scientific reports

OPEN

Transcriptomic profile comparison reveals conservation of ionocytes across multiple organs

Carla Pou Casellas[1,2], Cayetano Pleguezuelos-Manzano[1], Maarten B. Rookmaaker[2], Marianne C. Verhaar[2] & Hans Clevers[1,3]✉

Single-cell RNA sequencing has recently led to the identification of a flurry of rare, new cell types, such as the CFTR-high ionocytes in the airway epithelium. Ionocytes appear to be specifically responsible for fluid osmolarity and pH regulation. Similar cells exist in multiple other organs and have received various names, including *intercalated cell* in the kidney, *mitochondria-rich cell* in the inner ear, *clear cell* in the epididymis, and *ionocyte* in the salivary gland. Here, we compare the previously published transcriptomic profile of cells expressing *FOXI1*, the signature transcription factor expressed in airway ionocytes. Such *FOXI1+* cells were found in datasets representing human and/or murine kidney, airway, epididymis, thymus, skin, inner ear, salivary gland, and prostate. This allowed us to assess the similarities between these cells and identify the core transcriptomic signature of this ionocyte 'family'. Our results demonstrate that, across all these organs, ionocytes maintain the expression of a characteristic set of genes, including *FOXI1*, *KRT7*, and *ATP6V1B1*. We conclude that the ionocyte signature defines a class of closely related cell types across multiple mammalian organs.

Regulation of ion osmolality in epithelial surfaces is a key process for the correct functioning of many epithelial organs, being central to the maintenance of luminal fluid volume, mucus viscosity and regulation of pH. Different types of ion- and pH-regulating cells have been described in vertebrates. They were first reported in eel gill epithelium, and subsequently identified both in gills of fresh and seawater fish, as well as in *Xenopus* epidermis[1–3]. More recently, and despite the inconsistent terminology, ion-secreting cells have been described in several mouse and human organs using single-cell RNA sequencing (scRNA-seq) approaches. In the kidney, intercalated cells regulate urine pH in the collecting duct[4]. In the airway, ionocytes are suggested to regulate osmolarity, pH and viscosity of the airway surface liquid and mucus[5,6]. In the inner ear, mitochondria-rich cells (MRC) regulate the endolymphatic electron potential required for hearing[7]. In the epididymis, clear cell-mediated luminal acidification contributes to sperm maturation and fertility[8,9]. Furthermore, other ion-secreting populations have been recently described in the thymus, and in salivary and sweat glands[10–12]. For clarity, we hereinafter refer to all these mammalian ion-secreting cells as ionocytes.

Because of their central role to the functioning of many different organs, ionocytes have been implicated in the onset of various diseases. As a clear example, mutations in the chloride channel *CFTR*, an ionocyte marker, give rise to cystic fibrosis (CF). CF is characterized by a desiccated mucus layer in the airway, male infertility, salt imbalance in sweat, and sometimes, kidney disorders[13–16]. As another example, malfunction of other ionocyte-related genes, namely vacuolar ATPases, can cause distal renal tubular acidosis accompanied by hearing loss[17]. This highlights the functional (dis)similarities of ionocytes across different organs.

Despite the large number of scRNA-seq datasets available to date, a systematic comparison of the development and function of ionocytes across different mouse and human organs has not been reported to our knowledge. In this report, we compile and compare scRNA-seq datasets from human thymus, airway, kidney, epididymis and skin, and mouse prostate, epididymis, kidney, inner ear, airway and salivary gland. This analysis allows us to define a core ionocyte gene signature and to establish differences that are specific to ionocytes between tissues. In this way, the detailed characterization of ionocytes across several human and mouse tissues contributes to a better comprehension of this cell type and provides a framework to better understand ionocyte functional dysregulation, its involvement in disease pathophysiology and its potential as diagnostic biomarker (Fig. 1).

[1]Oncode Institute, Hubrecht Institute, Royal Dutch Academy of Science (KNAW) and University Medical Center Utrecht (UMCU), Utrecht, The Netherlands. [2]Department of Nephrology and Hypertension, University Medical Center Utrecht (UMCU), Utrecht, The Netherlands. [3]Present address: Roche Pharmaceutical Research and Early Development, 4058 Basel, Switzerland. ✉email: h.clevers@hubrecht.eu
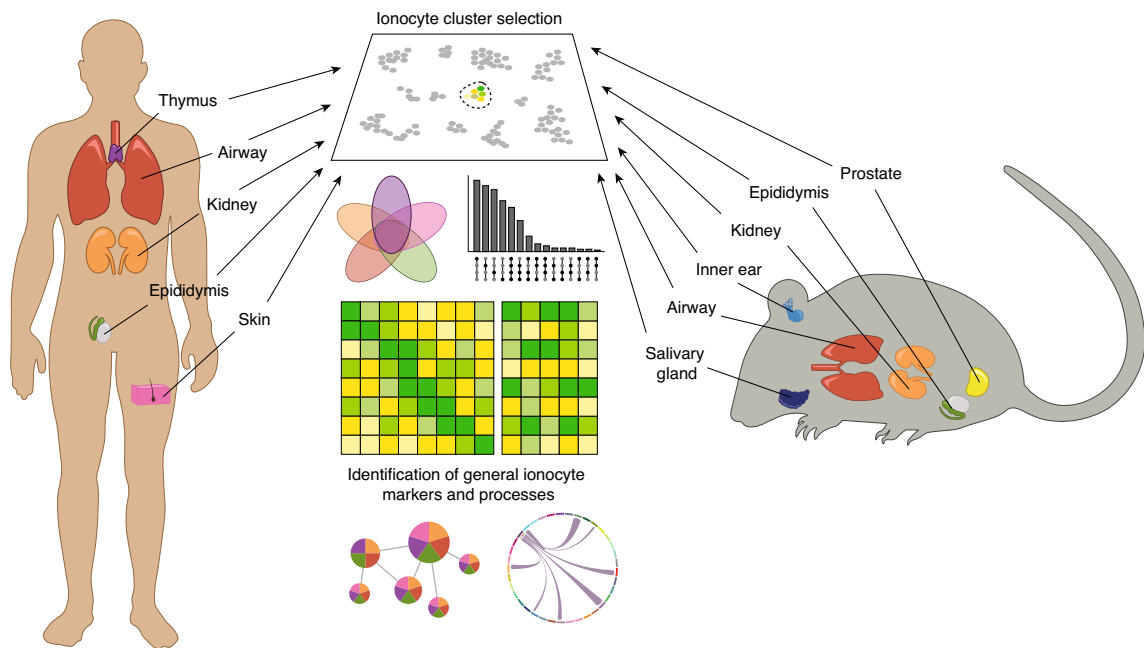
nature portfolio

Figure 1. Strategy applied in this study to identify the core ionocyte gene signature in humans and mice.

