

## Reviewed Preprint

v1 • April 20, 2026

Not revised

## ✉ For correspondence:

weijiankai@ouc.edu.cn

## Competing interests: No

competing interests declared

Funding: See [page 15](#)

Reviewing editor: Ariel Chipman,  
The Hebrew University of Jerusalem,  
Israel

© 2026, Ge et al. This article is  
distributed under the terms of the  
[Creative Commons Attribution  
License](#), which permits unrestricted  
use and redistribution provided that  
the original author and source are  
credited.

# The endoderm cell trajectory of urochordate *Styela clava* reveals the dual developmental origin and evolution of digestive tract

Yonghang Ge<sup>1</sup>, Wei Zhang<sup>1</sup>, Penghui Liu<sup>1</sup>, Jianqing Bi<sup>1</sup>, Haiyan Yu<sup>1</sup>, Bo Dong<sup>1,2,3</sup>, Jiankai Wei<sup>1,2,3</sup> ✉<sup>1</sup>Fang Zongxi Center for Marine EvoDevo, College of Marine Life Sciences, Ocean University of China, Qingdao, China •<sup>2</sup>Laboratory for Marine Biology and Biotechnology, Qingdao Marine Science and Technology Center, Qingdao, China •<sup>3</sup>MoE Key Laboratory of Evolution & Marine Biodiversity, Institute of Evolution and Marine Biodiversity, Ocean University of China, Qingdao, China

## eLife Assessment

This potentially **important** paper questions the evolutionary origin of the tunicate endoderm, using single-cell sequencing on a developmental series of the ascidian *Styela clava* that covers metamorphosis and gut development. The authors base their conclusions on a comparison with the development of mouse gut endoderm, where they point out similarities in the origin of tissues, perhaps representing a case of "deep homology". This work has the potential to make a significant contribution to the field of chordate evolution, but in its current form, the evidence it presents is **incomplete** and is limited by a problematic discussion of evolutionary implications and by major issues regarding the clarity and cogency of data presentation.

<https://doi.org/10.7554/eLife.110546.1.sa2>

## Abstract

The digestive system exhibits extensive diversity in developmental mechanisms and morphology across metazoans, yet the evolutionary origins underlying its organ differentiation remain unclear. Here, single-cell RNA sequencing was employed to investigate endodermal cell lineage specification during metamorphosis in the urochordate *Styela clava*, a newly established model for chordate evolution. By profiling 26,099 cells across five stages, we identified 21 major cell clusters and reconstructed the endodermal differentiation trajectories. Our analysis reveals two larval endodermal progenitor populations with distinct differentiation potentials. Pseudotime and RNA velocity analyses indicate that these progenitors give rise to stomach and intestinal lineages, respectively. Cross-species comparisons reveal putative homologous relationships between ascidian endodermal lineages and mouse definitive and visceral endodermal lineages, suggesting dual origins of digestive tract in chordates. We also identified conserved TGF- $\beta$  and FGF regulatory programs in digestive organ patterning and highlight earlier fate restriction of stomach and intestinal progenitors in ascidians compared to vertebrates. These findings provide insights into how chordate digestive organs evolved from ancestral endodermal patterning programs.

## Introduction

The digestive system, comprised of digestive tract and accessory organs, plays a central role in nutrient acquisition and processing across animals. While exhibiting morphological diversity depending on habitats and feeding strategies, digestive systems universally share conserved functional processes including ingestion, digestion, absorption, and elimination.<sup>1</sup>

Developmentally, the digestive tract and accessory organs are primarily originated from endodermal derivatives. However, the origins and formation patterns of these organs exhibit notable differences between invertebrates and vertebrates.<sup>2</sup> For example, in *Drosophila melanogaster*, the midgut derives from endoderm while foregut and hindgut originate from ectoderm.<sup>3</sup> In *Caenorhabditis elegans*, both ectodermal and mesodermal cells contribute to its foregut.<sup>4</sup> By contrast, in vertebrates especially mammals, digestive tracts are entirely endoderm-derived, with fore-, mid- and hindgut all arising from this germ layer.

One reason for this distinction lies in the profound differences of behaviors and origins of endodermal cells during early development. The endoderm cells of invertebrate remain internalized after gastrulation, whereas vertebrate endoderm undergoes dynamic morphogenetic movements.<sup>2,5</sup> Moreover, mammalian endoderm development displays additional complexity through its dual origin. The definitive endoderm (DE) arises during gastrulation from epiblast cells, while extra-embryonic (primitive endoderm) forms earlier from the inner cell mass. These two endodermal lineages collectively give rise to the embryonic gut endoderm and resulting gut tube.<sup>6</sup>

Furthermore, there is also distinct formation patterns during organogenesis of digestive organs between invertebrates and vertebrates. For example, in vertebrates, the gut tube first becomes regionalized into foregut, midgut, and hindgut. As development proceeds, The foregut gives rise to the esophagus, trachea, stomach, lungs, thyroid, liver, biliary system, and pancreas; whereas the midgut forms the small intestine and the hindgut forms the large intestine.<sup>7</sup> But unlike the highly compartmentalized and organ-specific digestive systems of vertebrates, invertebrate gut structures are generally less complex and exhibit more diffuse regional specialization.<sup>8–11</sup> These developmental differences underscore distinct cellular lineages between invertebrate and vertebrate digestive systems, which has hindered our understanding of how digestive tracts originated and evolved from invertebrates to vertebrates.

Urochordates, as the closest living relatives of vertebrates, offer a valuable model for exploring the developmental and evolutionary origins of the vertebrate organs.<sup>12</sup> Their unduplicated genomes and conserved developmental pathways make them especially informative for reconstructing ancestral chordate traits.<sup>13</sup> They also provide insights into the evolutionary origins of cell types in vertebrates.<sup>14</sup> Recent advances in comparative genomics using *Styela clava* have facilitated cross-species analyses of gene regulatory networks involved in organ evolution.<sup>15,16</sup>

In ascidians, the digestive system developed during metamorphosis and composed of pharynx, esophagus, stomach, intestine and pyloric gland, lacking complex accessory organs and glands.<sup>17</sup> Although the structure of ascidian digestive system is relatively simple, the basic function in nutrition processing is similar to vertebrates.<sup>18,19</sup> Previous studies have shown that the genes encoding enzymes involved in digestion, absorption, and immune function in ascidian are very similar to those of vertebrates.<sup>20–22</sup> Meanwhile, the expression patterns of some transcription factors are also conserved between ascidians and vertebrates.<sup>1,20</sup> However, the cellular heterogeneity and lineage dynamics of the urochordate digestive tract remain largely unexplored especially at single-cell resolution.

To investigate the evolution and diversification of digestive systems, it is essential to integrate not only morphological data but also information on cellular lineage and molecular regulation. While vertebrate models, particularly mammals, have provided extensive insights into digestive organ morphogenesis and function, the evolutionary history that gave rise to these systems are less well characterized. Here, we employ single-cell RNA sequencing to map the transcriptional landscapes and developmental trajectories of digestive tract cells in a urochordate model. Our analysis reveals conserved regulatory programs but divergent cell specification strategies underlying digestive tract formation, and provides new insights into the cell lineage evolution of digestive tract across chordates.

