophagus (Supplementary Fig. 7f). Wnt6 expression levels were comparable with those of Lgr5, displaying a clear proximal-distal gradient. Moreover, a significantly higher percentage of Lgr5+ basal layer cells in both proximal and distal regions co-express Wnt6 (Supplementary Fig. 7g), demonstrating that, despite their limited frequency, adult Lgr5+ cells still serve as a source of Wnt6. These findings collectively highlight a significant enrichment of Wnt6 in embryonic Lgr5+ cells throughout various developmental stages, which persists into adulthood. This identifies this cell population as a potentially crucial epithelial source of endogenous Wnt6 ligand during esophageal development.

We next investigated potential functions of the resident Lgr5+ epithelial cells in establishing and maintaining the esophageal epithelium via a targeted cell ablation approach using our Lgr5-2A-DTR mouse model[34]. The Lgr5-2A-DTR model expresses an Lgr5+ cell-driven Diphtheria Toxin (DT) receptor gene (Fig. 4e) to facilitate selective ablation of Lgr5-expressing cells via administration of DT. Ablation of Lgr5+ cells in developing embryos via DT administration to pregnant females causes embryonic lethality/miscarriage due to body-wide loss of essential Lgr5+ populations, precluding evaluation of Lgr5+ cell contributions to esophageal development in vivo. To circumvent this limitation, we established the explant culture model, which facilitates the study of esophageal development ex vivo[40]. Embryonic esophagi from Lgr5-2A-DTR and WT mice at E13.5 were isolated and cultured in a Matrigel matrix for up to 5 days, with the addition of 0.1 μg/ml DT to the growth media from day 0 and DT continually refreshed every 2 days (Fig. 4e). After 2 days of culture, Lgr5+ cell-ablated explants

showed marked alterations in the expression of epithelial and muscle layer markers and elevated cell death (marked by Cp3 expression), although proliferation levels remained unchanged (Supplementary Fig. 7h). By day 5, a significant reduction in the size of DT-treated Lgr5-2A- DTR explants compared to DT-treated WT explants was evident (Fig. 4f, g). Furthermore, explants with ablated Lgr5+ cells exhibited significant changes in epithelial morphology, with nearly complete absence of epithelial layering by day 5. This was accompanied by a marked disorganization of the external muscle layer and increased expression of the smooth muscle marker αSMA (Fig. 4h). Lgr5+ cell ablation also resulted in a complete loss of proliferation within the epithelial layer, and a significant reduction within the muscle layer (Fig. 4h). Dying cells (CP3+) were detected in the lumen, epithelium and muscle. To validate the previously observed correlation between Lgr5 and Wnt6 expression, we conducted immunofluorescence staining for Wnt6 and Krt14 on explants (Supplementary Fig. 7i). Lgr5-2A-DTR explants grown in medium containing DT showed a significant reduction in Wnt6 expression compared to control groups, further emphasizing the strong correlation between Lgr5 and Wnt6 in developing epithelium.

We next assessed the consequences of Lgr5+ cell ablation on organoid formation and maintenance using esophageal organoid cultures derived from adult Lgr5-2A-DTR mice (Fig. 4i). DT was introduced to the culture media either on the day of seeding or at later time points, and any resulting effect on organoid growth was evaluated for long-term and short-term treatments (Fig. 4j, Supplementary Fig. 8a, g, j). For long-term ablation, DT was added on day 0, day 4 or day 7 and replenished every second day (Supplementary Fig. 8a, g), while for short-term ablation, DT was left for 24 h before harvesting organoids on day 7 and day 12 (Supplementary Fig. 8j). To assess the ablation efficacy, RNAScope analysis was performed on organoids harvested on day 7 and day 12 for both short-term and long-term treatments (Supplementary Fig. 8b, k). Wild-type (WT) organoids lacking the DT receptor displayed comparable Lgr5 levels and comparable organoid size and morphology when cultured in the presence or absence of DT, confirming that the dose of DT used for Lgr5+ cell ablation did not cause general toxicity in these organoids. In contrast, Lgr5-2A-DTR organoids subjected to both short-term and long- term DT exposure demonstrated a noticeable reduction in Lgr5 signal, confirming efficient ablation of the resident Lgr5+ population (Supplementary Fig. 8b, k).

To evaluate the effect of Lgr5+ cell elimination on organoid formation efficacy and growth, we analyzed organoids treated long-term with DT from day 0, 4 and 7. Ablation initiated at day 0 resulted in significantly impaired organoid formation and growth compared to WT cultures grown in the same conditions, visible at early (day 7) and late (day 12) time-points (Fig. 4j, Supplementary Fig. 8c, d, e). Given the dynamic expression pattern of Lgr5 in the developing esophagus, we also investigated the consequences of long-term DT administration at later time points. When DT was added either 4 or 7 days after seeding, the impact on organoid growth was reduced and the organoid formation efficacy was no longer significantly affected. Moreover, as organoids matured, the reduction in Lgr5+ cells' impact became more pronounced, likely because other cells capable of compensating for Lgr5 elimination had emerged (Supplementary Fig. 8g, h, i).