Methods

Data processing and clustering. The human and murine scRNA-seq datasets used in this study can be found in Table 1. To ensure the highest data quality and comparability, we exclusively included datasets of healthy adult (or at least postnatal) subjects/animals and containing cells expressing forkhead box I1 (*FOXI1*). In line with this, from the dataset by Hauser et al.[18] we only used data from P30 and adult mice in our study.

All raw data files were processed and analyzed using the Seurat package (v.4.1.0)[19] in R (v.4.0.3). For all datasets, data filtering was performed equally, requiring a minimum of 5 cells per transcript, at least 300 transcripts per cell, and less than 5% expression of mitochondrial genes per cell. After filtering and normalization using the *NormalizeData* command, individual datasets or datasets belonging to the same organ were integrated using either Seurat's canonical correlation analysis (CCA) or reciprocal principal component analysis (PCA) methods. Upon subsequent data scaling, PCA and Uniform Manifold Approximation and Projection (UMAP) dimensionality reductions were run to create UMAP visualization plots. Clustering was performed using the functions *FindNeighbors* and *FindClusters*, and the ionocyte cluster(s) in each organ dataset were identified based on the highest expression of *FOXI1*.

Species	Tissue of Origin	Study ID	References
Human	Kidney	GSE151302	Muto et al.[26]
	Airway	GSE102580	Plasschaert et al.[27]
		GSE166766	Ravindra et al.[28]
		syn21041850	Travaglini et al. [29]
	Epididymis	GSE148963	Leir et al.[30]
	Thymus	GSE147520	Bautista et al.[10]
	Skin	GSE147424	He et al.[11]
Mouse	Kidney	GSE107585	Park et al.[31]
		GSE129798	Ransick et al.[32]
	Airway	GSE103354	Montoro et al.[5]
		GSE102580	Plasschaert et al.[27]
	Epididymis	GSE159713	Shi et al.[33]
		GSE145443	Rinaldi et al. [34]
	Inner ear	GSE87293	Honda et al.[35]
	Salivary gland	GSE150327	Hauser et al.[18]
	Prostate	GSE146811	Karthaus et al.[36]

Table 1. List of single-cell RNA sequencing datasets used in this study.

Marker differential expression and biological process analysis. The top 25 differentially expressed genes (DEGs) per ionocyte population were generated using Seurat's *FindMarkers* function selecting a minimum log2 fold-change of 0.25. The top 20 significantly upregulated biological processes based on these DEGs were computed using the clusterProfiler package (v.3.18.1)[20] and rendered using the packages enrichplot (v.1.10.2)[21] and DOSE (v.3.16.0)[22].

Common feature (gene intersection) analysis. Venn diagrams and gene intersection bar plots were created using the positive DEG lists per ionocyte population with the R packages VennDiagram (v.1.7.1)[23] and UpSetR (v.1.4.0)[24], respectively.

Integration of all human and murine datasets and correlation analysis of human data. All individual human datasets were integrated together by reciprocal PCA, using a k.weight of 80. Afterwards, the integrated object's data was scaled, dimensionality reductions were run, and clustering was computed. Differentially expressed markers in each cluster were calculated using a log fold-change threshold of 0.5. The same applies for the integration of all murine datasets. Correlation of average expression was subsequently performed comparing the different identities of the object (i.e., each organ-specific ionocyte subset) using the *cor* function.

Intercellular communication analysis. Cell–cell communication analysis was performed using the R package CellChat (v.1.1.3)[25]. The same workflow was applied for each ionocyte population, which were subsetted from the main integrated dataset. The computed communications were filtered based on a minimum of 10 cells.

Results

Ionocyte populations can be identified in at least five different human organs based on *FOXI1* expression.

To explore and compare the specific markers in ionocytes derived from different human tissues, we analyzed publicly available healthy tissue scRNA-seq datasets in which the presence of ionocytes was expected. To enable a precise identification of ionocytes in each dataset, we selected data files in which expression of the main ionocyte transcription factor, *FOXI1*[37], could be measured. This resulted in the acquisition of seven datasets, profiling a total of five human organs: kidneys, airway, epididymis, thymus, and skin. The kidney dataset consisted of 26,933 cells, among which 2064 (7.88%) were identified as *FOXI1+* intercalated cells. As expected, some of the top markers differentially expressed in these cells were the vacuolar ATPase subunit C2 (*ATP6V1C2*), the bicarbonate transporter *SLC4A9*, and the adhesion G protein-coupled receptor F5 (*ADGRF5*) (Fig. 2A), which are all well-known markers for this cell type.

Given that airway ionocytes have been so scarcely sequenced, we performed our analyses using three datasets in order to obtain a sufficiently large number of cells. The three datasets combined yielded a total of 91,226 cells, including 472 ionocytes (0.73%) defined by the expression of *FOXI1*. The top differentially expressed markers in this small population included the Achaete-Scute Family BHLH Transcription Factor 3 (*ASCL3*), insulin-like growth factor 1 (*IGF1*), and the integral membrane protein 2C (*ITM2C*) (Fig. 2B).

The employed epididymis dataset contained 5,299 cells, among which 68 (0.72%) were recognized as ionocytes, expressing high levels of markers such as the chloride voltage-gated channel Kb (*CLCNKB*), the vacuolar ATPase subunit G3 (*ATP6V1G3*), and keratin 7 (*KRT7*) (Fig. 2C). As a representation of the thymus, we found a dataset comprising 72,382 cells, and the lowest percentage of *FOXI1+* ionocytes (0.22%, 83 cells) among the tissues studied. Thymus ionocytes differentially expressed markers including the E74-like ETS transcription factor 3 (*ELF3*), claudin 3 (*CLDN3*), and *CD24* (Fig. 2D). Finally, we used a skin dataset comprising a total of 20,385 cells and 236 (1.93%) of ionocytes. This ionocyte population expressed the highest levels of aquaporin 5 (*AQP5*), the ATPase Na$^+$/K$^+$ transporting subunit beta 1 (*ATP1B1*), and carbonic anhydrase 6 (*CA6*) in human skin, among other markers (Fig. 2E). The clusters in each dataset can be found in Supplementary Figs. 1A–E.