## Results

### The single-cell atlas of ascidian *S. clava* across metamorphosis

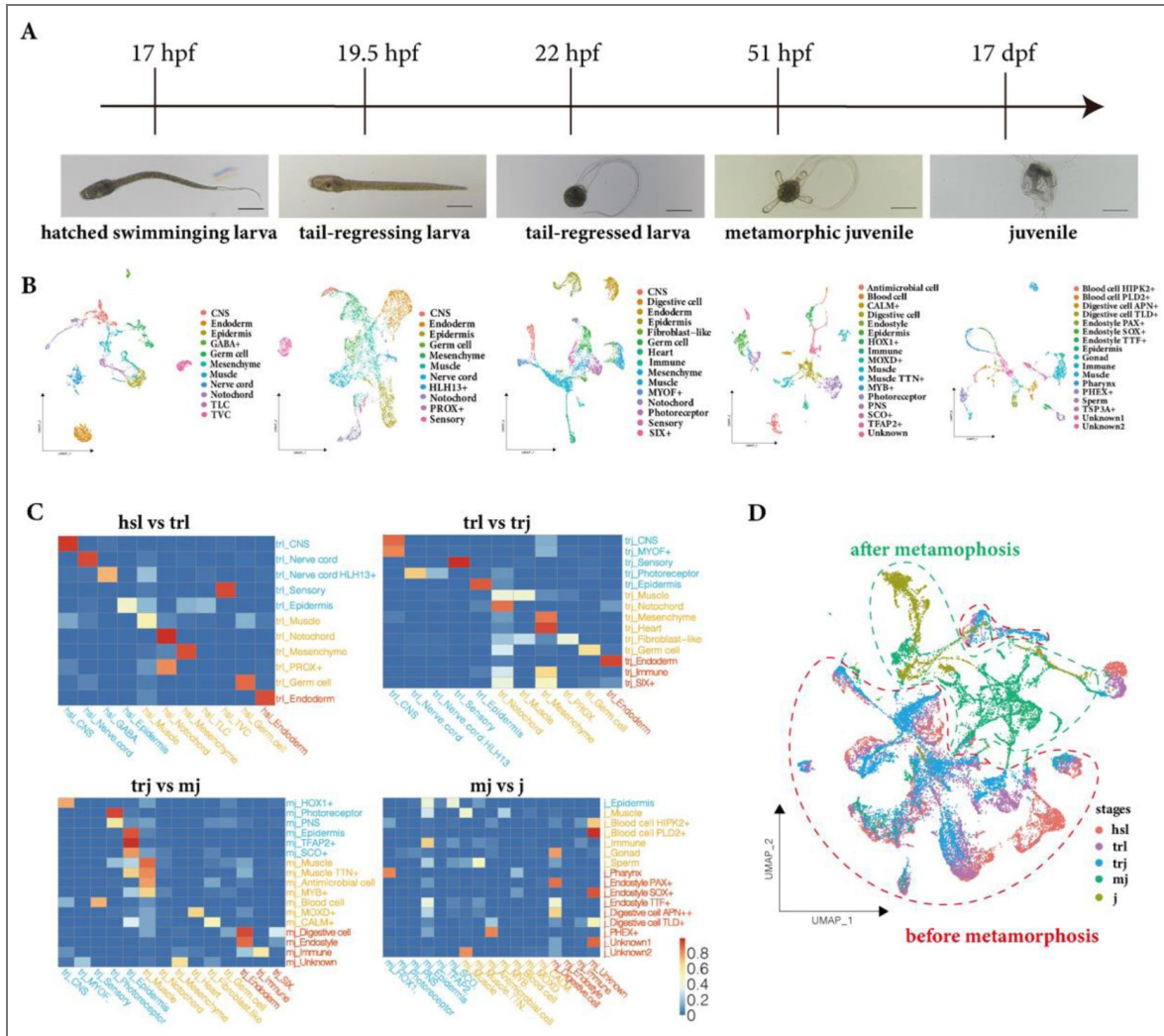
In ascidian, endodermal organogenesis is mainly occurred during metamorphosis. In order to trace the crucial developmental events, we performed single cell RNA-sequencing (scRNA-seq) using samples from five developmental stages including hatched swimming larva (hsl), tail regressing larva (trl), tail regressed larva (trj), metamorphic juvenile (mj), and juvenile (j) stages. These stages span the main metamorphic processes of *S. clava* (Figure 1A). After quality control, a total of 26,099 cells were acquired, with an average expression of 11,413 genes per stage (Figure S1). The major cell types were identified (11, 11, 15, 17, and 17 cell types at hsl, trl, trj, mj, and j stage, respectively) based on cluster-specific marker gene expressions (Figure S2, Table S1). The Uniform Manifold Approximation and Projection (UMAP) plots displayed the cell type composition at each stage (Figure 1B). And the cell type dynamic changes were observed during metamorphosis. For example, notochord cells disappeared at trj and mj stages, while endostyle and digestive tract cells appeared at mj and j stages (Figure 1B). To compare the cell type consistency, the cell type similarity heat map of adjacent stages was acquired based on the gene expression profiles. The results showed a higher one-to-one consistency between hsl and trl stage, and between trl and trj stage (Figure 1C). The integrated UMAP of sc-RNA seq data also revealed that the cell types from stages before metamorphosis (hsl, trl and trj stages) were distinctly separated from the ones after metamorphosis (mj and j stages) (Figure 1D).

### Increased proportion of endoderm-derived cells after *S. clava* metamorphosis

To gain a global single-cell atlas, we integrated the data of all stages and obtained a total of 72 cell-types (Figure 2A). According to the origin of germ layers, we divided them into ectoderm derived cells (including epidermal cells, central nervous system (CNS), peripheral nervous system (PNS), gamma-aminobutyric acid cells, nerve cord cells, sensory cells and photoreceptor cells), mesoderm derived cells (including mesenchyme cells, trunk lateral cells (TLCs), trunk ventral cells (TVCs), notochord cells, muscle cells (body wall muscle, heart muscle), blood cells, and germ cells) and endoderm derived cells (including endostyle cells, digestive cells, and pharynx cells). The expressions of germ layer restricted maker genes showed that endodermal makers, such as *Nkx2-1*, *Nkx2-6*, *A1CF* and *THA* were expressed in endoderm-derived cell types, and mesodermal markers (*Noto8*, *FGF18*, *KDM8*, *TPM*, *TNNC*) and ectodermal markers (*Ces*, *NRCAM*, *SCO*, *SSPO*) were also expressed in mesoderm and ectoderm derived cell-types, respectively (Figure 2B, Figure S2A, Table S1), which validated the cell type annotation and classification. In order to display the cell dynamics of each germ layer during metamorphosis, the cells were labeled based on their germ layer origins before (hsl, trl and trj stages) and after metamorphosis (mj and j stages) on UAMP (Figure 2C). The result not only showed a distinct cell composition but also an increasing of proportion of endodermal derived cells after metamorphosis. The ectodermal and mesodermal derived cell types occupied the majority before metamorphosis, while after metamorphosis the proportion of endodermal derived cells were increased from trj to j stage (Figure 2D).

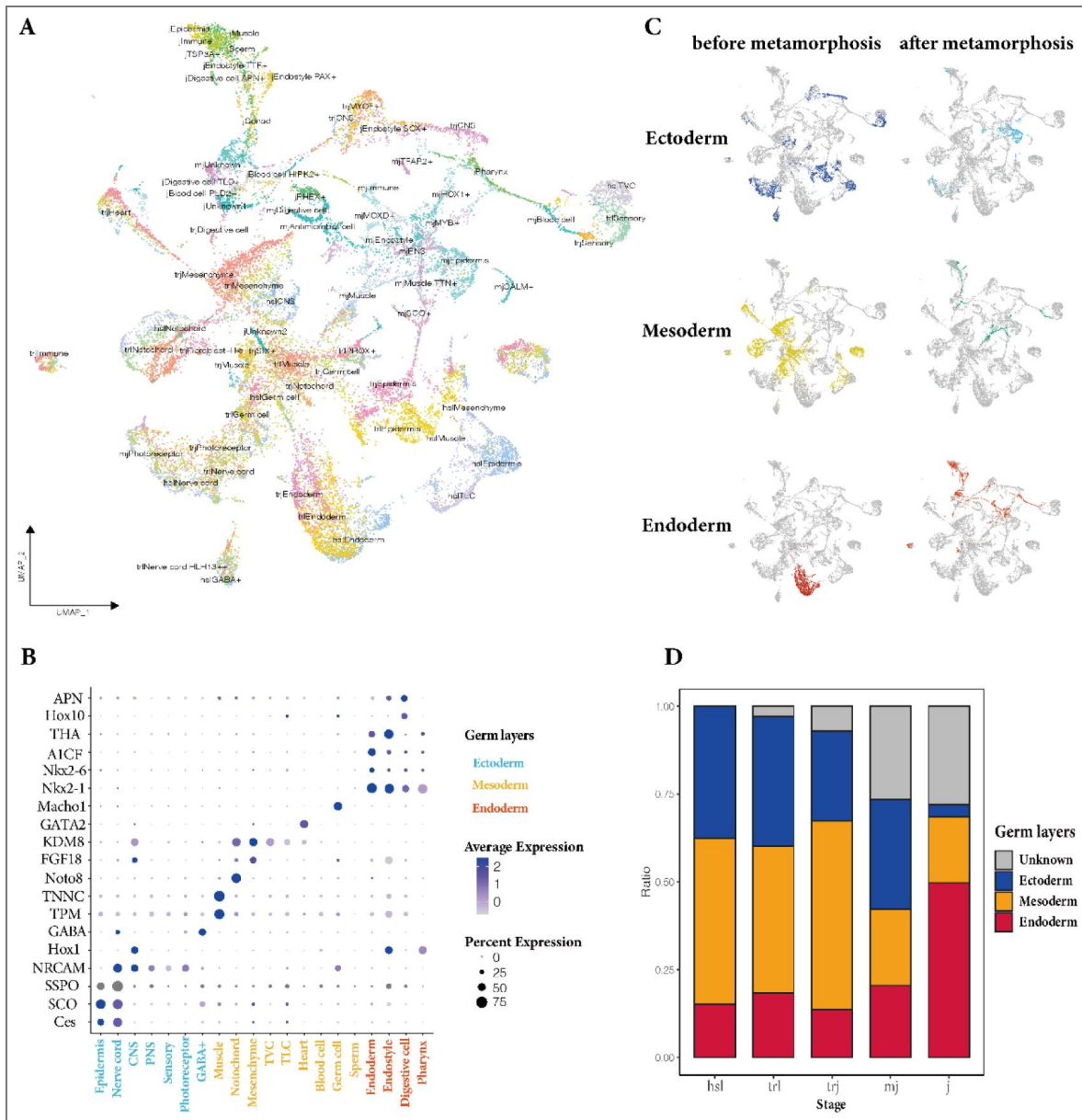
### Two distinct endodermal populations with different differentiation potential