To evaluate any influence of Lgr5+ cell ablation on epithelial maturation in vitro, we performed Co-IF labeling for embryonic cuboidal (Krt8) and squamous (Krt14) epithelial markers. Adult-derived organoid models have been previously found to recapitulate embryonic markers and signaling cascades as part of an injury response phenotype[36,37]. Indeed, 12 days after seeding, untreated organoids presented bi-layered epithelial structures harboring distinct basal Krt14+ and suprabasal Krt8+ layers consistent with ongoing maturation of the epithelial cells (Fig. 4k). In contrast, DT-treated Lgr5-2A-DTR cultures displayed an altered epithelial marker expression
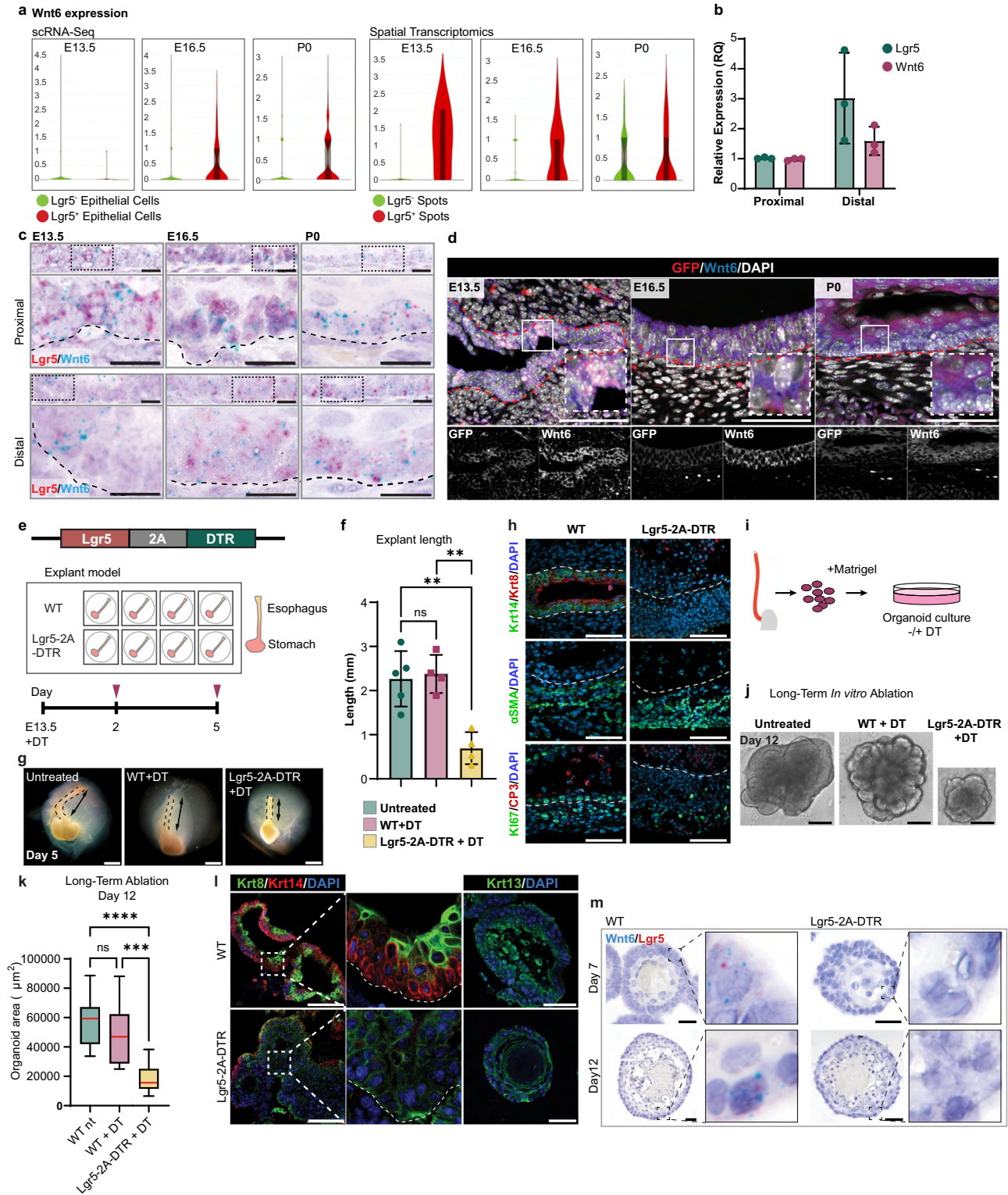
pattern, with a markedly expanded Krt8+ compartment and a concomitant reduction in expression of the suprabasal differentiated marker Krt13. These findings were confirmed by qPCR, demonstrating robust upregulation of Krt8 in DT-treated Lgr5-2A-DTR organoids (Supplementary Fig. 8f). qPCR analysis revealed no significant differences in the levels of squamous epithelial marker Krt14, likely because the adult epithelial cells used here already display an established Krt14 expression pattern. Prolonged expression of Krt8 in DT-treated organoids indicates a possible delay in the transition from cuboidal to squamous epithelium upon elimination of the Lgr5+ pool.

Given our profiling data indicates that Lgr5+ cells may serve as an endogenous source of Wnt6 ligands, we next considered whether ablation of Lgr5+ cells would also influence the expression of Wnt6 in organoid culture. Similar to our observations in the esophageal epithelium in vivo, Lgr5 and Wnt6 colocalized within the same cells in WT organoids (Fig. 4l). In Lgr5-2A-DTR organoids treated long-term with DT, there was a noticeable decrease in both Lgr5 and Wnt6 expression (Fig. 4l), supporting their co-expression in the same cell compartments. Wnt6 immunofluorescence on the same DT-treated Lgr5-2A-DTR organoids confirmed the reduction of Wnt6 following Lgr5+ cell

Fig. 4 | Epithelial Lgr5⁺ cells are a source of Wnt6 during esophageal development and organoid formation. a Violin plots depicting Wnt6 expression in Lgr5- and Lgr5⁺ single cells and spots detected using scRNAseq and Spatial Transcriptomics. Single set of sequencing data was generated for each analyzed timepoint. The central box shows the interquartile range (IQR), and the line within the box denotes the median expression level. Whiskers extend to 1.5 times the IQR from the quartiles, capturing the range of the data within this range. Data are presented as median ± IQR. **b** Relative gene expression in proximal and distal epithelium isolated from the esophagus at P0. n = 3, biological replicates. Data are presented as mean values ± SD. **c** Duplex RNA ISH assay showing Lgr5 and Wnt6 expression in embryo sections at E13.5, E16.5, and P0. Scale bar: 25 μm. **d** Co-IF labeling for GFP and Wnt6 expression in embryo sections at E13.5, E16.5, and P0. Scale bar: 50 μm. **e** Schematic representation of the Lgr5-2A-DTR mouse model, the explant culture assay and the experimental design for ex vivo cell ablation. **f** Statistical analysis of explant length after 5 days of ex vivo culture. Statistical significance: p = 0.001 (*WT + DT vs Lgr5-2A-DTR + DT*), p = 0.0031 (*WT nt vs Lgr5-2A-DTR + DT*), Sample size: n = 5 (WT nt), n = 4 (WT + DT, Lgr5-2A-DTR + DT), biological replicates. Data are

presented as mean values ± SD with two-tailed unpaired t-test. **g** Five-day-old WT and Lgr5-2A-DTR explants. Scale bar: 1 mm. **h** Co-IF labeling for Krt14/Krt8, αSMA, and Ki67/Cp3 expression in 5-day-old explants. Scale bar: 50 μm. **i** Schematic representation of the experimental design for in vitro ablation assay. **j** Organoids grown with and without diphtheria toxin (DT) treatment for 12 days. Data are presented as mean values ± SEM. Statistical significance: p < 0.0001, n = 5 (biological replicates). Scale bar: 50 μm. **k** Box plots showing the distribution of organoid size grown with or without DT treatment for 12 days. Each box plot represents the distribution of the data, with the box indicating the IQR and the line within the box denoting the median value. Whiskers extend to the 2.5th and 97.5th percentiles, capturing the range of the data within this interval. Outliers beyond these percentiles are shown as individual points. Data are presented as median ± IQR. Statistical significance: p < 0.0001, n = 3 biological replicates. Statistical significance was assessed using Kruskal-Wallis test. **l** Co-IF labeling for Krt8 and Krt14 expression in WT and Lgr5-2A-DTR 12-day-old organoids. Scale bar: 100 μm. **m** Duplex RNA ISH assay showing Lgr5 and Wnt6 expression in WT and Lgr5-2A-DTR 12-day-old organoids. Scale bar: 25 μm.

ablation, both at early (day 7) and late (day 12) stages (Supplementary Fig. 9a).