Ionocytes display a conserved function and transcriptomic profile across different human organs.

Upon the preliminary identification of ionocytes in each human tissue dataset, we sought out to investigate how comparable these cells are in terms of predicted function and gene expression. Using a gene set enrichment analysis (GSEA) based on biological processes, we observed that all five ionocyte populations share similar predicted processes or functions, mainly those concerning transmembrane transport and pH regulation (Fig. 3A). As for their transcriptomic profile, we noted that the average expression of all genes was highly correlated (Pearson coefficient > 0.5) across ionocyte populations derived from thymus, airway, epididymis, and kidney. In contrast, the average gene expression in skin *FOXI+* cells correlated poorly with those of the other four organs (Fig. 3B).

To compare the human ionocyte populations in more detail, we integrated all five whole-tissue datasets and observed a distinct ionocyte cluster expressing *FOXI1*, composed of cells emerging from the five different organs (Fig. 3C, D). All clusters of the resulting integrated UMAP can be seen in Supplementary Fig. 1F, together with an overview of the markers expressed per cluster, each cluster's identity, and the cluster composition based on tissue of origin (Supplementary Fig. 1G). After computing the list of upregulated genes in each ionocyte population, we identified 30 markers shared between all of them (Fig. 3E). Notably, each ionocyte population shared at least 40% of their differentially expressed genes with other ionocyte populations (Fig. 3F), confirming the transcriptomic similarity between these populations. The 30 markers commonly shared among all ionocytes included several transcriptional regulators (e.g., calmodulin binding transcription activator 1 (*CAMTA1*), *FOXI1*, peroxisome proliferator-activated receptor gamma coactivator 1 alpha (*PPARGC1A*), and transcription factor CP2-like 1 (*TFCP2L1*)); membrane ATPases such as *ATP6V0A4*, *ATP6V0B*, and *ATP6V1B1*; signal transduction-related genes (e.g., guanine nucleotide binding protein alpha (*GNAS*), and regulator of calcineurin 2 (*RCAN2*)), and
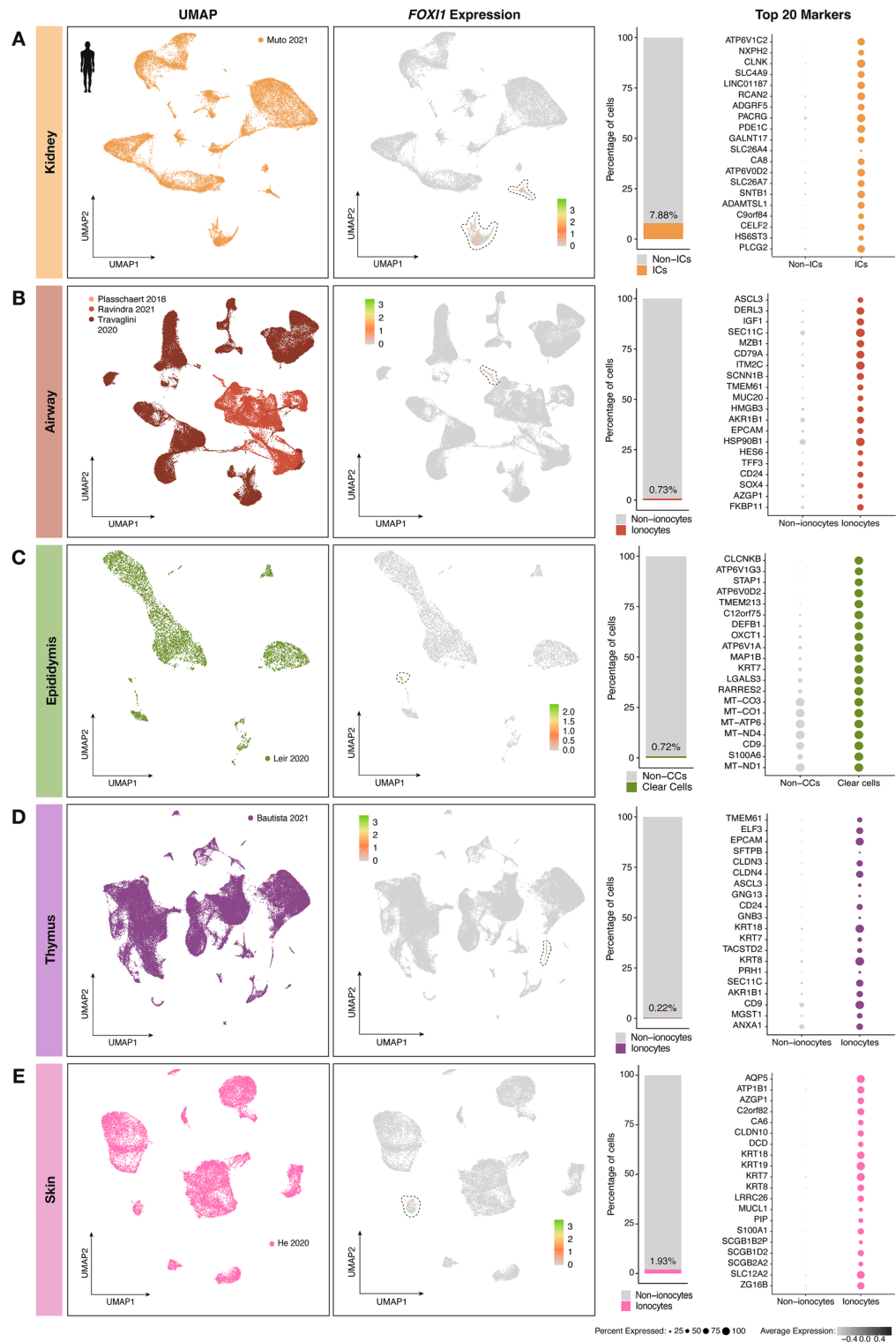
Figure 2. Ionocytes can be identified in several human tissues by having the highest *FOXI1* expression. Depicted are the Uniform Manifold Approximation and Projection (UMAP) plots per organ (**A**: kidney, **B**: airway, **C**: epididymis, **D**: thymus, **E**: skin), together with feature plots showing where each *FOXI1*-expressing cluster is located, and the percentage of cells found in these ionocyte clusters. Finally, the top 20 differentially expressed markers per ionocyte population can be observed.