To investigate the differentiation of endoderm-derived cells during metamorphosis, we selected 1,815 cells from trj stage (before metamorphosis) and mj stage (after metamorphosis) with expression of endoderm-specific marker genes (Figure S2B). The sub-cluster analysis identified 12 cell types, including four cell types of endoderm cells from trj stage, and three cell types of stomach progenitors, three cell types of intestine progenitors, endostyle cells, immune cells from mj stage, representing the endodermal cell lineage during metamorphosis of *S. clava* (Figure 3A). Then, the differentiation potential levels of each cell type were predicted using CytoTRACE



**Figure 1. Identification of major cell types across five developmental stages in *S. clava*.**

(A) Representative sampling stages and corresponding larval and juvenile morphologies. (B) Cell type annotations derived from scRNA-seq data, aligned with the respective developmental stages. Hpf, hours post-fertilization; Dpf, days post-fertilization. (C) Inferred corresponding relationships between cell types across successive stages. Each cell type at a later stage (y-axis) is matched to its most similar cell type from the previous stage (x-axis) based on transcriptomic similarity calculated via K-nearest neighbor (KNN) analysis. A similarity score > 0.7 is used as the criterion for inferring the corresponding relationship. Cell types are color-coded by their germ layer of origin: ectoderm (blue), mesoderm (yellow), and endoderm (red). (D) UMAP plot integrating scRNA-seq data from all five stages, with cells colored by developmental stage, illustrating the cellular state transitions before and after metamorphosis. Cells from pre-metamorphic stages are outlined by red dashed lines, while post-metamorphic cells are enclosed by green dashed lines.



**Figure 2. Germ layer-based cell type composition and dynamics during metamorphosis of *S. clava*.**

(A) UMAP plot of all 26,099 cells integrated across five developmental stages, showing the distribution of cell types. Cluster labels represent a combination of developmental stage and cell type identity. (B) Dot plot displaying the expression of representative germ layer-specific marker genes. Dot color indicates the average expression level within each cluster, and dot size reflects the proportion of cells expressing the gene. (C) UMAP plot highlighting cells derived from the three germ layers, showing the distribution of ectoderm-, mesoderm-, and endoderm-derived populations before and after metamorphosis. (D) Bar plot illustrating the proportional changes in cell populations derived from each germ layer across the five developmental stages.

analysis, with higher scores indicating less differentiated, progenitor-like cells and lower scores corresponding to highly differentiated, lineage-committed cells (Figure 3B). The results showed that cell types with different differentiation potentials (ranging from high to low) are observed at both stages (trj stage and mj stage) (Figure 3C and 3D). The differentiation potential level of four endoderm cell types at trj stage showed significant differences.

To reveal the difference between them, the gene expression profiles were exhibited (Figure 3E). The trunk endodermal marker genes *Lhx3* and *Nkx2-1* (*TTF1*) were expressed in the high differentiation potential endoderm 1 and endoderm 2. *Foxp4*, *Otx1* expressed in the high differentiation potential endoderm 1 and *Sox7*, *FGF16* expressed in the high differentiation potential endoderm 2. *Hox10*, the marker of the endodermal strand cells in larval tail,<sup>23</sup> was detected in low differentiation potential endoderm 1 and 2. Meanwhile, *Amyloid precursor-like protein* (*APLP*) and *Carboxypeptidase Z* (*CBPZ*) were detected in the low differentiation potential endoderm 1; *CBPE*, *Six1*, *Low-density lipoprotein receptor 1* (*LRP1*) and *LRP5* were expressed in the low differentiation potential endoderm 2 (Table S2). The results indicated that the high differentiation potential endoderm cells derived from endoderm in larval trunk and the low differentiation potential endoderm derived from the endodermal strand in larval tail.

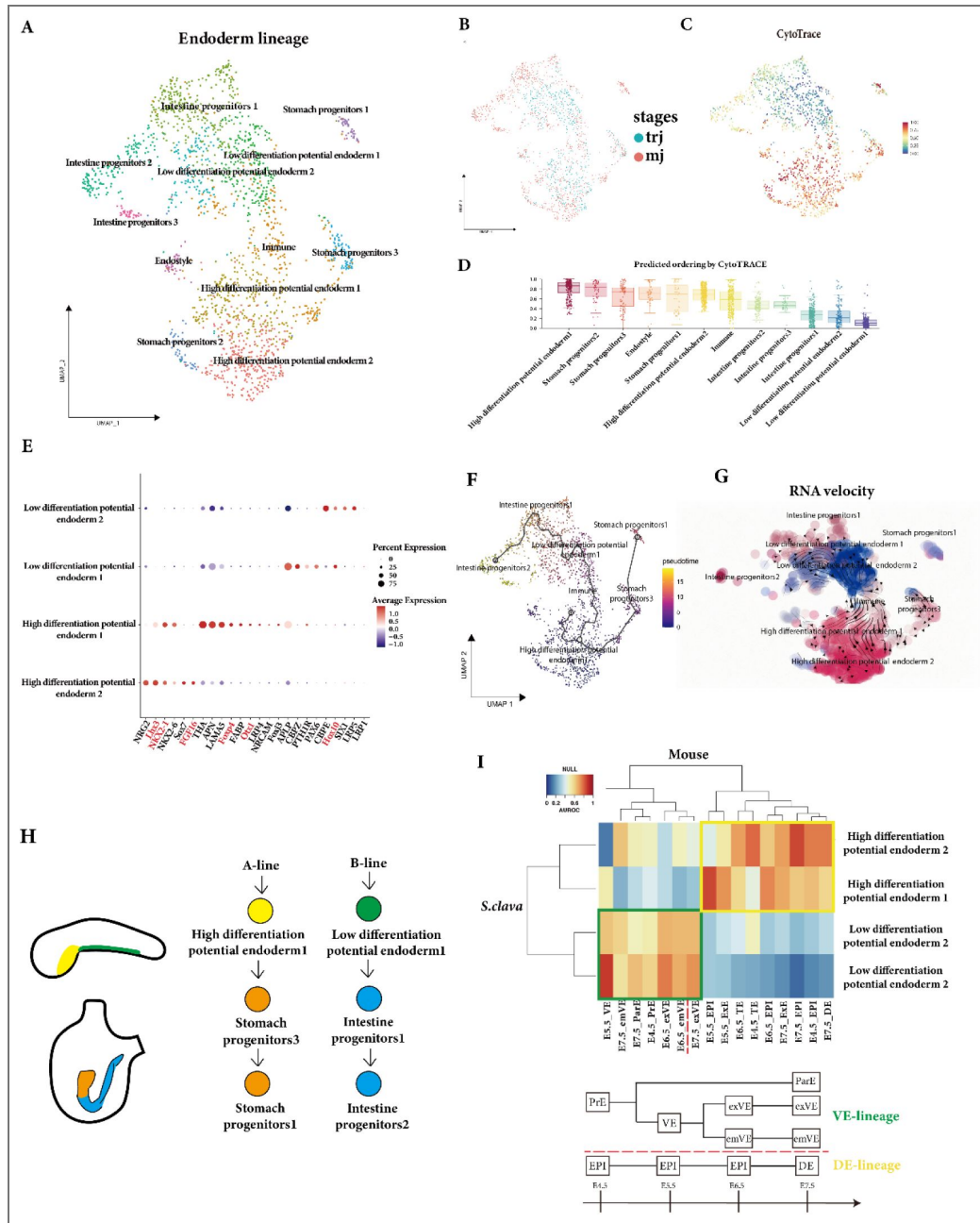
In order to further reveal the cell lineage relationships, pseudotime trajectory of these cells were constructed using Monocle3. We designated the high differentiation potential endoderm 1 population, which exhibited the highest differentiation potential, as root in pseudotime ordering, and ordered all cell types of endoderm lineage. The results showed that one branch developed from high differentiation potential endoderm clusters to the stomach progenitor clusters; and the other one was from the low differentiation potential endoderm clusters to the intestine progenitor clusters (Figure 3F). RNA velocity analysis also indicated that stomach progenitor clusters, endostyle cluster, and immune cluster originated from the high differentiation potential endoderm cells, while intestine progenitor clusters originated from the low differentiation potential endoderm cells (Figure 3G). Combining the endodermal cell lineages and pseudo-time analysis results, we inferred that high differentiation potential endoderm cells are corresponded to trunk endoderm lineage which could generate the stomach progenitors, while low differentiation potential endoderm cells are corresponded to endoderm strand lineage which could generate the intestine progenitors (Figure 3H).