To directly evaluate the functional importance of Wnt6 for epithelial organoid formation, we employed CRISPR/CAS9 gene editing to generate Wnt6 knockout (KO) organoids. Validated KO organoids failed to expand in culture (Supplementary Fig. 9b), highlighting a critical dependency on endogenous Wnt6 expression. We then evaluated whether supplementation with recombinant Wnt6 protein can rescue the growth of Lgr5-2A-DTR cultures following long-term Lgr5⁺ cell ablation, assessed at day 7 and day 12 (Supplementary Fig. 9c). This revealed a robust reversal of the impaired organoid growth driven by loss of the resident Lgr5⁺ cells, further highlighting their key role as endogenous sources of Wnt6 (Supplementary Fig. 9c, d). Co-IF labeling confirmed that supplementing Wnt6 also reversed the effects of Lgr5 ablation on the distribution and expression of Krt14 and Krt13 (Supplementary Fig. 9e).

Together, our findings reveal a robust correlation between endogenous Lgr5 and Wnt6 expression within the developing esophageal epithelium. This suggests that Lgr5⁺ cells likely play a pivotal role as a significant source of Wnt6 during the early stages of development, representing a substantial subset of the Wnt6⁺ cell population within the epithelium. Significantly, in vitro and ex vivo experiments using organoids and explants confirmed that elimination of Lgr5⁺ cells led to a marked decrease in Wnt6 expression. Furthermore, the absence of Wnt6 substantially impaired organoid formation and resulted in dysregulation of Krt8⁺ and Krt14⁺ population dynamics and distribution. Remarkably, external supplementation of Wnt6 successfully reverted the phenotype of organoids, restoring the Krt13/Krt14 distribution observed in WT organoids. These findings underscore the functional importance of Lgr5⁺ cells in maintaining the integrity of the developing esophageal epithelium.

Discussion

Here, we employ single-cell RNA sequencing, spatial transcriptomics and faithful reporter mouse models to document the presence of phenotypically heterogenous populations of Lgr5⁺ cells predominantly within squamous epithelial and striated muscle compartments of the developing mouse esophagus. Lgr5 is detected soon after foregut separation at E11.5 and subsequently displays a dynamic expression pattern within the developing epithelium until birth. Comparative expression profiling of Lgr5⁺ epithelial populations during latter stages of embryogenesis highlights the transcriptional diversity of E13.5 populations relative to that of E16.5 and P0 populations. Notably, late embryonic/early neonatal populations display elevated expression of stem/progenitor markers such as Trp63, Sox2, and a concomitant reduction in differentiation markers such as Krt13, Ivl, hinting at a

potential epithelial stem/progenitor cell identity. In accordance with this, in vivo lineage tracing assays using our well-validated Lgr5-2A-CreERT2;R26-tdTomato mouse model establish a transient contribution of proliferating E11.5-13.5 Lgr5⁺ cells to early epithelial development, whilst late embryonic/early neonatal Lgr5⁺ populations function as a pool of dedicated progenitors contributing to the establishment and maintenance of the adult epithelium (Fig. 5a, b).

A highly dynamic Lgr5⁺ expression pattern is also observed within the developing external muscle layer, with expression becoming increasingly apparent as development advances from E11.5 to E16.5, predominantly within striated muscle lineages enriched for stem cell markers such as Pax7. The observed pattern correlates well with known developmental dynamics of the embryonic muscle layer. Early in development, external muscle primarily comprises smooth muscle cells, but after the gut/airway separation the muscle layer undergoes proximal to distal replacement by cranial mesoderm-derived migratory striated muscle cells[27], which infiltrate the proximal region of developing external muscle and initiate the formation of the striated muscle layer[41]. Importantly, our in vivo lineage tracing assays identify these embryonic Lgr5⁺ populations as a major pool of stem/progenitors contributing to external muscle development and maintenance in the mouse esophagus (Fig. 5a, b).

Comparative profiling of late embryonic/early neonatal epithelial populations reveals a marked enrichment of several Wnts, including the canonical ligand Wnt6. Wnt signaling is a critical regulator of epithelial development and maintenance in many organs, where it forms a key component of the local niche for resident stem/progenitor cells[42]. In the embryonic mouse esophagus, Wnt6 is known to be selectively expressed in the developing esophageal epithelium, but its specific source and function are poorly defined[40]. Here, we demonstrate that CRISPR-mediated KO of Wnt6 in adult epithelial organoids prevents organoid growth, highlighting a critical dependence on endogenous Wnt6. We further identify nascent Lgr5⁺ populations within the developing epithelium as a key source of endogenous Wnt6 required for the formation and maintenance of esophageal organoids in vitro. It remains to be established whether the Wnt6 ligands supplied by the nascent Lgr5⁺ epithelial populations serve an autocrine function to regulate endogenous Lgr5⁺ stem/progenitor cell activity, or instead act in a paracrine fashion to influence proliferation/cell fate decisions of neighboring populations during esophageal development (Fig. 5a).

Essential contributions of the newly documented Lgr5⁺ populations to the nascent esophagus are further highlighted by the severe impairment to ex vivo embryonic explant development following targeted ablation of resident Lgr5⁺ populations. The elimination of resident Lgr5⁺ cells from cultured epithelial esophageal organoids induces
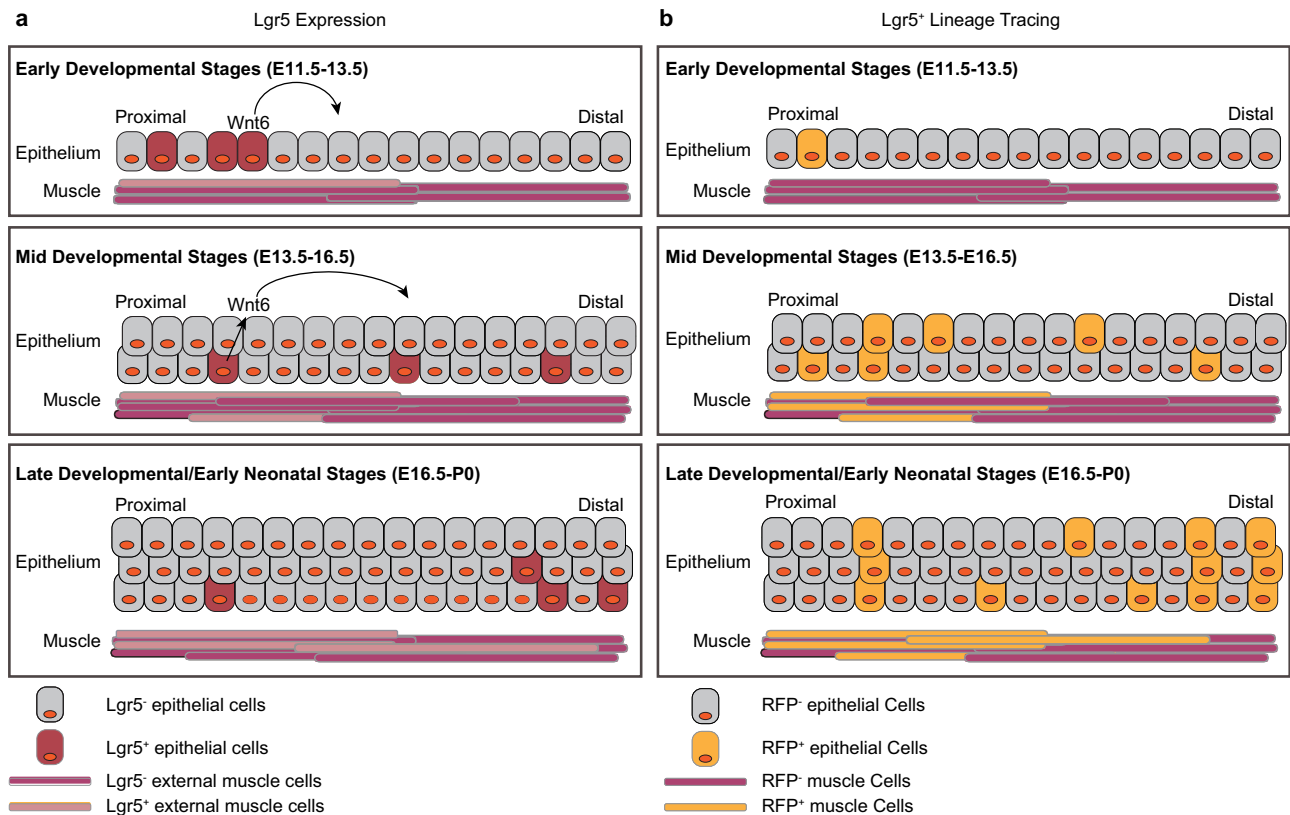
Fig. 5 | Localization and contribution of embryonic Lgr5⁺ cells during esophageal development. a Schematic illustration depicting Lgr5 expression during development of epithelial and muscle layers in embryonic esophagus, acting as a major source of Wnt6. **b** Schematic illustration of Lgr5-driven lineage tracing in the epithelium and muscle following induction at different developmental stages.