Figure 3. Human ionocytes derived from kidney, airway, epididymis, thymus, and skin are functionally and transcriptionally very similar. (**A**) Enrichment map depicting the most highly enriched biological processes based on the differentially expressed genes (DEGs) in each ionocyte population. (**B**) Correlation heatmap of average gene expression between ionocytes found in thymus, airway, epididymis, kidney, or skin. The correlation coefficient indicates the results of a Pearson's correlation test. (**C**) Uniform Manifold Approximation and Projection (UMAP) plot outlining the integration of five whole-tissue human datasets. The traced cell cluster indicates the ionocyte cluster, and the bar graph shows the number of cells comprising this cluster derived from each organ. (**D**) Feature plot showing relative *FOXI1* expression in the ionocyte cluster. (**E**) Venn diagram illustrating the shared DEGs between the five ionocyte populations. (**F**) Bar graph indicating the percentage of overlapping genes between each ionocyte population and the rest. (**G**) Heatmaps depicting the relative expression levels of the 30 shared DEGs between ionocytes and non-ionocytes per tissue of origin. (**H**) Violin plots of exemplar DEGs shared between kidney, epididymis, thymus, and airway ionocytes.

surface molecules (e.g., *CA12, KIT*, and *KRT7*) (Fig. 3G). Given that *CFTR* is a well-known ionocyte marker in airways and kidney[6] but was not found in the list of shared markers among the five populations, we explored which ionocyte populations shared a significant differential expression of this gene. We found that *CFTR* differential expression was shared between kidney, epididymis, thymus, and airway ionocytes, along with other important markers such as *CLCNKA*, doublesex and mab-3 related transcription factor 2 (*DMRT2*), forkhead box P1 (*FOXP1*), and phospholipase C gamma 2 (*PLCG2*) (Fig. 3H). The complete list of overlapping genes between each ionocyte population can be found in Supplementary File 1.

Human ionocyte interactions and signaling are preserved among different tissues. To further investigate the general ionocyte functions, we computed both the incoming and outgoing signaling pathways in ionocytes derived from the five organs in question: kidney, airway, epididymis, thymus, and skin. A summary of all involved pathways detected in each population can be observed in Fig. 4A. Signaling pathways that were both incoming and outgoing in ionocytes, mostly referring to autocrine pathways, included e-cadherin (CDH1) and myelin protein zero (MPZ) signaling. In terms of prediction of intercellular communication, we observed that the main outgoing signaling pathways from ionocytes included those capable of affecting endothelium and mesenchyme, for instance, the semaphorin 3 (SEMA3) and ephrin-A (EPHA) pathways. These findings were also reported in a recent publication where salivary glands were studied[12]. Other major pathways found to be signaled by ionocytes revolve around immune system regulation. Two clear examples are the macrophage migration inhibitory factor (MIF) and CD46 signaling. MIF is a pro-inflammatory cytokine that greatly regulates innate immunity[38], while CD46 is thought to regulate T cell-driven inflammation[39]. The putative ability of ionocytes to modulate the immune system was exemplified in an in vivo study that showed the chemoattraction power of neutrophils by kidney intercalated cells[40]. Finally, we observed that ionocytes globally receive signals from other cells through collagen, midkine (MK), pleiotrophin (PTN), and fibronectin (FN1) signaling pathways, among others. Notably, both MK and PTN are heparin-binding growth factors[41], indicating that heparin might be an important regulator of ionocyte function or differentiation. Among the incoming signaling pathways we also found insulin-like growth factor (IGF), which has been found to trigger salt secretion in fish[42]. Surprisingly, while several studies have reported that ionocyte development and cell state is dependent on Notch signaling[27,43], our analysis did not find Notch to be an important signaling pathway in ionocytes (Fig. 4A).

To get an indication of the kind of receiver and source cell types that interact with ionocytes, we generated a communication probability heatmap showing examples of main pathways found in all ionocytes with a clear outgoing or incoming trend. Figure 4B depicts an illustrative heatmap showing the cellular communicators that kidney-derived ionocytes have. Figure 4C, D show examples of two of these pathways, namely SEMA3 and EPHA, in ionocyte populations derived from other organs. As observed, ionocytes derived from different human organs have similar inferred signal-receiving cell targets.

Exploration of murine datasets reveals the presence of ionocytes in other organs. Since human datasets were not available for all desired organs meeting our inclusion criteria, we opted for repeating the analysis using murine datasets. Thus, we could also include healthy inner ear, salivary gland, and prostate datasets. To ensure the highest data quality and representability in our study, we integrated two different datasets per organ of interest. However, for inner ear, salivary gland, and prostate we could only find one available dataset. This resulted in the inclusion of two murine kidney datasets, two airway datasets, two epididymis datasets, one inner ear, one salivary gland, and one prostate dataset. All UMAPs and computed clusters for each separate organ can be found in Supplementary Figs. 2A–F. Just as for human, ionocyte clusters in each murine organ could be identified based on the highest expression of *Foxi1* (Fig. 5A–F). Observing the top 20 markers present in each ionocyte population, it was readily apparent that many of these overlapped (e.g., *Tfcp2l1*, *Hepacam2*, and *Cd24a*) (Fig. 5A–F). To further explore the overlapping genes in ionocytes derived from these six different organs, we created a matrix layout showing all intersecting DEGs (Fig. 5G). Remarkably, we found that murine ionocytes from kidney, epididymis, inner ear, prostate, airway, and salivary gland share the differential (positive) expression of 113 genes. These markers included many membrane ATPases (e.g., *Atp5h*, *Atp6ap1*, and *Atp6v1e1*) (Fig. 5H), and mitochondrial genes (e.g., *Acss1*, *Cox8a*, and *Mdh1*) (Fig. 5I); transcription factors including those found in the human data analysis (i.e., *Foxi1*, *Ppargc1a*, and *Tfcp2l1*); a variety of surface molecules (e.g., *Cldn7*, *Itpr2*, and *Kit*) and signal transduction-related markers (e.g., *Met*, *Ralgapa2*, and *Stap1*); and other markers such as dicarbonyl and l-xylulose reductase (*Dcxr*) and PDZ and LIM domain 3 (*Pdlim3*) (Fig. 5J). The full list of overlapping genes between each murine ionocyte population can be found in Supplementary File 2.

Using the DEGs computed per ionocyte population, we also evaluated the most enriched biological processes and found a clear convergence between ionocytes from the six organs studied. The most prominent processes were those involved in ion transmembrane transport, which we similarly observed using human data (Supplementary Fig. 2G).