To determine the evolutionary relationships of the endodermal cells between ascidian and vertebrates, we compared ascidian endodermal cells at trj stage with mouse embryonic endoderm cell types from E4.5 to E8.75 stage.<sup>6</sup> The results showed that the ascidian high differentiation potential endoderm cell types 1 and 2 are both corresponded to the mouse definitive endoderm (DE) lineages (including epiblast) and ectoderm lineages (including trophoctoderm and extra-embryonic ectoderm). While the low differentiation potential endoderm cell types 1 and 2 are both correspond to the visceral endoderm (VE) lineages including primitive endoderm, embryonic visceral endoderm, extra-embryonic visceral endoderm and parietal endoderm (Figure 3I).

## The identification of stomach and intestine progenitor cells

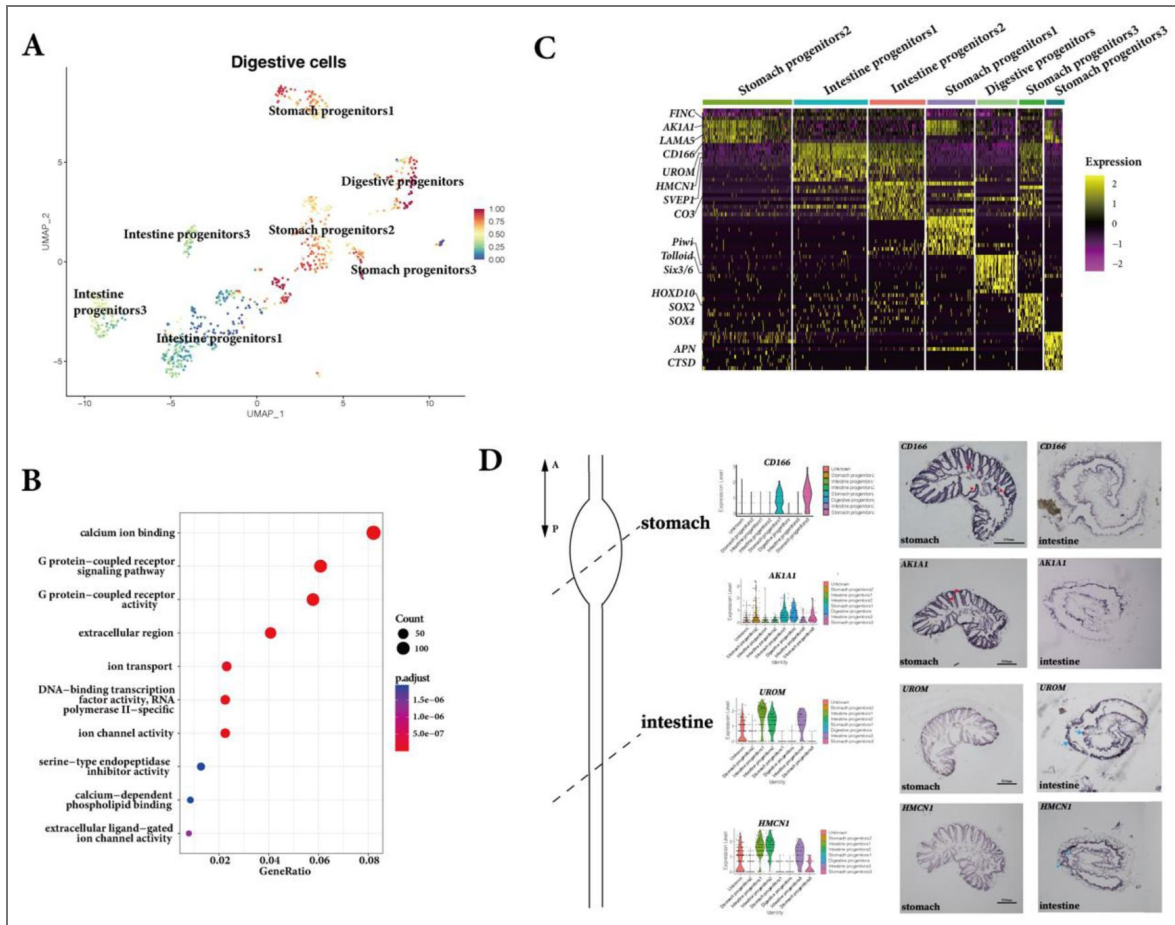
To investigate the potential differentiation signals of stomach and intestine progenitor cells, we selected the digestive tract related cells from mj stage and further divided them into sub-types. Seven sub-types were identified based on marker genes, including three sub-types of stomach progenitors (aminopeptidase N+), three sub-types of intestine progenitors (*Hox10*+), and a group of cells expressing stem-cell markers such as *Piwi*, *Tolloid*, *Six3/6* (labeled as digestive progenitors) (Figure 4A, Table S3). To verify the cell type annotation, we compared our scRNA-seq data with the published bulk RNA-seq data of digestive tract in ascidian *Ciona savignyi*.<sup>24</sup> The results showed that stomach progenitors of *S. clava* had the highest expression similarity with the stomach and anterior intestine of *C. savignyi*, while the intestine progenitors had the highest expression similarity with the middle intestine and posterior intestine (Figure S3).

To identify the biological function of these seven sub-types, the top100 differentially expressed genes (DEGs) were identified by comparing these cells against all other cells at mj stage for Gene Ontology (GO) analysis. The result showed that the genes of digestive cell types were enriched in



**Figure 3. Developmental trajectories of endoderm lineage cells.**

(A) UMAP plot showing sub-clustering of endodermal lineage cells and their distribution during metamorphosis. (B) Box plot summarizing the predicted differentiation levels of endodermal lineage cells based on CytoTRACE scores. (C) UMAP overlaid with CytoTRACE differentiation scores of endodermal lineage cells, where red indicates higher differentiation potential states and blue indicates more differentiated states. (D) UMAP plot of endodermal lineage cells colored by developmental stages of *trj* and *mj*. (E) Dot plot showing expression levels of representative marker genes distinguishing the two major endodermal lineages. Dot color reflects average expression within each cluster, and dot size indicates the proportion expressing cells. (F) Pseudotime trajectory analysis visualized on the UMAP plot. Cells are colored by pseudotime score. (G) RNA velocity streamlines overlaid on the UMAP plot, indicating the predicted directions of cell state transition within the endodermal lineage. (H) Schematic illustration summarizing the two major endodermal lineages and their respective differentiation fates. (I) (Top) Heatmap showing cross-species transcriptomic similarity between ascidian and mouse endodermal lineages. AUROC scores were calculated using expression of one-to-one orthologous genes. (Bottom) Schematic diagram of mouse definitive endoderm and visceral endoderm lineages. PrE, primitive endoderm; ParE, parietal endoderm; VE, visceral endoderm; emVE, embryonic visceral endoderm; exVE, extra-embryonic visceral endoderm; TE, trophoctoderm; ExE, extra-embryonic ectoderm; EPI, epiblast; DE, definitive endoderm.



**Figure 4. Characterization of stomach and intestine progenitor cells.**

(A) UMAP plot showing sub-clustering of digestive tract cells at the metamorphic juvenile stage. Cells are colored according to CytoTRACE-inferred differentiation scores, with red indicating higher differentiation potential. (B) Dot plot showing enriched Gene Ontology (GO) terms for the top 100 differentially expressed genes in digestive cells compared to all other cells at a certain stage. (C) Heatmap displaying the expression of representative marker genes in the major digestive tract cell types. Each column represents a cell type, and each row represents a marker gene. (D) Schematic diagram of the adult digestive tract (stomach and intestine), with dotted lines indicating the anterior-posterior levels of transverse sections (left). Violin plots showing the expression levels of marker genes (CD166 and AK1A1) in stomach progenitor cells and marker genes (UROM and HMCN1) in intestinal progenitor cells (middle). *In situ* hybridization results validating marker gene expression in ascidian digestive tract (right). Red and blue arrows indicate signals in stomach and intestinal regions, respectively.

calcium ion binding, G protein-coupled receptor activity, G protein-coupled receptor signaling pathway, extracellular matrix region, ion transport and DNA-binding transcription factor activity (Figure 4B [↗](#)). The gene expression analysis showed that extracellular matrix and digestive enzymes genes such as *FINC*, *CD166*, *LAMA5*, *AK1A1*, *APN*, *CTSD* were preferentially expressed in stomach progenitors. While, *HPGDs*, *UROM*, *HMCN1*, *SVEP1*, *SOX2*, *Hox10* were expressed in intestine progenitors (Figure 4C [↗](#), Table S3). The expression patterns of these markers were experimentally verified in digestive tract of adult ascidians by *in situ* hybridization. The signals of *CD166* and *AK1A1* (stomach progenitor markers) were mainly expressed in the folds of adult stomach, while *UROM* and *HMCN1* (intestine progenitor markers) showed stronger signals in the inner fold of intestinal epithelium (Figure 4D [↗](#)). The expression pattern of marker genes further supported the contribution of these digestive progenitors to the digestive organs in adult.