a substantial delay in maturation. This delay is marked by prolonged expression of the embryonic basal layer marker Krt8, reduced expression of suprabasal marker Krt13, and a decrease in organoid size. Our study underscores the crucial role of Lgr5⁺ cells as both embryonic stem/progenitors and as a major source of endogenous Wnt6, which plays a significant role in regulating epithelial patterning. While Wnt6 is a key contributor to epithelial organoid formation in vitro and likely also during epithelial development, it is unlikely to be the sole critical player. Lgr5⁺ cells within the developing external muscle express various key ligands and growth factors, such as Notch, Shh, and BMPs, which could potentially influence epithelial development through short-range signaling interactions. Further studies are needed to elucidate potential crosstalk between the Lgr5⁺ pool in the muscle layer and epithelial development.

In the adult esophagus, minor epithelial pools of Lgr5⁺ cells are found predominantly within basal compartments of more distal regions, where they make limited, yet significant, contributions to epithelial maintenance throughout life. Strikingly, isolated adult Lgr5⁺ epithelial cells display potent stem cell potential in organoid initiation assays modeling epithelial injury responses, potentially reflecting an important role for adult Lgr5+ role in driving epithelial repair in vivo.

Taken together, the findings presented in this study provide insights into the role of Lgr5⁺ stem/progenitor cells in the development of the different esophageal tissue compartments during embryogenesis. This discovery significantly advances our understanding of esophageal development by revealing the presence of previously unrecognized populations of epithelial Lgr5⁺ cells with stem/progenitor properties, and Lgr5⁺ external muscle cells that critically contribute to striated muscle formation. In addition to their stem/progenitor functions, these Lgr5⁺ populations likely supply key ligands to modulate local signaling pathways necessary for orchestrating proper organogenesis and tissue maturation in the mouse esophagus. Reactivation of key developmental programs is associated with tissue regeneration and cancer development in adults. It will therefore be interesting to explore potential contributions of Lgr5⁺ populations to tissue regeneration and disease in the esophagus. The comprehensive findings presented here establish a foundation for future studies aimed at deciphering the functional significance and therapeutic potential of Lgr5⁺ cells in esophageal development and disease.

Methods
Mouse models

Animal experiments

For induction of lineage tracing from embryonic stages, 0.2 mg/g body weight of tamoxifen (Sigma, T5648) was injected intraperitoneally to the pregnant female mice at E11.5-E16.5. Timed pregnancies were staged relative to a vaginal plug detection representing embryonic day 0.5 (E0.5). Progesterone (Sigma, PHR1142) was dissolved in sesame oil (Sigma, S3547) and injected s.c. at 50 µg/g body weight to pregnant females at the same time points.

Cell dissociation and in vitro/ex vivo culture

For epithelial cell dissociation and organoid culturing. Esophagi were harvested from adult mice in cold phosphate-buffered saline (PBS) with 1% penicillin/streptomycin (Gibco, 15140122), 5 ml/esophagus. To isolate epithelial cells, muscle and epithelial layer were mechanically separated as previously described[44]. Epithelium separated from external muscle layer was finely chopped using scalpel blades and collected in a 15 ml Falcon tube with 0.25% Trypsin-EDTA (Gibco, 25200056)/2000 U/ml DNAse I (Qiagen, #79256), digestion solution at 37 °C for 45 min with vigorous vortexing every 15 min (~3 ml digestion solution per sample). After digestion, Trypsin was inactivated with a 3× volume of 10% FBS + growth medium (growth medium details indicated below), followed by vigorous pipetting. The cell suspension was passed through a 70-µm sterile strainer attached to a 50 ml Falcon tube. Subsequently, the suspension was filtered through a 40-µm sterile strainer attached to a 50 ml Falcon tube to obtain a single-cell suspension. Obtained single cell suspension was centrifuged at 430 g for 3 min at 4 °C. After centrifugation, the supernatant was discarded, and the pellet was resuspended in growth factor-reduced Matrigel mixed with growth medium at a 1:1 ratio. Isolated primary esophageal cells were suspended in a 15-50 µl Matrigel/Growth Medium mix, with the volume adjusted according to the specific experiment and plate used. Cell counts ranged from 150 to 5000 cells per well, with adjustments made depending on the experiment and well volume. The cells were then cultured in growth medium.

Organoid growth medium. 50% WRNF/Advanced DMEM/F-12 media (Gibco, 12634010) supplemented with 1X Glutamax (Thermo Fisher, 35050061), 1% Penicillin/Streptomycin (Gibco, 15140122), 1X B27 (Invitrogen, 17504044), 10 mM Nicotinamide (Sigma, N0636), 10 µM p38 MAPK inhibitor- SB202190 (StemCell Technologies, SB202190), 50 ng/ml EGF (Invitrogen, PHG0311), 2 µM A83-01 (Tocris, 2939). 10 µM Rock-inhibitor -Y27632 (Sigma, Y0503) was added to single- cell cultures for the first three days of culture.

Culture medium was changed every second day. Organoids were analyzed on day 1, 4, 7, 11 or 12 of culture. For ablation assays DT (Sigma, D0564) was added at a concentration of 0.05–0.075 µg/ml to the culture medium to induce Lgr5+ cell ablation at the indicated time points. For Wnt6 supplementation assays recombinant Wnt6 protein 0.1 µg/ml (Novus, H00007475-P01) was added to growth medium containing 0.075 µg/ml DT. The treatment was replenished every second day.