Finally, when assessing predicted signaling from and to murine ionocytes, we observed that the main pathways signaled by ionocyte populations included amyloid precursor protein (APP), MIF, vascular endothelial growth factor (VEGF), and SEMA3 signaling, which were also found relevant in human ionocytes. As for incoming signaling, we found that KIT, PTN, FN1, and MK signaling remains relevant in both species, among others (Supplementary Fig. 3).

Discussion

Most individual members of the ionocyte cell type 'family' described in this study have only recently been discovered. Consequently, there is still limited understanding of their transcriptional profile and biological function. Last year, by comparing the transcriptome of human epithelial cells in nasal, bronchial, and epididymal samples,
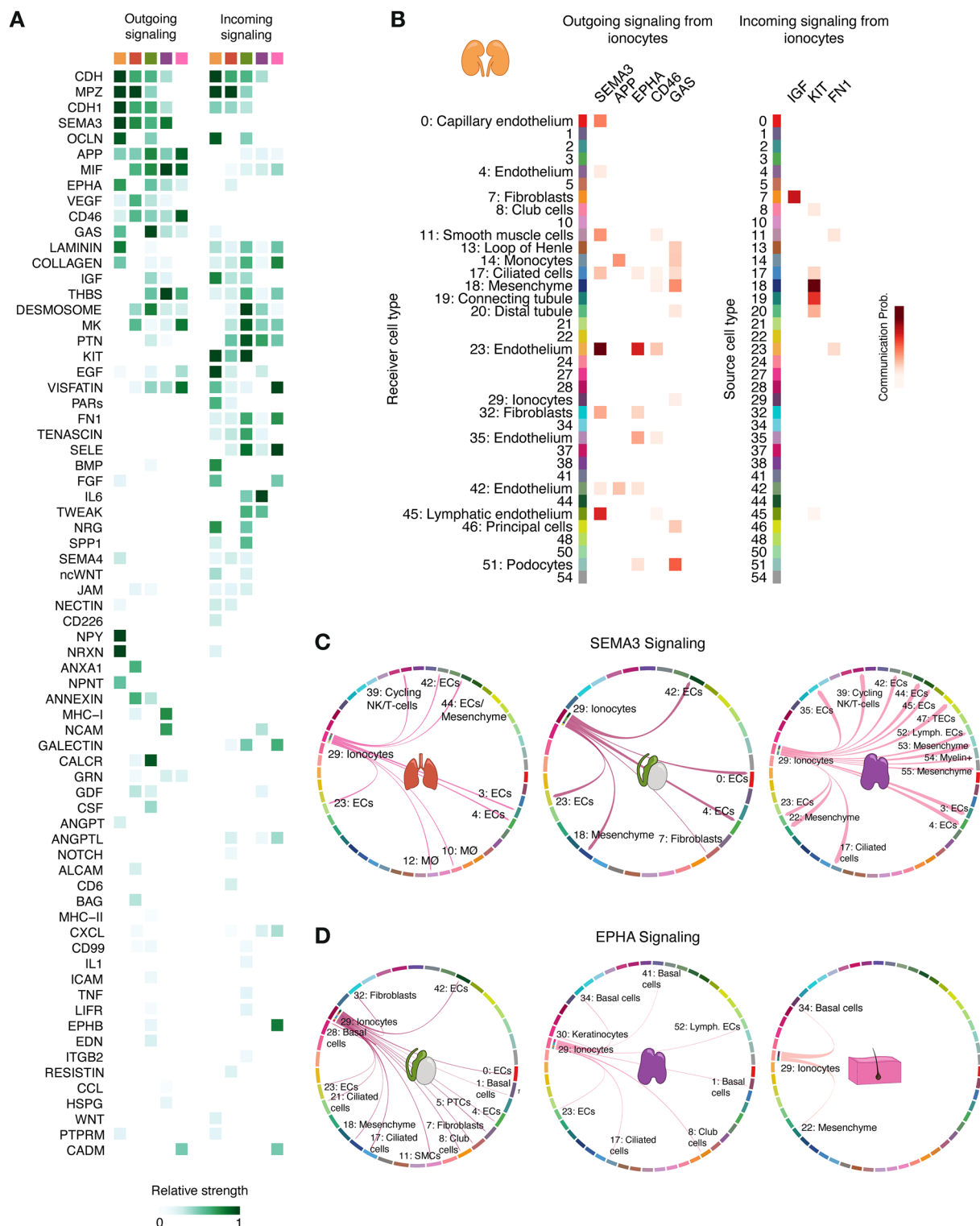
Figure 4. Main outgoing and incoming signaling in human ionocytes. (**A**) Heatmap showing the list of pathways identified as relevant outgoing or incoming signals in human ionocytes derived from different organs (orange: kidney, red: airway, green: epididymis, purple: thymus, pink: skin). (**B**) Heatmap depicting examples of probable communications between kidney ionocytes and other cell types through different pathways. Examples of communications between ionocyte populations from other organs and other cell types based on the signaling pathways semaphorin 3 (SEMA3) (**C**) and ephrin-A (EPHA) (**D**) can be seen as chord diagrams.
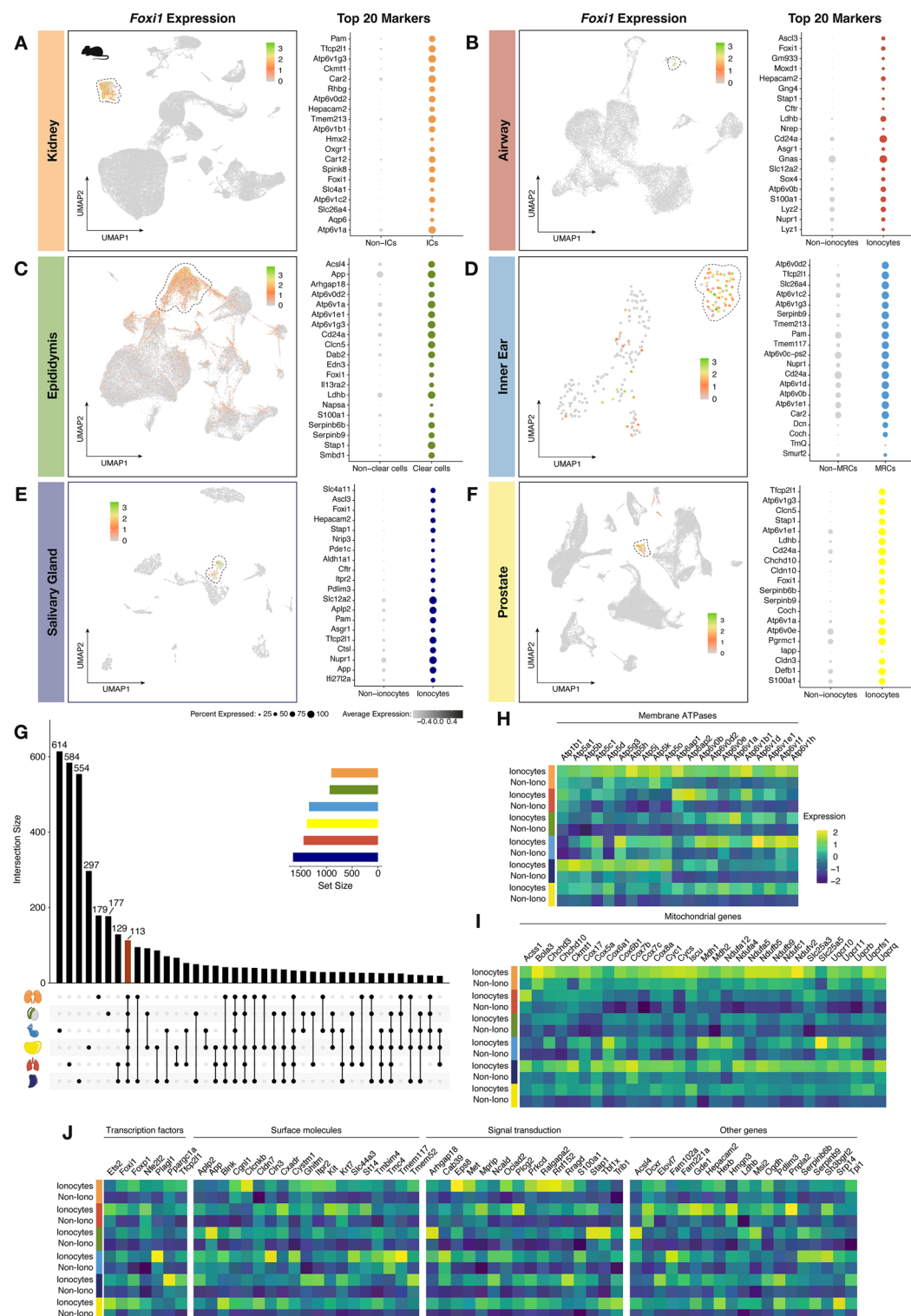
Figure 5. Murine ionocyte identification and comparison. Depicted are feature plots from datasets of murine tissues (**A**: kidney, **B**: airway, **C**: epididymis, **D**: inner ear, **E**: salivary gland, **F**: prostate), showing the presence of ionocyte clusters (*Foxi1*-expressing cells). Next to each feature plot, we show the top 20 differentially expressed genes (DEGs) per ionocyte population. (**G**) Bar graph revealing the intersecting DEGs between ionocyte populations. The 113 genes commonly shared in all five ionocyte populations can be observed in the heatmaps, which were divided by category (**H**, membrane ATPases; **I**, mitochondrial genes; **J**, transcription factors, surface molecules, signal transduction-related genes, and other genes).