## Cell differentiation signals of stomach and intestine progenitors

To investigate the cell differentiation of stomach and intestine lineages, a pseudo-time developmental trajectory of digestive cells at mj stage was constructed using Monocle2. The trajectory showed two distinct directions, with one differentiating into stomach progenitors and the other into intestine progenitors (Figure 5A [↗](#), Figure S4A). The branched gene expression analysis on the distinct directions of stomach and intestine progenitors highlighted the genes on each direction (Figure 5B [↗](#), Table S4). *MASP1*, *ITLN2* and *LARP1* were highly expressed in stomach direction. *KLF7*, *LRP4* and *Notch2* were highly expressed in intestine direction. *FOS* was differently expressed in the two branches (Figure 5B [↗](#)).

To construct the cell-cell communication during the development of the digestive tract, the CellChat analysis of all cell types to digestive-related cells at mj stage were conducted (Figure 5C [↗](#)). The results showed that the interactions mediated by ligands *TGFB2* and *FGF9* present in digestive related cell types, which belong to the *TGF- $\beta$*  signaling pathway and *FGF* signaling pathway respectively (Figure 5D [↗](#)). Furthermore, the expression level of *FGF* signaling was higher in digestive progenitors than in stomach and intestine progenitors, and *TGF- $\beta$*  signaling was higher in digestive progenitors and stomach progenitors than in intestinal progenitors (Figure S4B). Meanwhile, *PSAP-GPR37* and *NRG2-ERBB4* interaction signal were also found in cell-cell interaction of digestive related cells (Figure 5D [↗](#)).

To explore the homologous relationship of the digestive progenitors between ascidian and vertebrate, we compared the expression feature of digestive progenitor cells of ascidian with mouse digestive organ progenitors (Table S5). The stomach progenitors 1, 2 and 3 from ascidian exhibited higher similarity with DE-derived progenitors in mouse including DE cells in thymus, lung, liver, small intestine and colon. The intestine progenitor 1 exhibited higher similarity with both DE and VE progenitors from pancreas in mouse, and the intestine progenitor 2 and 3 exhibited higher similarity with VE-derived progenitors in mouse, including VE cells in thymus, thyroid, lung, pancreas, small intestine and colon (Figure 5E [↗](#) and 5F [↗](#)), which indicated a different cell lineage composition of digestive organs between ascidian and mouse.

## Discussion

In ascidian, there is a unique retrogressive metamorphosis process.<sup>25,26</sup> Due to the unique lifestyle strategy of ascidian, the development processes of endoderm-derived organs are unsynchronized with ectoderm and mesoderm derived organs, which differs from most vertebrates.<sup>27–29</sup> This allows us to better trace the formation of digestive tract. In this study, we present an extensive analysis of single-cell transcriptomic of ascidian *S. clava* from five developmental stages that cover the metamorphosis process.



## Ascidian digestive tract progenitor cells originate from different endodermal cell populations with distinct differentiation potentials

In ascidians, the endodermal cells of the tadpole larva are distributed in the trunk and tail, corresponding to the trunk endoderm and tail endodermal strand, respectively.<sup>28,30</sup> The digestive tract develops during metamorphosis from these two primordial tissues. The endodermal cells undergo epithelialization to form a tubular structure from stage 27 (tail regressing larva stage); and the development of the digestive tract is completed at stage 37 (metamorphic juvenile stage) after body axis rotation.<sup>31</sup> According to previous cell tracing experiment, the anterior and posterior parts of the digestive tract are formed through distinct processes from the posterior endoderm and the endodermal strand, respectively. However, it still has been largely unknown about the characteristics of these primordial cells. Here, we identified high and low differentiation potential endoderm cells at trj stage. The further pseudo-time trajectories analyses inferred that the high differentiation potential endoderm cells generated into the stomach progenitor, and the low differentiation potential endoderm cells generated into the intestine progenitor. These results revealed the different developmental origins of stomach and intestine organs and a significant difference in differentiation potentials between two groups of endoderm cells. These distinct differentiation potentials may have functional consequences for the following organ specialization, which also could help to explain different regions of the digestive tract exhibit distinct cellular compositions and regenerative capacities.<sup>32,33</sup>

## Conserved molecular signaling in digestive tract regionalization between ascidians and vertebrates

During organ formation, cell identity and the tissue morphogenesis must be tightly coordinated. Several signaling pathways, including those of FGF, Wnt, hedgehog, BMP and retinoic acid are used reiteratively during endoderm organ development in vertebrates.<sup>34–37</sup> Among them, FGF signaling plays a particularly pivotal role in the early regionalization of the gut tube, where it is required for establishing distinct domains along the anterior-posterior (A-P) axis.<sup>38</sup> The cell-cell communication analysis of ascidians also identified the FGF signaling among ascidian digestive progenitor cells. This observation suggests evolutionary conservation of FGF-mediated patterning mechanisms in gut development, highlights the conserved molecular but not cellular logic underlying digestive tract regionalization across chordate evolution.

## Dual origins of digestive tract progenitor cells in both ascidians and mammals

Recent study demonstrates both embryonic and extraembryonic cells contribute to the emerging gut endoderm in mouse.<sup>39</sup> Our work also showed dual origins of digestive tract progenitors, with some cell populations corresponded to the mouse DE lineages (embryonic cells), the other cell populations correspond to the VE lineages (extraembryonic cells). The lineage tracing experiments in sea urchin<sup>40,41</sup> also show that digestive tract is derived from the endoderm, but the foregut (together with the oral midgut) and the hindgut (together with the aboral midgut) are derived from different macromere lineages. These indicated conserved dual origins of digestive tract progenitor cells in deuterostomes.

Nevertheless, the cells contribute to the landscape of the specialized endodermal organs seems different between ascidians and mammals. The stomach progenitor cells and intestine progenitor cells in ascidians corresponded to the DE and VE lineages that constitute the digestive organs, rather than corresponding to specific organs in mouse. This suggests that the cell fates of ascidian stomach and intestine progenitors restricted earlier, which is distinct from vertebrates. Moreover, recent study showed the extraembryonic gut cells (VE lineage) of mouse undergo programmed cell

death and neighboring embryonic cells clear their remnants via non-professional phagocytosis, so that VE lineage cells are no longer detected in the wild-type gut.<sup>42</sup> These indicated a different cell fate determination and contribution to the digestive tract between ascidians and mammals.

Overall, the dual lineage origin of the digestive tract appears to be a fundamental feature established during development in deuterostome. The complex digestive tract development observed in vertebrates may represent an extension and elaboration of the cellular heterogeneity already present in the endodermal cells of chordate ancestors.

## Conclusions

This study provided the single-cell atlas during metamorphosis of ascidians, and elucidated the pseudo-trajectories of endoderm differentiation in *S. clava*. Our work revealed that the digestive tract originated from two endodermal progenitor populations with distinct differentiation potentials in ascidians. Comparative analyses with vertebrates support a shared dual origin but divergent differentiation trajectories of digestive tract lineages. Together, these findings provide a cellular framework for understanding digestive tract specialization and the evolutionary diversification of endodermal lineages across chordates.

## Materials and Methods

### Animals and sample collection

Adult *S. clava* were obtained from Xunshan Company (Weihai, China) and temporarily maintained in a recirculating aquarium system at 18 °C in the laboratory, with continuous aeration provided by air pumps. Healthy adults exhibiting normal gamete release were selected. Sperm and eggs were collected separately and mixed for fertilization, which was allowed to proceed for 30 minutes before transferring the zygotes to culture dishes. The culture dishes were subsequently incubated at 18 °C for further development. Embryos were monitored periodically according to the established developmental staging for ascidians.<sup>43</sup> At the desired developmental stages, larvae or juveniles were collected for subsequent experiments. Samples were collected at five representative developmental stages: hatched swimming larva, tail-regressing larva, tail-regressed larva, metamorphic juvenile, and juvenile stages.

### Single-cell isolation and scRNA-seq library construction

Samples from five developmental stages were used for single-cell isolation and library construction. Three biological replicates were prepared. For scRNA-seq library preparation, gel beads in emulsion (GEMs) were generated using the Chromium Controller (10x Genomics). Libraries were constructed using the Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (10x Genomics), following the manufacturer's protocol. The resulting libraries were sequenced on the Illumina NovaSeq platform.