For esophageal explant culture. All dissection steps were performed on ice. Embryos were harvested from Lgr5-2A-DTR and C57BL/6J mice at E13.5. Embryonic esophagi were dissected under a dissection microscope using needles and microdissection scissors. Embryos were placed in a petri dish on ice, and the ventral side of each embryo was opened to expose the heart, lung, and gut. Using a fine needle, the esophagus was carefully separated from the surrounding organs. Together with the stomach, it was transferred to a 1.5 ml Eppendorf tube containing ice-cold 1x PBS/1% PenStrep (1 ml per esophagus). The stomach and pharyngeal muscles were left attached. Matrigel matrix was prepared as previously described[40] with slight modifications. Matrigel was mixed with explant culture medium in a 1:1 ratio. Explant culture medium composition: Opti-MEM (Gibco, 14025), 20% WRNF,

1% Penicillin/Streptomycin (Gibco, 15140122), 10% FBS, and 1 ng/ml FGF (StemCell Technologies, 78003.1). The prepared Matrigel/culture medium mix was placed on the dish bottom (~50 µl/well), allowing it to set for 15–30 min at 37 °C. Dissected embryonic esophagi were transferred directly from ice onto the Matrigel dome and overlaid with additional Matrigel prepared in the same manner as the bottom layer, then left to set for an additional 30 minutes. Embedded explants were covered with growth medium. The medium was refreshed every 2 days. For the DT-treated groups, 0.1 µg/ml DT was added to the growth medium from day 0 onwards.

Generation of Wnt6 knockout in mouse epithelial organoids

WT mouse organoids were cultured in Matrigel (Corning, #356231) domes in standard esophageal organoid growth medium. After 5 passages organoids were dissociated using 0.25%Trypsin (Gibco, #25200056) to obtain single cell suspension, and viability determined using 0.4% Trypan Blue (Gibco, #15250061) staining. 37.5 pmole sgRNA (IDT), 1 µg spCas9 protein (IDT, #1081059) and 1.25 µl R buffer (Neon transfection kit, Invitrogen, #MPK1096) was mixed and incubated at room temperature for 20 min to form RNP complexes. RNP complexes were electroporated into 450,000 cells using Neon transfection system (Invitrogen, #MPK5000) at 1700 V, 20 ms, 1 pulse. Electroporated cells were equally seeded into two wells of 48-well dish in WRNF growth medium with Y-27632 (Stemcell Technologies, #72304). 72 h post-electroporation, gDNA was isolated using Quickextract™ DNA extraction solution (Lucigen, # QE09050). PCR was performed using the Q5® high-fidelity PCR kit (NEB # M0491L), following manufacturer's instructions, to amplify target site. Primers are provided in Supplementary Table 2. Target site was then Sanger sequenced to check for indels. Extensive overlapping peaks at the sgRNA binding site show high editing efficiency. Pooled population organoids were then dissociated using 0.25%Trypsin/EDTA and seeded at low densities ranging from 50 to 1000 cells per well in 48-well dish to generate clonal organoids. Five days later, clonal organoids were manually picked and transferred to individual wells in 96-well dish.

RNA isolation and qPCR

RNA was isolated from cells using QIAzol Lysis Reagent (Qiagen, #79306) or RNeasy Universal Plus Kit (Qiagen, #73404) according to the manufacturer's instructions. cDNA was generated using Superscript III (Life Technologies, #18080044) according to the manufacturer's instructions. qPCR was performed in triplicate per gene for at least three biological replicates using GoTaq SYBR Green dye (Promega, A6002) according to the manufacturer's instructions. Relative quantification of gene expression was analyzed with Step One Software v2.1 (Applied Biosystems) or Bio-Rad CFX Manager using the ΔΔCT method with Gapdh as an endogenous reference. The qPCR primers are described in Supplementary Table 2.

Immunohistochemistry and immunofluorescence

All IHC/IF assays were performed according to the standard protocol. Tissues were fixed in 4% paraformaldehyde (PFA) overnight at 4 °C before paraffin embedding. IHC/IF was performed on deparaffinized and rehydrated 6-8 µm tissue sections. Antigen retrieval was carried out by heating slides at 121 °C for 20 min using 2100 Antigen Retriever (Aptum Biologics) either in a modified citrate buffer, pH 6.1 (S1699, DAKO) or Tris/EDTA buffer, pH 9.0 (S2367, DAKO). The following primary antibodies were employed: rabbit anti-Krt8 (1:400; Abcam, ab53280, lot 1008419-3), rabbit anti-cleaved Caspase3 (1:200; Cell Signaling, 9661, lot 47), rabbit anti- Ki67 (1:200; ThermoFisher, MA5-14520, lot YB3824981), mouse anti-Ki67 (1:200; BD Transduction Laboratories, 550609, lot 6195670), chicken anti-GFP (1:100; Abcam, ab13970, lot ar3190550-17), rabbit anti-GFP (1:200; Cell Signaling, 2956 S, lot 2), rabbit anti-RFP (1:200; Rockland, 600-401-379, lot 46510), mouse anti-RFP (1:100; Abcam, ab125244, lot ar3426475-1),

mouse anti-Krt14 (1:200; Abcam, ab7800, lot 1013337-1), mouse anti-MyoD1 (1:100, Agilent Dako, M3512, lot 10115416), mouse anti-Sox2 (1:200, Millipore, MAB4343, lot 3464576), mouse anti-αSMA (1:200, Invitrogen, MA511547, lot W03247394), rabbit anti-Krt13 (1:200, Abcam, ab92551, lot 1016806-6), rabbit anti-Wnt6 (1:200, Elabscience, E-AB-17612, lot DK2984), mouse anti-Pax7 (1:100, DSHB, AB_528428). The peroxidase-conjugated secondary antibodies used were mouse/rabbit EnVision + (DAKO) for HRP IHC or anti-chicken/rabbit/mouse Alexa Fluor 488 (1:500, Invitrogen, #A21202, lot 2563848; #A11008, lot 2521157), Alexa Fluor 594 (1:500, Invitrogen, #A11005, lot 49818A; #11012, lot 1704538), Alexa Fluor 647 (1:500; Invitrogen, #A21236, lot 2482947; #A31573, lot 2674379) IgG for IF. IHC slides were mounted using DPX (Sigma 1.07979.0500) and IF slides were mounted using Hydromount (National Diagnostics, HS-106) with Hoescht or DAPI as nuclear counterstain. Immunostainings were repeated on at least three tissue sections per tissue block and with at least three biological repeats. Only representative immunostainings were included in the manuscript. H&E staining was performed on deparaffinized and rehydrated 6 μm tissue sections which were stained with Haematoxylin 2 (Richard-Allan Scientific, 7231L) followed by Scott's blue reagent (0.2% NaHCO3 (w/v), 2% MgSO4 (w/v) in water), then Eosin (Sigma, HT110132). Stained sections were dehydrated and mounted using DPX (Sigma 1.07979.0500).