Paranjapye et al. described conservation of ionocytes between these tissues[44]. However, a complementing cross comparison of these cell types between other organs has not previously been given. In this study, we report the presence of cells related to airway ionocytes across different human (i.e., kidneys, airways, epididymis, thymus, and skin) and murine (i.e., kidneys, airways, epididymis, inner ear, salivary gland, and prostate) organs using publicly available datasets. By comparing the transcriptomic profile of ionocytes from different organs, we observed a considerable overlap of specific DEGs (30 markers in human ionocytes and 113 in murine ionocytes). It should be noted, however, that the lists of genes characteristic of ionocytes might not be restricted to the ones found in this study. Considering technical matters, markers might have been omitted or underrated in the analysis due to several factors, including different sequencing depths between datasets, a different barcoding procedure, or the predefined p-value cutoff for DEG computation. This means that some markers, albeit not sufficiently differentially expressed, could still be relevant ionocyte signature markers. The presence of exclusive markers in the respective ionocyte populations might be the consequence of different environmental niche conditions that each population is exposed to (e.g., oxygenation, flow, or neighboring cells).

By further exploring the DEGs found in both human and murine ionocytes, we found that the main biological process enriched in ionocytes is that of ion transmembrane transporter activity, which is in accordance with their function of regulating intracellular and extracellular pH and osmolarity in aquatic vertebrate species[45–47]. Since ion transmembrane transport and regulation of acid–base homeostasis are crucial throughout the body, the presence of ionocytes is likely not limited to the organs included in our analysis. Last year, our group and others identified a population of CFTR*high*-expressing cells in the human duodenum. This rare cell population is characterized by high expression levels of *BEST4*, *GUCA2A*, and *OTOP2*, among others[48,49]. However, after identifying core ionocyte markers in the current study, we believe that these cells do not belong to the family of ionocytes as they do not express any other characteristic marker. According to the Human Protein Atlas, breast glandular cells and some cells in the thyroid express *FOXI1*[50]. Future studies should investigate whether the cells express more markers identified here.

Finally, we computed the most dominant incoming and outgoing signaling pathways in ionocytes, which were largely congruent between humans and mice (e.g., APP, SEMA3 and MK signaling), confirming a high transcriptomic conservation. Exploring ionocyte transcriptional characterization in both humans and mice enabled us to unravel the core ionocyte signature conserved in both species. Such signature included the transcription factors *FOXI1*, *PPARGC1A*, and *TFCP2L1*; the membrane ATPases *ATP6V0B*, *ATP6V1A*, *ATP6V1B1*, and *ATP6V1H*; and the surface molecules encoded by *ITPR2*, *KIT*, *KRT7*, and *ST14*.

In a recent study, Morris proposed that cell identity can be defined using three pillars: phenotype, lineage, and state[51]. Based on phenotype, which includes transcriptomic profile and functionality, it could be defended that the ionocytes identified in different organs share the same cell identity. However, the diverse developmental origins of this cell type family are an interesting observation: the kidneys and epididymis are derived from the mesodermal lineage[52,53]; the airways, thymus, and prostate are endoderm-derived[54]; and the epithelial cells of the inner ear[55], sweat glands of the skin[56], and the major salivary glands[57] arise from the ectoderm germ layer. This could indicate that a convergent development has taken place, in which distinct progenitor populations give rise to the same (or a very similar) cell type.