### Data analysis for scRNA-seq

For the scRNA-seq data generated using the 10x Genomics Chromium platform, each of the biological replicate was processed independently. Raw sequencing data were processed using Cell Ranger to generate gene-barcode matrices for each sample, employing the default parameters. A previously published genome assembly was used as the reference.<sup>15</sup> Gene expression matrices were then imported into Seurat v4.0 to create Seurat objects.<sup>44</sup> Standard preprocessing steps including quality filtering, normalization, integration, and clustering were performed. Cells were initially filtered based on expression metrics, retaining those with nCount between 0 and 30,000 and nFeature between 300 and 1,000. Cells exhibiting high mitochondrial gene content (defined as >15% of total counts) were excluded. High-quality datasets were normalized and scaled using SCTransform, followed by dimensionality reduction. The biological replicates were then integrated using the IntegrateData function to remove batch effects and enable joint analysis. The integrated

dataset was subjected to principal component analysis (PCA), and unsupervised clustering was performed using the shared nearest neighbor algorithm with a resolution of 0.5. Marker genes for each cluster were identified using the FindAllMarkers function.

## Cell type annotation

For both the single-stage and integrated single-cell datasets covering five developmental stages, differentially expressed genes (DEGs) were identified using the FindAllMarkers function in Seurat with default parameters. For each dataset, the top 30 DEGs were selected based on fold change and statistical significance (adjusted p-value), and their functional annotations were retrieved from the Swiss-Prot database (<https://www.ebi.ac.uk/uniprot/index>). Given that our dataset spans five distinct stages of ascidian metamorphosis, we performed cell clustering and annotation independently for each stage. Cell types or cluster identities were manually assigned based on the expression of cluster-enriched marker genes in each stage-specific single-cell dataset. Final cell type annotations were consolidated based on the clustering results from the integrated dataset, allowing for consistent comparison across stages.

## GO analysis

GO analysis was performed with the package clusterProfiler4.0.<sup>45</sup> Only biological process GO terms with a *p*-value < 0.05 were selected. The results were further visualized with the ggplot2 package.

## RNA velocity analysis

Loom files containing unspliced and spliced reads were generated from velocity.py for downstream analysis with default parameters. The python package scVelo<sup>46</sup> was applied to estimate RNA velocity with a dynamic model by default parameters. The RNA velocities were projected into UMAP embedding and visualized by cell types.

## Cell trajectory analysis of the endodermal lineage cells and digestive progenitors

Pseudotime analysis of endodermal lineage cells was performed using the Monocle3 package.<sup>47</sup> The gene expression matrix and UMAP coordinates were imported into a CellDataSet object. Cell trajectories were constructed using the learn\_graph function with the parameter minimal\_branch\_len = 3. To infer pseudotime, the order\_cells function was applied, and root cells were manually selected to anchor the trajectory. For digestive progenitor cells, pseudotime analysis was conducted using Monocle2.<sup>47</sup> Genes highly expressed in these clusters ( $\log_{2} \text{FC} > 0.25$ ) were identified using the differentialGeneTest function to detect dynamic gene expression changes along pseudotime. To analyze branch-specific expression patterns, the BEAM function was applied with branch\_point = 1. Genes exhibiting branch-dependent expression were exported, and transcription factors were selectively filtered from this list. The plot\_genes\_branched\_pseudotime function was then used to visualize temporal expression patterns of these TFs along divergent trajectories.

## Cell-cell interaction analysis

Cell-cell communication analysis was performed using the CellChat package.<sup>49</sup> The analysis utilized the CellChatDB.human database, which includes interactions from secreted signaling, extracellular matrix-receptor binding, and direct cell-cell contact ligand-receptor pairs. The resulting Seurat object was converted into a CellChat object and processed according to the recommended workflow in the CellChat manual.

## Comparison across species between ascidian and vertebrates

Single-cell RNA-seq data for mouse endodermal tissues were obtained from the GEO database (accession number: GSE123046). Cross-species comparison was performed based on one-to-one orthologous genes between *S. clava* and mouse. The MetaNeighbor package<sup>50</sup> was used to assess cell-type similarity across species based on gene expression profiles. Sankey diagrams were generated using the networkD3 package (<https://christophergandrud.github.io/networkD3/>) to visualize correspondence between cell clusters in *S. clava* and mouse.

## *In situ* hybridization experiment

The digestive tract of healthy individuals was carefully dissected using sterilized instruments. After isolating the entire digestive tract, the stomach was separated from the intestinal region based on distinct external morphological features. Fresh samples were fixed in 4% paraformaldehyde in artificial seawater at 4 °C overnight. Subsequently, the tissues were cryoprotected by immersion in 30% sucrose in PBS. The tissues were embedded in O.C.T. compound (Sakura 4583) and snap-frozen in liquid nitrogen. Cryosections were prepared using a Leica CM3050 cryostat and mounted on adhesive microscope slides. Digoxigenin (DIG)-labeled RNA probes were designed (Table S6) and synthesized by *in vitro* transcription using SP6 and T7 RNA polymerases (Thermo Fisher, EP0131 and EP0111). *In situ* hybridization (ISH) was performed in a humidified dark chamber to prevent tissue desiccation. Tissue sections were permeabilized with 10 µg/mL Proteinase K for 30 minutes at 37 °C, followed by washing in PBST (0.1% Tween-20 in PBS). Sections were then post-fixed with 4% PFA in PBS for 30 minutes at room temperature, followed by additional PBST washes. For hybridization, tissues were first incubated in pre-hybridization buffer (50% formamide, 5× SSC, 0.1% Tween-20) for 10 minutes at room temperature, and then transferred to hybridization buffer (50% formamide, 5× SSC, 0.1% Tween-20, 5× Denhardt's solution, 100 µg/mL yeast tRNA, 50 µg/mL heparin) at the hybridization temperature for 4 hours. Fresh hybridization buffer containing RNA probes (0.5 ng/µL) was then applied, and the tissues were incubated for 20 hours at the same temperature. Post-hybridization, sections were washed four times in hybridization wash buffer (50% formamide, 5× SSC, 0.1% Tween-20), followed by four washes in PBST. Slides were then incubated overnight at 4 °C in 1% blocking buffer (Roche) containing the anti-DIG antibody. Finally, chromogenic detection was performed using NBT/BCIP substrate.

## Statistics analysis

All the statistical analyses were performed using R (version 3.6). Student's *t*-test and Spearman's rank correlation analysis were utilized in this study.

## Availability of data

All data generated or analyzed during this study are included in the manuscript and supporting files. The RNA-sequencing datasets generated in this study have been deposited at NCBI under the BioProject accession number PRJNA1122959.

## Acknowledgements

The authors thank Professor Sebastian Shimeld (Department of Zoology, University of Oxford) for constructive comments and suggestions on the manuscript.

## Additional information

### Funding

This work was funded by the National Natural Science Foundation of China (Grant No. 32522017, 32370461), the Science & Technology Innovation Project of Laoshan Laboratory (No. LSKJ202203001, LSKJ202203002), the Natural Science Foundation of Shandong Province

(ZR2023MD029), and the Taishan Scholar Program of Shandong Province, China (B.D.).

## Authors' contributions

Conceptualization: B.D., J.W., and H.Y. Project supervision: B.D., and J.W. scRNA-seq library generation: Y.G., W. Z., J.W., Data analyze: Y.G., and P.L. Technical and experimental support: Y.G., J.B., and H.Y.. Writing—original draft: Y.G. Writing—review and editing: J.W., and B.D.

## Funding

Funder	Grant reference number	Author
MOST   National Natural Science Foundation of China (NSFC)	32522017	Jiankai Wei
MOST   National Natural Science Foundation of China (NSFC)	32370461	Jiankai Wei
Science & Technology Innovation Project of Laoshan Laboratory	LSKJ202203002	Bo Dong
山东省科学技术厅   Natural Science Foundation of Shandong Province (山东省自然科学基金)	ZR2023MD029	Haiyan Yu
Taishan Scholar Foundation of Shandong Province (山东省泰山学者特聘教授)		Bo Dong

## Author ORCID iDs

**Bo Dong:** <https://orcid.org/0000-0003-1616-5363>

**Jiankai Wei:** <https://orcid.org/0000-0003-3215-5579>

## Additional files

**Supplementary material.** [↗](#)

**Table S1.** [↗](#) Cell markers for single-cell RNA-seq data sets of five stages.

**Table S2.** [↗](#) Cell markers for single-cell RNA-seq of the endodermal cells at trj stage.

**Table S3.** [↗](#) Cell markers for single-cell RNA-seq of the digestive cells at mj stage.

**Table S4.** [↗](#) Differential expressed genes for three branches on the digestive pseudo-time trajectory.

**Table S5.** [↗](#) The one-to-one homologous genes list between *S. clava* and *Mus musculus*.

**Table S6.** [↗](#) Primers used for in situ hybridization probe synthesis.

## References

1. Nakayama S., Sekiguchi T., Ogasawara M (2019) Molecular and evolutionary aspects of the protochordate digestive system. *Cell Tissue Res* **377**:309-320 <https://doi.org/10.1007/s00441-019-03035-5>
2. Nowotschin S., Hadjantonakis A.-K., Campbell K (2019) The endoderm: a divergent cell lineage with many commonalities. *Development* **146** <https://doi.org/10.1242/dev.150920>
3. Lemaitre B., Miguel-Aliaga I (2013) The Digestive Tract of *Drosophila melanogaster*. *Annu Rev Genet* **47**:377-404 <https://doi.org/10.1146/annurev-genet-111212-133343>
4. Maduro M.F., Rothman J.H (2002) Making worm guts: The gene regulatory network of the *Caenorhabditis elegans* endoderm. *Dev Biol* **246**:68-85 <https://doi.org/10.1006/dbio.2002.0655>
5. Tremblay K.D (2010) Formation of the murine endoderm: lessons from the mouse, frog, fish, and chick. *Prog Mol Biol Transl Sci* **96**:1-34 <https://doi.org/10.1016/b978-0-12-381280-3.00001-4>