Flow cytometry

Cells were dissociated as previously described, and the suspension was centrifuged at 430 × g for 3 min at 4 °C. The resulting pellet was resuspended in HBSS containing 5% fetal bovine serum (FBS, Hyclone, SH30071.04) and 1% PenStrep (Gibco, 15140122) (solution volume: 500–1000 μl per esophagus processed) before being filtered through a 40 μm strainer. CD104 Antibody (FITC-conjugated, 1:750, BioLegend 123606, lot B193078) was added to the suspension and incubated on ice for 30 min. Following incubation, the samples were washed three times with PBS containing 1% Pen/Strep and centrifuged at 430 × g for 3 minutes at 4 °C. The resulting pellet was then resuspended in a 5% FBS/1% PenStrep/HBSS solution and sorted using a BD Influx Cell Sorter (BD Biosciences). Cells were collected in a 5% FBS/1% PenStrep/HBSS solution for subsequent organoid culture.

RNA in situ hybridization (ISH)

RNAscope experiments were performed on deparaffinized and rehydrated 6-8 μm thick sagittal embryo sections (time points E11.5-E16.5), neonatal esophagus sections (time point P0), and adult esophagus sections from 8-week-old samples. RNAscope 2.0 FFPE Brown Reagent Kit was used for singleplex experiments and RNAscope 2.5 HD Duplex Assay for duplex staining (Advanced Cell Diagnostics, ACD). All procedures were performed according to manufacturer's instructions. Probes used in this study were: Lgr5 (C1, C2), Wnt6 (C1). Adjacent sections were hybridized for DapB as negative control and PPIB as positive control for each experiment. All procedures were performed according to manufacturer's instructions. Probes used are listed in Supplementary Table 3.

Mouse embryo collection and sample preparation for single cell RNA-sequencing

For the single-cell analysis, timed breedings were set up between C57BL/6J mice, and the day of plug detection was considered embryonic day 0.5. The esophagi were microdissected at E13.5/E16.5/P0 from 12, 9, and 7 embryos/neonates, respectively, and dissected tissues pooled for each time point. Samples for each time point were isolated and pooled from 2 to 3 litters. Single-cell dissociation was done by cold active protease digestion as described previously[21]. Dissected tissues were transferred to ice-cold PBS with 5 mM CaCl2, 10 mg/ml of Bacillus Licheniformis protease (Sigma, P5380-250MG), and 125 U/ml DNAse I (Qiagen, #79256), and incubated on ice with mixing by pipet every 2 min. Samples were incubated on ice for 7 min (E13.5), 9 min (E16.5) and 12 min (P0) and single-cell dissociation was confirmed under a brightfield microscope. Cells were then transferred to a 15 ml conical tube, and 3 ml ice-cold PBS with 10% FBS (FBS/PBS) was added. Cells were pelleted at 4 °C (300 × g for 5 min) and resuspended in 2 ml PBS/FBS. Cells were washed three times in 5 ml PBS/0.01%BSA (PBS/BSA) and resuspended in a final cell concentration of 100,000 cells/ml for scRNA-seq. Single-cell suspensions of each stage were loaded onto the Chromium Instrument (10x Genomics) to generate single-cell gel beads in emulsion. Single-cell RNA- Seq libraries for high-throughput sequencing were prepared using the Chromium Single Cell Library and Gel Bead Kit (10x Genomics). All samples were singleplexed together and sequenced in an Illumina NovaSeq 6000 (HWI-ST1276).

Single cell RNA sequencing data processing

Library preparation and sequencing were performed by Novogene using their standard workflow. FASTQ files for each sample were generated and processed by Novogene according to their protocol. Briefly, reads from single-cell RNA sequencing (scRNA-seq) were processed using Cell Ranger software (10x Genomics) with default parameters. This processing included alignment to the reference genome, unique molecular identifier (UMI) collapsing, UMI counting, and initial quality control. Filtered gene expression matrices, containing only cellular barcodes, were generated. Data quality control was subsequently performed using the Seurat R package.

Alignment summary. Splicing-aware alignment of reads to the genome was performed using the STAR aligner in Cell Ranger. The transcript annotation GTF file was utilized to categorize reads into exonic, intronic, and intergenic regions, based on their alignment to the genome. A read was classified as exonic if at least 50% of it intersects with an exon; as intronic if it is non-exonic and intersects with an intron; and as intergenic if it falls outside these regions. Exonic reads were further aligned to the transcriptome using the provided annotation. Reads that aligned to exons of annotated transcripts and were on the same strand as the transcript were considered mapped to the transcriptome. Only uniquely mapped reads among the mapped set were used for unique molecular identifier (UMI) counting.

Cell calling and UMI counting. The cell-calling algorithm, based on the EmptyDrops method within Cell Ranger, was employed to identify cell-associated barcodes. This was achieved by evaluating their unique molecular identifier (UMI) counts and RNA profiles.

Quality control summary. For each cell, quality control metrics including the total number of counts and the proportion of counts derived from mitochondrial genes were calculated. Cells were filtered out from downstream processing if they met any of the following criteria: fewer than 200 genes detected, genes with non-zero counts in at most 3 cells, more than 8000 features (indicative of potential multiplets), or if the proportion of counts attributable to mitochondrial genes exceeded 50%.

Identification of highly variable genes (HVGs). Following cell filtering, gene expression matrices from each sample were imported into Seurat. Gene expression values were normalized by dividing the total counts for each cell, multiplied by a scale factor (the median UMI counts for all cells in the sample), and then log-transformed using log1p. The gene expressions were subsequently scaled. Highly variable genes (HVGs) were identified using the FindVariableGenes function, selecting genes with the highest standardized variance via the selection.method = 'vst'. The top 3000 most variable genes identified by Seurat in each sample were used to compute principal components (PCs). The counts-per-median ($CPM_{i,j}$) were calculated as $Count_{i,j}/$

Totalcountj * M, and the expression values (Ei,j) were computed as log10(CPMi,j + 1).

Single cell RNA sequencing data analysis

Data was collected from Novogene experimental output after cellranger analysis (10X Genomics). Three separate datasets corresponding to three data points (E13.5, E16.5 and P0) were merged using Seurat's merge function[45]. From the resulting dataset, a subset was generated with the following criteria: cells were retained if they exhibit a gene count of at least 300, a total count of UMIs of more than 1500, and no more than 15% of mitochondrial RNA. Additionally, only genes exhibiting the top 2000 highest cell-to-cell variation in the datasets used in the downstream analysis. Surviving feature counts for each cell were divided by the total counts for that cell and multiplied by a scale.factor of 10,000; this was then natural-log transformed to obtain the final LogNormalisation. Data was subsequently scaled using a Z-score scaling. Principal component analysis (PCA) was run on the scaled, normalized dataset in order to apply a UMAP projection (15 dimensions)[46]. A 20-nearest-neighbor graph was constructed in order to apply Seurat's shared nearest neighbor (SNN) modularity optimization based clustering algorithm (resolution = 0.5). Clusters were merged based on observed expression levels of several markers, including Lgr5 (using a threshold of 1.5 in scaled magnitude to differentiate "high" from "low" expression). Heatmaps were produced using the ComplexHeatmap package[47], with no clustering on the rows nor columns. Pseudotimes were computed using monocle, as described below. After casting to a monocle cds object[48] using as.cell_data_set (from the SeuratWrapper package), cells were reclustered using monocle's own clustering function (Leiden community detection method[49], UMAP projection). The principle graph on the clusters was extracted using the reversed graph embedding algorithm with partition[50]. Cells were ordered using the Lgr5+ (highly expressing, see above) cells as roots, thus generating the pseudotime trajectories. Loupe Browser (version 6.4.0, 10x Genomics) and cloupe files (generated by Novogene) were used to identify differentially expressed genes in Lgr5+ and Lgr5- cells at E13.5, E16.5, and P0. Epithelial and muscle clusters were defined in each dataset using common markers (Supplementary Table 1a). Lgr5+ cells were identified within each epithelial cluster using the "Gene/Feature Expression" function in Loupe Browser. Cells were considered positive if the expression level was greater than 0, ensuring the inclusion of all cells with detectable Lgr5 levels. The "Significant Feature Comparison" function of Loupe Browser was then employed to compare the expression profiles of Lgr5+ and Lgr5- cells, generating a list of upregulated genes for both groups. These lists were compared across timepoints to identify genes consistently expressed in Lgr5+ cells across all stages. Violin plots were generated using the "Violin Plots of Gene/Feature Expression" Loupe Browser function.