In summary, we have identified a core signature for ionocytes, an apparent key cell type in regulating fluid ion homeostasis and pH, throughout the body. This finding can provide a better understanding or even facilitate the discovery of multi-organ disease states. One example of such an ionocyte-related disease is that caused by mutations in *ATP6V1B1*, in which patients present with severe renal tubular acidosis often accompanied by hearing loss[58,59]. A similar phenotype can be observed in patients with mutations in *ATP6V0A4*[17,60], another core ionocyte marker, and non-surprisingly, in patients carrying *FOXI1* mutations[61]. With the recognition of ionocytes in multiple mammalian organs, additional symptoms in patients with ionocyte-related diseases might be found in previously unsuspected organs.

Data availability
The datasets used in this study can be found in Table 1, and can be accessed either through Gene Expression Omnibus (GEO) or Synapse.

References
1. Dymowska, A. K., Hwang, P. P. & Goss, G. G. Structure and function of ionocytes in the freshwater fish gill. *Respir. Physiol. Neurobiol.* **184**(3), 282–292 (2012).
2. Esaki, M. *et al.* Mechanism of development of ionocytes rich in vacuolar-type H(+)-ATPase in the skin of zebrafish larvae. *Dev. Biol.* **329**(1), 116–129 (2009).
3. Quigley, I. K., Stubbs, J. L. & Kintner, C. Specification of ion transport cells in the Xenopus larval skin. *Development* **138**(4), 705–714 (2011).
4. Purkerson, J. M. *et al.* Adaptation to metabolic acidosis and its recovery are associated with changes in anion exchanger distribution and expression in the cortical collecting duct. *Kidney Int.* **78**(10), 993–1005 (2010).
5. Montoro, D. T. *et al.* A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* **560**(7718), 319–324 (2018).
6. Shah, V. S. *et al.* Cystic fibrosis and the cells of the airway epithelium: What are ionocytes and what do they do?. *Annu. Rev. Pathol.* **17**, 23–46 (2022).
7. Mori, N. *et al.* Ion transport its regulation in the endolymphatic sac: Suggestions for clinical aspects of Meniere's disease. *Eur. Arch. Otorhinolaryngol.* **274**(4), 1813–1820 (2017).
8. Blomqvist, S. R. *et al.* Epididymal expression of the forkhead transcription factor Foxi1 is required for male fertility. *EMBO J.* **25**(17), 4131–4141 (2006).
9. Breton, S. *et al.* Acidification of the male reproductive tract by a proton pumping (H+)-ATPase. *Nat. Med.* **2**(4), 470–472 (1996).