6. Nowotschin S., Setty M., Kuo Y.-Y., Liu V., Garg V., Sharma R., Simon C.S., Saiz N., Gardner R., Boutet S.C., *et al.* (2019) The emergent landscape of the mouse gut endoderm at single-cell resolution. *Nature* **569**:361-367 <https://doi.org/10.1038/s41586-019-1127-1>
7. Zorn A.M., Wells J.M (2009) Vertebrate Endoderm Development and Organ Formation. *Annu Rev Cell Dev Biol* **25**:221-251 <https://doi.org/10.1146/annurev.cellbio.042308.113344>
8. Boyle M.J., Yamaguchi E., Seaver E.C (2014) Molecular conservation of metazoan gut formation: evidence from expression of endomesoderm genes in *Capitella teleta* (Annelida). *EvoDevo* **5**:39 <https://doi.org/10.1186/2041-9139-5-39>
9. Lengyel J.A., Iwaki D.D (2002) It takes guts:: The *Drosophila* hindgut as a model system for organogenesis. *Dev Biol* **243**:1-19 <https://doi.org/10.1006/dbio.2002.0577>
10. Martín-Durán J.M., Hejnal A (2015) The study of *Priapulus caudatus* reveals conserved molecular patterning underlying different gut morphogenesis in the Ecdysozoa. *BMC Biol* **13**:29 <https://doi.org/10.1186/s12915-015-0139-z>
11. McGhee J.D (2013) The *Caenorhabditis elegans* intestine. *Wiley Interdiscip Rev Dev Biol* **2**:347-367 <https://doi.org/10.1002/wdev.93>
12. Corbo J.C., Di Gregorio A., Levine M. (2001) The ascidian as a model organism in developmental and evolutionary biology. *Cell* **106**:535-538 [https://doi.org/10.1016/S0092-8674\(01\)00481-0](https://doi.org/10.1016/S0092-8674(01)00481-0)
13. Dehal P., Satou Y., Campbell R.K., Chapman J., Degnan B., De Tomaso A., Davidson B., Di Gregorio A., Gelpke M., Goodstein D.M., *et al.* (2002) The draft genome of *Ciona intestinalis*: Insights into chordate and vertebrate origins. *Science* **298**:2157-2167 <https://doi.org/10.1126/science.1080049>
14. Cao C., Lemaire L.A., Wang W., Yoon P.H., Choi Y.A., Parsons L.R., Matese J.C., Wang W., Levine M., Chen K (2019) Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* **571**:349-354 <https://doi.org/10.1038/s41586-019-1385-y>
15. Wei J., Zhang J., Lu Q., Ren P., Guo X., Wang J., Li X., Chang Y., Duan S., Wang S., *et al.* (2020) Genomic basis of environmental adaptation in the leathery sea squirt (*Styela clava*). *Mol Ecol Resour* **20**:1414-1431 <https://doi.org/10.1111/1755-0998.13209>
16. Jiang A., Han K., Wei J., Su X., Wang R., Zhang W., Liu X., Qiao J., Liu P., Liu Q., *et al.* (2024) Spatially resolved single-cell atlas of ascidian endostyle provides insight into the origin of vertebrate pharyngeal organs. *Sci Adv* **10**:eadi9035 <https://doi.org/10.1126/sciadv.adi9035>
17. Chiba S., Sasaki A., Nakayama A., Takamura K., Satoh N (2004) Development of *Ciona intestinalis* juveniles (through 2nd ascidian stage). *Zool Sci* **21**:285-298 <https://doi.org/10.2108/zsj.21.285>
18. Ermak T.H (1977) Glycogen deposits in pyloric gland of ascidian *Styela clava* (urochordata). *Cell Tissue Res* **176**:47-55 <https://doi.org/10.1007/BF00220343>
19. Gaill F (1980) Glycogen and degeneration in the pyloric gland of dendrodoa-grossularia (ascidiacea, tunicata). *Cell Tissue Res* **208**:197-206 <https://doi.org/10.1007/BF00234870>
20. Nakayama S., Ogasawara M (2017) Compartmentalized expression patterns of pancreatic- and gastric-related genes in the alimentary canal of the ascidian *Ciona intestinalis*: evolutionary insights into the functional regionality of the gastrointestinal tract in Olfactores. *Cell Tissue Res* **370**:113-128 <https://doi.org/10.1007/s00441-017-2627-7>
21. Hayashibe M., Nakayama S., Ogasawara M (2017) Shared hemocyte-and intestine-dominant expression profiles of intelectin genes in ascidian *Ciona intestinalis*: insight into the evolution of the innate immune system in chordates. *Cell Tissue Res* **370**:129-142 <https://doi.org/10.1007/s00441-017-2647-3>
22. Orito W., Ohhira F., Ogasawara M (2015) Gene expression profiles of FABP genes in protochordates, *Ciona intestinalis* and *Branchiostoma belcheri*. *Cell Tissue Res* **362**:331-345 <https://doi.org/10.1007/s00441-015-2198-4>
23. Kawai N., Ogura Y., Ikuta T., Saiga H., Hamada M., Sakuma T., Yamamoto T., Satoh N., Sasakura Y (2015) Hox10-regulated endodermal cell migration is essential for development of the ascidian intestine. *Dev Biol* **403**:43-56 <https://doi.org/10.1016/j.ydbio.2015.03.018>

24. Zhang W., Jiang A., Yu H.Y., Dong B (2023) Comparative Transcriptomic Analysis Reveals the Functionally Segmented Intestine in Tunicate Ascidian. *Int J Mol Sci* **24**:6270 <https://doi.org/10.3390/ijms24076270>
25. Cloney R.A (1982) Ascidian larvae and the events of metamorphosis. *Am Zool* **22**:817-826 <https://doi.org/10.1093/icb/22.4.817>
26. Nakayama-Ishimura A., Chambon J.P., Hofie T., Satoh N., Sasakura Y (2009) Delineating metamorphic pathways in the ascidian *Ciona intestinalis*. *Dev Biol* **326**:357-367 <https://doi.org/10.1016/j.ydbio.2008.11.026>
27. Hirano T., Nishida H (1997) Developmental fates of larval tissues after metamorphosis in ascidian *Halocynthia roretzi* I. Origin of mesodermal tissues of the juvenile. *Dev Biol* **192**:199-210 <https://doi.org/10.1006/dbio.1997.8772>
28. Hirano T., Nishida H (2000) Developmental fates of larval tissues after metamorphosis in the ascidian, *Halocynthia roretzi* II. Origin of endodermal tissues of the juvenile. *Dev Genes Evol* **210**:55-63 <https://doi.org/10.1007/s004270050011>
29. Davidson B., Wallace S.E.S., Howsmon R.A., Swalla B.J (2003) A morphological and genetic characterization of metamorphosis in the ascidian *Boltenia villosa*. *Dev Genes Evol* **213**:601-611 <https://doi.org/10.1007/s00427-003-0363-3>
30. Nishida H (1987) Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme .3. up to the tissue restricted stage. *Dev Biol* **121**:526-541 [https://doi.org/10.1016/0012-1606\(87\)90188-6](https://doi.org/10.1016/0012-1606(87)90188-6)
31. Nakazawa K., Yamazawa T., Moriyama Y., Ogura Y., Kawai N., Sasakura Y., Saiga H (2013) Formation of the digestive tract in *Ciona intestinalis* includes two distinct morphogenic processes between its anterior and posterior parts. *Dev Dyn* **242**:1172-1183 <https://doi.org/10.1002/dvdy.24009>
32. Dzama M.M., Nigmatullina L., Sayols S., Kreim N., Soshnikova N (2017) Distinct populations of embryonic epithelial progenitors generate Lgr5+ intestinal stem cells. *Dev Biol* **432**:258-264 <https://doi.org/10.1016/j.ydbio.2017.10.012>
33. Fawcner-Corbett D., Antanaviciute A., Parikh K., Jagielowicz M., Gerós A.S., Gupta T., Ashley N., Khamis D., Fowler D., Morrissey E., et al. (2021) Spatiotemporal analysis of human intestinal development at single-cell resolution. *Cell* **184**:810-826.e823. <https://doi.org/10.1016/j.cell.2020.12.016>
34. McLin V.A., Rankin S.A., Zorn A.M (2007) Repression of Wnt/ $\beta$ -catenin signaling in the anterior endoderm is essential for liver and pancreas development. *Development* **134**:2207-2217 <https://doi.org/10.1242/dev.001230>
35. van den Brink G.R. (2007) Hedgehog signaling in development and homeostasis of the gastrointestinal tract. *Physiol Rev* **87**:1343-1375 <https://doi.org/10.1152/physrev.00054.2006>
36. Tiso N., Filippi A., Pauls S., Bortolussi M., Argenton F (2002) BMP signalling regulates anteroposterior endoderm patterning in zebrafish. *Mech Dev* **118**:29-37 [https://doi.org/10.1016/s0925-4773\(02\)00252-6](https://doi.org/10.1016/s0925-4773(02)00252-6)
37. Wang Z.X., Dollé P., Cardoso W.V., Niederreither K (2006) Retinoic acid regulates morphogenesis and patterning of posterior foregut derivatives. *Dev Biol* **297**:433-445 <https://doi.org/10.1016/j.ydbio.2006.05.019>
38. Dessimoz J., Opoka R., Kordich J.J., Grapin-Botton A., Wells J.M (2006) FGF signaling is necessary for establishing gut tube domains along the anterior-posterior axis in vivo. *Mech Dev* **123**:42-55 <https://doi.org/10.1016/j.mod.2005.10.001>
39. Nowotschin S., Hadjantonakis A.K (2020) Guts and gastrulation: Emergence and convergence of endoderm in the mouse embryo. *Cur Top Dev Bio* **36**:429-454 <https://doi.org/10.1016/bs.ctdb.2019.11.012>
40. Logan C.Y., McClay D.R (1997) The allocation of early blastomeres to the ectoderm and endoderm is variable in the sea urchin embryo. *Development* **124**:2213-2223 <https://doi.org/10.1242/dev.124.11.2213>