Spatial transcriptomics sample preparation and RNA quality evaluation

Embryos were collected from pregnant females (C57/BL6) at E13.6 and E16.5 stage and esophagi were dissected from P0 neonates. The sex of embryos and neonates was not determined. All samples were dissected on ice and fresh frozen in Tissue-Tek® O.C.T. Compound (#4583). After tissue freezing and prior to sample processing, approximately 8 tissue sections (8 µm thickness) were collected for RNA quality evaluation using the RNeasy Mini kit (Qiagen, #74104). Extracted total RNA was measured using the Agilent Bioanalyzer (Agilent, RNA 6000 Pico kit, Part number 5067–1513) to obtain RIN scores. The remaining portions of the same samples were cryo-sectioned at 8 µm thickness and placed onto Visium Spatial Gene Expression Slides (Visium Spatial Gene Expression Slide & Reagent Kit, 16 rxns PN-1000184). Slides with sample sections were stored at −80 °C for 24 h before processing. Spatial gene expression libraries were generated following the 10× Genomic Visium Spatial Gene Expression protocol (User Guide,

CG000239 Rev E, Product number 1000184). Libraries were sequenced on the Illumina NovaSeq (HWI-ST1276).

Data processing

Raw sequencing data were processed by 10x Genomic using the 10x Genomic Space Ranger pipeline(version-2.0.0) to generate fastq files. Sequences were aligned to mm10 genome to gene expression count using Space Ranger default settings. Gene expression level of Refseq coding genes were quantified using Space Ranger's default settings.

Data analysis

Spatial transcriptomics data were analyzed using cloupe files generated by 10x Genomics and Loupe Browser (version 6.4.0, 10x Genomics). Spots covering the embryonic esophagus were manually selected based on histological features using the Loupe Browser. This selection involved visually inspecting the sections to identify spots corresponding to the esophagus. The expression levels of genes of interest were evaluated across the selected spots. To detect spots expressing each analyzed gene, we employed the "Gene/Feature Expression" function of Loupe Browser. The Scale and Attribute menu within this software was used for filtering and selecting features based on gene expression levels. For each gene analyzed the parameters in the Scale and Attribute menu were set to "Log2" and "Feature Average," with the selection criterion of "> 0 counts." For all genes, the "Select by Count" feature was utilized to set a threshold by selecting spots with an expression level greater than 0. This ensured that all spots with detectable expression of analyzed genes were included in the analysis. Violin plots generated to compare the expression levels of different genes between spots were created using the Violin Plots feature of the Gene/Feature Expression tool in Loupe Browser.

Microscopy imaging

The following microscopes were used for image acquisition: Nikon Ni-E microscope/DS-Ri2 16.25MP color camera with 2.5x F-mount coupler was used for IHC, H&E and RNAscope slides. IF samples were imaged on Confocal Microscopes Zeiss LSM 780/LSM 800/LSM 880. Cultured cells and organoids were captured with Evos AMEX1000 and Evos M5000 Imaging Systems.

Image processing and analysis

Brightfield and immunofluorescent images were processed using ImageJ or ZEN 3.4 Blue Edition Software. Adobe Photoshop and Adobe Illustrator were used for white balance and contrast adjustments.

Quantification and statistical analysis

Statistical analyses were performed using Microsoft Excel and GraphPad Prism. Datasets were evaluated for normality using the Shapiro-Wilk test. Variables that follow a normal distribution were analyzed by the unpaired, two-tailed Student's t-test and Two-way ANOVA tests. Datasets that do not follow a normal distribution were analyzed using Kruskal–Wallis test. A p-value less than 0.05 was considered statistically significant.

Statistics and reproducibility

All samples discussed in this manuscript were measured and analyzed as biological replicates. Reproducibility was confirmed by at least three independent experiments. All IHC/IF/ISH assays were done on at least 3 biological replicates, at least 3 sections were analyzed for each replicate. For scRNAseq samples were collected from n = 12 (E13.5), n = 9 (E16.5), n = 7 (P0) and pulled together for sequencing. For Spatial Transcriptomics one capture area was analyzed for each timepoint.

Reporting summary

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

Data availability

The scRNAseq and Spatial Transcriptomics datasets generated in this study are available in the Gene Expression Omnibus (GEO) database under accession codes GSE272184 and GSE271795. Source data are provided with this paper.

Code availability

No new codes were generated in this study. List of all code packages used in the analysis is provided in Source Data file.

