scientific reports

Check for updates

OPEN

A temporal single cell transcriptome atlas of zebrafish anterior segment development

Oliver Vöcking & J. K. Famulski✉

Anterior segment dysgenesis (ASD), resulting in vision impairment, stems from maldevelopment of anterior segment (AS) tissues. Incidence of ASD has been linked to malfunction of periocular mesenchyme cells (POM). POM cells specify into anterior segment mesenchyme (ASM) cells which colonize and produce AS tissues. In this study we uncover ASM developmental trajectories associated with formation of the AS. Using a transgenic line of zebrafish that fluorescently labels the ASM throughout development, Tg[*foxc1b*:GFP], we isolated GFP+ASM cells at several developmental timepoints (48–144 hpf) and performed single cell RNA sequencing. Clustering analysis indicates subdifferentiation of ASM as early as 48 hpf and subsequent diversification into corneal epithelium/ endothelium/stroma, or annular ligament (AL) lineages. Tracking individual clusters reveals common developmental pathways, up to 72 hpf, for the AL and corneal endothelium/stroma and distinct pathways for corneal epithelium starting at 48 hpf. Spatiotemporal validation of over 80 genes found associated with AS development demonstrates a high degree of conservation with mammalian trabecular meshwork and corneal tissues. In addition, we characterize thirteen novel genes associated with annular ligament and seven with corneal development. Overall, the data provide a molecular verification of the long-standing hypothesis that POM derived ASM give rise to AS tissues and highlight the high degree of conservation between zebrafish and mammals.

The development of the eye is a complex process that includes the orchestration of tissue types from different origins. One of these tissue types is the periocular mesenchyme, a subgroup of neural crest cells, that gives rise to the tissues of the anterior segment (AS) of the eye. Two major components of the anterior segment are the cornea and iridocorneal angle (ICA). The former's main task is to focus incoming light onto the retina, while the latter is critically involved in regulating the intraocular pressure of the eye[1–6]. Maldevelopment of the anterior segment and particularly these two components, can have severe consequences for eye function and can lead to blindness in the form of cataracts or glaucoma[4,7].

Amongst several well-established model organisms used in the study of AS development, zebrafish embryos offer a unique opportunity to not only visualize AS development in real time but also quarry its components at numerous intervals and in high quantity. The mature zebrafish cornea is organized in five layers like the mammalian cornea: (1) the epithelium, (2) Bowman's layer, (3) stroma, (4) Descement's layer and (5) the endothelium[8–10]. The corneal epithelium connects the cornea with the outside world. Underneath the epithelium and anterior to the stroma is the Bowman's layer, which mainly consists of collagens (Collagen IV and VII). This is followed by the actual stroma (stroma proper), which is mainly made of extracellular matrix proteins, for the most part specifically organized collagens (mainly collagen I). The stroma itself is acellular but contains keratinocytes, which originate from neural crest cells. This is followed by the acellular Descement's layer which is the basement membrane of the monolayered corneal endothelium[8,10,11]. The main differences between the mammalian and zebrafish cornea are, that the corneal epithelium of zebrafish is thicker (circa 60% of the entire corneal thickness), likely due to the absence of eyelids and immediate exposure to the surrounding water medium. Also, in mammals the corneal stroma is the main contributor to the corneal thickness, whereas it makes up only approximately 30% of the thickness in zebrafish[10]. The initial development of the zebrafish cornea appears very rapid. At 30 hpf, a two-layered epithelium and a rudimentary stroma are already present. However, further differentiation and maturation is a much slower process. For instance, it takes 5 days after fertilization for the Bowman's layer to appear, and keratinocytes only start invading the stroma between 14 and 28 days post fertilization. Full maturity of the cornea is reached after approximately 2 months[10]. The cornea has been a matter of research interest for some time and several marker genes are known, whose mutation can cause maldevelopments[3,12]. However, a detailed analysis of gene expression during early development is largely missing.