41. Ransick A., Davidson E.H (1998) Late specification of Veg1 lineages to endodermal fate in the sea urchin embryo. *Dev Biol* **195**:38-48 <https://doi.org/10.1006/dbio.1997.8814>
42. Batki J., Hetzel S., Schifferl D., Bolondi A., Walther M., Wittler L., Grosswendt S., Herrmann B.G., Meissner A (2024) Extraembryonic gut endoderm cells undergo programmed cell death during development. *Nat Cell Biol* **26**:868-877 <https://doi.org/10.1038/s41556-024-01431-w>
43. Lin B.-Y., Shi W.J., Lu Q.-X., Shito T.T., Yu H.-Y., Dong B (2023) Establishment of a developmental atlas and transgenic tools in the ascidian *Styela clava*. *Mar Life Sci Tech* <https://doi.org/10.1007/s42995-023-00200-2>
44. Hao Y.H., Hao S., Andersen-Nissen E., Mauck W.M., Zheng S.W., Butler A., Lee M.J., Wilk A.J., Darby C., Zager M., et al. (2021) Integrated analysis of multimodal single-cell data. *Cell* **184**:3573-3587 <https://doi.org/10.1016/j.cell.2021.04.048>
45. Wu T.Z., Hu E.Q., Xu S.B., Chen M.J., Guo P.F., Dai Z.H., Feng T.Z., Zhou L., Tang W.L., Zhan L., et al. (2021) clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2**:100141 <https://doi.org/10.1016/j.xinn.2021.100141>
46. La Manno G., Soldatov R., Zeisel A., Braun E., Hochgerner H., Petukhov V., Lidschreiber K., Kastrioti M.E., Lönnberg P., Furlan A., et al. (2018) RNA velocity of single cells. *Nature* **560**:494-498 <https://doi.org/10.1038/s41586-018-0414-6>
47. Cao J.-Y., Spielmann M., Qiu X.J., Huang X.F., Ibrahim D.M., Hill A.J., Zhang F., Mundlos S., Christiansen L., Steemers F.J., et al. (2019) The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**:496-502 <https://doi.org/10.1038/s41586-019-0969-x>
48. Qiu X.J., Hill A., Packer J., Lin D.J., Ma Y.A., Trapnell C (2017) Single-cell mRNA quantification and differential analysis with Census. *Nat Methods* **14**:309-315 <https://doi.org/10.1038/nmeth.4150>
49. Jin S., Guerrero-Juarez C.F., Zhang L., Chang I., Ramos R., Kuan C.-H., Myung P., Plikus M.V., Nie Q (2021) Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**:1088 <https://doi.org/10.1038/s41467-021-21246-9>
50. Crow M., Paul A., Ballouz S., Huang Z.J., Gillis J (2018) Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat Commun* **9**:12 <https://doi.org/10.1038/s41467-018-03282-0>

## Peer reviews

### Reviewer #1 (Public review):

Summary:

The authors employ state-of-the-art single-cell sequencing technologies to map the gene expression profiles of the developing digestive tract in the ascidian *Styela clava*, a member of the invertebrate sister group to vertebrates. This data has the potential to provide a new perspective on the relationships between the guts of an invertebrate like this ascidian relative to vertebrate systems. Key findings include the elaboration of our understanding that the *Styela* gut arises from two distinct cellular origins, with this being comparable to the dual embryogenic origin of vertebrate guts (at least, as exemplified by the mouse digestive tract arising from both definitive and visceral endoderm).

Strengths:

The resolution that can be achieved from the series of developmental stages analysed by the authors through the metamorphosis and early gut specification and development is vital to the strength of this new dataset. This new scRNAseq data is likely to provide a useful foundation for future work that delves into the functions of various genes within regions of the ascidian gut.

### Weaknesses:

The main weakness of the manuscript as it currently stands is the lack of clarity about the genetic comparisons between ascidian and mouse, and what the precise genetic underpinnings are for any statements of similarity.

<https://doi.org/10.7554/eLife.110546.1.sa1>

### Reviewer #2 (Public review):

This manuscript explores endodermal lineage specification during metamorphosis in *Styela clava*. As biphasic lifestyle organisms, the endoderm exists as a rudiment in the non-feeding larvae that differentiates throughout metamorphosis to build the digestive components of the adult body plan. The authors of this manuscript use scRNA sequencing of individuals throughout the metamorphic process, as well as maturing juveniles, to follow the trajectories of the endodermal precursors. They identify two distinct populations that give rise to the stomach and intestinal lineages, and they suggest that there are homologous relationships between tunicate & vertebrate dual-origin endodermal lineages. Additionally, the authors highlight the role of conserved FGF signal-dependent programs in digestive organ patterning and suggest that endodermal fate restriction occurs earlier in *Styela* in comparison with the mouse gut.

Overall, the paper is the first in-depth look at tunicate endodermal fate from a single-cell sequencing perspective and provides a robust framework for understanding the evolutionary origins of the deuterostome/chordate gut. The data is substantial and of great interest. However, we find their discussion of evolutionary implications to be highly problematic, and there are also numerous major issues regarding the clarity and cogency of their data presentation. Thus, we consider that substantial revision is required to provide a more accurate analysis of this data and its evolutionary implications. This revision would not require further experimentation.

<https://doi.org/10.7554/eLife.110546.1.sa0>

### Author response:

We sincerely thank the Reviewing Editor, Senior Editor, and both reviewers for their careful and constructive assessment of our manuscript. We are encouraged that the reviewers recognize the value of our dataset and its potential contribution. We greatly appreciate the thoughtful comments and have carefully considered the reviews. We plan to revise the manuscript accordingly.

First, we will revise and refine the cross-species comparative analysis, with particular attention to clarifying the basis of the comparisons between ascidian and mouse endodermal lineages. In particular, we will adopt a more cautious and precise comparative framework, clarify the scope and limitations of the mouse comparison, and broaden the context by incorporating additional vertebrate and invertebrate deuterostome systems where relevant.

Second, we will strengthen the gene-level interpretation of the identified endodermal populations and clarify the molecular basis for the similarities and differences. In particular, we will more clearly identify the key marker genes defining each population, better explain their relationship to previously described developmental sources.

Third, we will improve the clarity of the Results presentation, including the description of the two major endodermal progenitor populations and their subcategories, as well as the organization of the text, figures, and figure legends.

Fourth, we will substantially rewrite the Discussion, especially the sections dealing with evolutionary implications, to ensure that our interpretations are presented in a more cautious manner.

These revisions are intended to address the reviewers' concerns regarding both the evolutionary framing and the presentation of the data. We believe that these revisions, which will include both rewriting and additional analyses, will improve the clarity and rigor of the manuscript. We look forward to submitting a revised version.

We thank the editors and reviewers again for their time and expertise.

<https://doi.org/10.7554/eLife.110546.1.sa3>