10. Bautista, J. L. *et al.* Single-cell transcriptional profiling of human thymic stroma uncovers novel cellular heterogeneity in the thymic medulla. *Nat. Commun.* **12**(1), 1096 (2021).
11. He, H. *et al.* Single-cell transcriptome analysis of human skin identifies novel fibroblast subpopulation and enrichment of immune subsets in atopic dermatitis. *J. Allergy Clin. Immunol.* **145**(6), 1615–1628 (2020).
12. Mauduit, O. *et al.* A mesenchymal to epithelial switch in Fgf10 expression specifies an evolutionary-conserved population of ionocytes in salivary glands. *Cell Rep.* **39**(2), 110663 (2022).
13. Ahmad, A., Ahmed, A. & Patrizio, P. Cystic fibrosis and fertility. *Curr. Opin. Obstet. Gynecol.* **25**(3), 167–172 (2013).
14. Brown, M. B., McCarty, N. A. & Millard-Stafford, M. High-sweat Na+ in cystic fibrosis and healthy individuals does not diminish thirst during exercise in the heat. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **301**(4), R1177–R1185 (2011).
15. Kreda, S. M., Davis, C. W. & Rose, M. C. CFTR, mucins, and mucus obstruction in cystic fibrosis. *Cold Spring Harb. Perspect. Med.* **2**(9), a009589 (2012).
16. Shahin, W. *et al.* Early renal involvement in children with cystic fibrosis. *Eur. Respir. J.* **52**(62), PA4627 (2018).
17. Stover, E. H. *et al.* Novel ATP6V1B1 and ATP6V0A4 mutations in autosomal recessive distal renal tubular acidosis with new evidence for hearing loss. *J. Med. Genet.* **39**(11), 796–803 (2002).
18. Hauser, B. R. *et al.* Generation of a single-cell RNAseq atlas of murine salivary gland development. *iScience* **23**(12), 101838 (2020).
19. Hao, Y. H. S. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**(13), 3573-3587.e29 (2021).
20. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**(3), 100141 (2021).
21. Yu, G. *enrichplot: Visualization of Functional Enrichment Result. R package version 1.18.3.* https://yulab-smu.top/biomedical-knowledge-mining-book/ (2022).
22. Yu, G. *et al.* DOSE: An R/Bioconductor package for Disease Ontology Semantic and Enrichment analysis. *Bioinformatics* **31**(4), 608–609 (2015).
23. Chen, H. B. & Paul, C. VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinform.* **12**, 35 (2011).
24. Conway, J. R. L. A. & Gehlenborg, N. UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**(18), 2938–2940 (2017).
25. Jin, S.G.-J.C.F. *et al.* Inference and analysis of cell-cell communication using Cell Chat. *Nat. Commun.* **12**, 1088 (2021).
26. Muto, Y. *et al.* Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. *Nat. Commun.* **12**(1), 2190 (2021).
27. Plasschaert, L. W. *et al.* A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **560**(7718), 377–381 (2018).
28. Ravindra, N. G. *et al.* Single-cell longitudinal analysis of SARS-CoV-2 infection in human airway epithelium identifies target cells, alterations in gene expression, and cell state changes. *PLoS Biol.* **19**(3), e3001143 (2021).
29. Travaglini, K. J. *et al.* A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**(7835), 619–625 (2020).
30. Leir, S. H. *et al.* An atlas of human proximal epididymis reveals cell-specific functions and distinct roles for CFTR. *Life Sci. Alliance* **3**, 11 (2020).
31. Park, J. *et al.* Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**(6390), 758–763 (2018).
32. Ransick, A. *et al.* Single-cell profiling reveals sex, lineage, and regional diversity in the mouse kidney. *Dev. Cell* **51**(3), 399–413 (2019).
33. Shi, J. *et al.* Spatio-temporal landscape of mouse epididymal cells and specific mitochondria-rich segments defined by large-scale single-cell RNA-seq. *Cell Discov.* **7**(1), 34 (2021).
34. Rinaldi, V. D. *et al.* An atlas of cell types in the mouse epididymis and vas deferens. *Elife* **2020**, 9 (2020).
35. Honda, K. *et al.* Molecular architecture underlying fluid absorption by the developing inner ear. *Elife* **2017**, 6 (2017).
36. Karthaus, W. R. *et al.* Regenerative potential of prostate luminal cells revealed by single-cell analysis. *Science* **368**(6490), 497–505 (2020).
37. Vidarsson, H. *et al.* The forkhead transcription factor Foxi1 is a master regulator of vacuolar H-ATPase proton pump subunits in the inner ear, kidney and epididymis. *PLoS ONE* **4**(2), e4471 (2009).
38. Calandra, T. & Roger, T. Macrophage migration inhibitory factor: A regulator of innate immunity. *Nat. Rev. Immunol.* **3**(10), 791–800 (2003).
39. Arbore, G. *et al.* Complement receptor CD46 co-stimulates optimal human CD8(+) T cell effector function via fatty acid metabolism. *Nat. Commun.* **9**(1), 4186 (2018).
40. Azroyan, A. *et al.* Renal intercalated cells sense and mediate inflammation via the P2Y14 receptor. *PLoS ONE* **10**(3), e0121419 (2015).
41. Xu, C. *et al.* Functional receptors and intracellular signal pathways of midkine (MK) and pleiotrophin (PTN). *Biol. Pharm. Bull.* **37**(4), 511–520 (2014).
42. Yan, J. J. *et al.* Insulin-like growth factor 1 triggers salt secretion machinery in fish under acute salinity stress. *J. Endocrinol.* **246**(3), 277–288 (2020).
43. Mukherjee, M. *et al.* Endogenous notch signaling in adult kidneys maintains segment-specific epithelial cell types of the distal tubules and collecting ducts to ensure water homeostasis. *J. Am. Soc. Nephrol.* **30**(1), 110–126 (2019).
44. Paranjapye, A. *et al.* Cell function and identity revealed by comparative scRNA-seq analysis in human nasal, bronchial and epididymis epithelia. *Eur. J. Cell Biol.* **101**(3), 151231. https://doi.org/10.1016/j.ejcb.2022.151231 (2022).
45. Fridman, S. Ontogeny of the osmoregulatory capacity of teleosts and the role of ionocytes. *Front. Mar. Sci.* **2020**, 7 (2020).
46. Shir-Mohammadi, K. & Perry, S. F. Expression of ion transport genes in ionocytes isolated from larval zebrafish (Danio rerio) exposed to acidic or Na(+)-deficient water. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **319**(4), R412–R427 (2020).
47. Yew, H. M., Zimmer, A. M. & Perry, S. F. Assessing intracellular pH regulation in H(+)-ATPase-rich ionocytes in zebrafish larvae using in vivo ratiometric imaging. *J. Exp. Biol.* **223**, 5 (2020).
48. Busslinger, G. A. *et al.* Human gastrointestinal epithelia of the esophagus, stomach, and duodenum resolved at single-cell resolution. *Cell Rep.* **34**(10), 108819 (2021).
49. Elmentaite, R. *et al.* Cells of the human intestinal tract mapped across space and time. *Nature* **597**(7875), 250–255 (2021).
50. *Human Protein Atlas* (accessed 1 Jun 2022); http://www.proteinatlas.org (2022).
51. Morris, S. A. The evolving concept of cell identity in the single cell era. *Development* **146**, 12 (2019).
52. Joseph, A., Yao, H. & Hinton, B. T. Development and morphogenesis of the Wolffian/epididymal duct, more twists and turns. *Dev. Biol.* **325**(1), 6–14 (2009).
53. Mugford, J. W. *et al.* Osr1 expression demarcates a multi-potent population of intermediate mesoderm that undergoes progressive restriction to an Osr1-dependent nephron progenitor compartment within the mammalian kidney. *Dev. Biol.* **324**(1), 88–98 (2008).
54. Hogan, B. L. M. & Zaret, K. S. 15—development of the endoderm and its tissue derivatives. In *Mouse Development* (eds Rossant, J. & Tam, P. P. L.) 301–330 (Academic Press, 2002).
55. van der Valk, W. H. *et al.* Building inner ears: Recent advances and future challenges for in vitro organoid systems. *Cell Death Differ.* **28**(1), 24–34 (2021).
56. Cui, C. Y. & Schlessinger, D. Eccrine sweat gland development and sweat secretion. *Exp. Dermatol.* **24**(9), 644–650 (2015).

57. Rothova, M. *et al.* Lineage tracing of the endoderm during oral development. *Dev. Dyn.* **241**(7), 1183–1191 (2012).
58. Karet, F. E. *et al.* Mutations in the gene encoding B1 subunit of H+-ATPase cause renal tubular acidosis with sensorineural deafness. *Nat. Genet.* **21**(1), 84–90 (1999).
59. Mohebbi, N. *et al.* Homozygous and compound heterozygous mutations in the ATP6V1B1 gene in patients with renal tubular acidosis and sensorineural hearing loss. *Clin. Genet.* **83**(3), 274–278 (2013).
60. Vargas-Poussou, R. *et al.* Genetic investigation of autosomal recessive distal renal tubular acidosis: Evidence for early sensorineural hearing loss associated with mutations in the ATP6V0A4 gene. *J. Am. Soc. Nephrol.* **17**(5), 1437–1443 (2006).
61. Enerback, S. *et al.* Acidosis and deafness in patients with recessive mutations in FOXI1. *J. Am. Soc. Nephrol.* **29**(3), 1041–1048 (2018).

Author contributions
C.P.C. gathered the data and performed the analyses; C.P.C. and C.P.M. conceptualized and wrote the manuscript; M.B.R., M.C.V., and H.C. reviewed and edited the manuscript.

Competing interests
HC is inventor of several patents related to organoid technology and currently an employee of Roche, Basel. HC's full disclosure is given at https://www.uu.nl/staff/JCClevers/. The other authors declare no conflicts of interest.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-30603-1.

Correspondence and requests for materials should be addressed to H.C.

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

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