References

1. Rosekrans, S. L., Baan, B., Muncan, V. & van den Brink, G. R. Esophageal development and epithelial homeostasis. *Am. J. Physiol. Gastrointest. Liver Physiol.* **309**, G216–G228 (2015).
2. Mittal, R. K. *Motor Function of the Pharynx, Esophagus, and its Sphincters* (Morgan & Claypool Life Sciences Copyright © 2011 by Morgan & Claypool Life Sciences., 2011).
3. Gopalakrishnan, S. et al. A cranial mesoderm origin for esophagus striated muscles. *Dev. Cell* **34**, 694–704 (2015).
4. Yazaki, E. & Sifrim, D. Anatomy and physiology of the esophageal body. *Dis. Esophagus* **25**, 292–298 (2012).
5. Fausett, S. R. & Klingensmith, J. Compartmentalization of the foregut tube: developmental origins of the trachea and esophagus. *WIREs Dev. Biol.* **1**, 184–202 (2012).
6. Yu, W. Y., Slack, J. M. & Tosh, D. Conversion of columnar to stratified squamous epithelium in the developing mouse oesophagus. *Dev. Biol.* **284**, 157–170 (2005).
7. DeWard, A. D., Cramer, J. & Lagasse, E. Cellular heterogeneity in the mouse esophagus implicates the presence of a nonquiescent epithelial stem cell population. *Cell Rep.* **9**, 701–711 (2014).
8. Kabir, M. F. et al. Single cell transcriptomic analysis reveals cellular diversity of murine esophageal epithelium. *Nat. Commun.* **13**, 2167 (2022).
9. Barker, N. et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003–1007 (2007).
10. Barker, N. et al. Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. *Cell Stem Cell* **6**, 25–36 (2010).
11. Ng, A. et al. Lgr5 marks stem/progenitor cells in ovary and tubal epithelia. *Nat. Cell Biol.* **16**, 745–757 (2014).
12. Leushacke, M. et al. Lgr5-expressing chief cells drive epithelial regeneration and cancer in the oxyntic stomach. *Nat. Cell Biol.* **19**, 774–786 (2017).
13. Theodosiou, N. A. & Tabin, C. J. Wnt signaling during development of the gastrointestinal tract. *Dev. Biol.* **259**, 258–271 (2003).
14. Verzi, M. P. & Shivdasani, R. A. Wnt signaling in gut organogenesis. *Organogenesis* **4**, 87–91 (2008).
15. Thompson, C. A., DeLaForest, A. & Battle, M. A. Patterning the gastrointestinal epithelium to confer regional-specific functions. *Dev. Biol.* **435**, 97–108 (2018).
16. Neyaz, A. et al. LGR5 in Barrett's esophagus and its utility in predicting patients at increased risk of advanced neoplasia. *Clin. Transl. Gastroenterol.* **12**, e00272 (2020).
17. Zhang, X., Yuan, A., Zhao, X., Li, Z. & Cui, G. Tumoral expression of CD166 in human esophageal squamous cell carcinoma: implications for cancer progression and prognosis. *Cancer Biother. Radiopharm.* **35**, 214–222 (2020).
18. Lv, Z. et al. Expression and functional regulation of stemness gene Lgr5 in esophageal squamous cell carcinoma. *Oncotarget* **8**, 26492–26504 (2017).
19. Yoon, J. Y., Brezden-Masley, C. & Streutker, C. J. Lgr5 and stem/progenitor gene expression in gastric/gastroesophageal junction carcinoma - significance of potentially retained stemness. *BMC Cancer* **20**, 860 (2020).
20. Quinlan, J. M., Yu, W. Y. & Tosh, D. Isolation, culture, and characterisation of mouse embryonic oesophagus and intestine. *Methods Mol. Biol.* **633**, 81–90 (2010).
21. Adam, M., Potter, A. S. & Potter, S. S. Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: a molecular atlas of kidney development. *Development* **144**, 3625–3632 (2017).
22. Zhang, Y. et al. Development and stem cells of the esophagus. *Semin Cell Dev. Biol.* **66**, 25–35 (2017).
23. Seishima, R. et al. Neonatal Wnt-dependent Lgr5 positive stem cells are essential for uterine gland development. *Nat. Commun.* **10**, 5378 (2019).
24. Rishniw, M. et al. Molecular aspects of esophageal development. *Ann. N. Y. Acad. Sci.* **1232**, 309–315 (2011).
25. Romer, A. I., Singh, J., Rattan, S. & Krauss, R. S. Smooth muscle fascicular reorientation is required for esophageal morphogenesis and dependent on Cdo. *J. Cell Biol.* **201**, 309–323 (2013).
26. Patapoutian, A., Wold, B. J. & Wagner, R. A. Evidence for developmentally programmed transdifferentiation in mouse esophageal muscle. *Science* **270**, 1818–1821 (1995).
27. Kablar, B., Tajbakhsh, S. & Rudnicki, M. A. Transdifferentiation of esophageal smooth to skeletal muscle is myogenic bHLH factor-dependent. *Development* **127**, 1627–1639 (2000).
28. Jaks, V. et al. Lgr5 marks cycling, yet long-lived, hair follicle stem cells. *Nat. Genet.* **40**, 1291–1299 (2008).
29. Barker, N. et al. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608–611 (2009).
30. Leung, C. et al. Lgr5 Marks Adult Progenitor Cells Contributing to Skeletal Muscle Regeneration and Sarcoma Formation. *Cell Rep.* **33**, 108535 (2020).
31. Que, J. et al. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. *Development* **134**, 2521–2531 (2007).
32. Crowe, A. R. & Yue, W. Semi-quantitative determination of protein expression using immunohistochemistry staining and analysis: an integrated protocol. *Bio Protoc.* **9**, e3465 (2019).
33. Baguma-Nibasheka, M., Fracassi, A., Costain, W. J., Moreno, S. & Kablar, B. Striated-for-smooth muscle replacement in the developing mouse esophagus. *Histol. Histopathol.* **34**, 457–467 (2019).
34. Suarez-Martinez, E., Suazo-Sanchez, I., Celis-Romero, M. & Carnero, A. 3D and organoid culture in research: physiology, hereditary genetic diseases and cancer. *Cell Biosci.* **12**, 39 (2022).
35. Li, X. et al. Organoid cultures recapitulate esophageal adenocarcinoma heterogeneity providing a model for clonality studies and precision therapeutics. *Nat. Commun.* **9**, 2983 (2018).
36. Drakhlis, L. et al. Human heart-forming organoids recapitulate early heart and foregut development. *Nat. Biotechnol.* **39**, 737–746 (2021).
37. Montenegro-Miranda, P. S. et al. A novel organoid model of damage and repair identifies HNF4α as a critical regulator of intestinal epithelial regeneration. *Cell Mol. Gastroenterol. Hepatol.* **10**, 209–223 (2020).
38. Kasagi, Y. et al. The esophageal organoid system reveals functional interplay between notch and cytokines in reactive epithelial changes. *Cell. Mol. Gastroenterol. Hepatol.* **5**, 333–352 (2018).
39. Fuhrmann-Benzakein, E., García-Gabay, I., Pepper, M. S., Vassalli, J. D. & Herrera, P. L. Inducible and irreversible control of gene expression using a single transgene. *Nucleic Acids Res.* **28**, E99 (2000).
40. Dixon, K. J., Turbic, A., Turnley, A. M. & Liebl, D. J. Explant methodology for analyzing neuroblast migration. *Bio Protoc.* **7** https://doi.org/10.21769/BioProtoc.2249 (2017).
41. Krauss, R. S., Chihara, D. & Romer, A. I. Embracing change: striated-for-smooth muscle replacement in esophagus development. *Skelet. Muscle* **6**, 27 (2016).

42. Clevers, H. Wnt/beta-catenin signaling in development and disease. *Cell* **127**, 469–480 (2006).
43. Sun, X., Jackson, L., Dey, S. K. & Daikoku, T. In pursuit of leucine-rich repeat-containing G protein-coupled receptor-5 regulation and function in the uterus. *Endocrinology* **150**, 5065–5073 (2009).
44. Ko, K.-P., Zhang, J. & Park, J.-I. Establishing transgenic murine esophageal organoids. *STAR Protoc.* **3**, 101317 (2022).
45. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
46. McInnes, L., Healy, J. & Melville, J. UMAP: uniform manifold approximation and projection for dimension reduction. https://arxiv.org/abs/1802.03426v3 (2020).
47. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
48. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
49. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
50. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).

Acknowledgements

Author contributions

L.K. designed and performed experiments, analyzed data, and wrote the manuscript, CL. designed and performed experiments, K.A.M. performed histology, RNAScope, and qPCR experiments. S.K. performed organoid culture, RNAScope and histology experiments. S.P. and B.L. performed bioinformatics data analyses. N.B. supervised the project and revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-51559-4.

Correspondence and requests for materials should be addressed to Nick Barker.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

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