Department of Biology, University of Kentucky, Lexington, USA. ✉email: jakub.famulski@uky.edu

In contrast to the cornea, the iridocorneal angle (ICA) of zebrafish is significantly less studied. In mammals, the main parts of the ICA are the trabecular meshwork (TM) and the Schlemm's canal (SC). These tissues have important roles in regulating the flow and drainage of the aqueous humor of the eye and thereby maintaining proper intraocular pressure (IOP)[4,7,13]. In the absence of proper flow and drainage, the IOP can increase which in turn can lead to glaucoma, one of the leading causes of blindness in the US and worldwide[4,7,13–15]. Accordingly, an increasing amount of attention has been paid to understand the trabecular meshwork (TM) and the Schlemm's canal (SC), specifically the underlying genetics[4,16–19]. These studies catalogued gene expression from adult tissues at the singe cell level and in doing so, uncovered numerous potential targets for the study of ICA function. A potential roadblock to taking advantage of this treasure trove of data is the laborious nature of ICA functional examination in current models such as mice or primates. Zebrafish have become a well-accepted model for ocular disease modeling, in particular for retinopathies, however their use in ASD research has been less pronounced[8,20–25]. One potential issue is the physiological difference between mammalian and zebrafish ICA. The tissue within the zebrafish ICA is called the annular ligament (AL), which is similarly organized to the TM and SC of mammals and fulfills the same function[8,9,26,27]. It is a porous tissue, filled with glycoproteins and leads to a network of channels ending in the episcleral vasculature, like the SC. Although the morphological similarity implies a similar function regarding drainage of the aqueous humor, a drainage system without the involvement of the AL has also been suggested[27].

In contrast to mammalian ICA which drains circumferentially, the AL facilitates drainage dorsally and aqueous humor production ventrally. Furthermore, previous studies suggested that the AL may not be a reliable a model for the human TM due to differences in the proteins tested to reside in the tissues[26]. However, at present we know very little about the actual genetic and molecular profile of the AL and a detailed characterization would help to determine it's usefulness as a model for human TM diseases like glaucoma[28].

While recent studies have revealed some of the molecular composition of the adult cornea and ICA, there is still a lot more to uncover, particularly regarding the molecular mechanism regulating their development[16–19,29–32]. Work from various model systems, including zebrafish, chick, mouse, and human tissue, has firmly established the physiological landscape and developmental timing for the assembly of the AS[3,29,33–37]. It has also affirmed the conservation and physiological function between such distinct species. The missing piece of the puzzle is how each of the tissues comprising the AS becomes functional and how this is regulated during development. Having a molecular atlas of these critical developmental events would go a long way to aid in the detection and possibly prevention of congenital blinding disorders associated with malformation of the AS. Novel approaches and advancements in the field of transcriptomics give an opportunity to finally begin such investigations. Recently, there has been an explosion of single cell transcriptomic analysis pertaining to retinal development, however, data for the development of the AS and particularly the TM are still sparse[17–19,29]. Furthermore, the current single cell transcriptomic data of the vertebrate eye, mainly include tissue from a single, mostly adult time point. This includes recent work on human as well as monkey and pig AS samples which were also limited to only mature tissue that has already fully differentiated[16–19,30–32]. As such, in our current study we used the highly versatile zebrafish embryo system to systematically analyze transcriptomic changes within ASM isolated from the AS over the course of early development. From our analysis of single cell transcriptomes, we have established clear developmental trajectories for progenitor cells of the cornea and annular ligament while uncovering new biological pathways that may be involved in the development of the zebrafish AS. Furthermore, we highlight that the annular ligament has been discounted in previous studies as divergent from the mammalian drainage structures, while it does exhibit gene expression patterns conserved with those of the trabecular meshwork[8,9,26]. Overall, our study yields the first developmental atlas of AS development with single cell resolution and further supports that zebrafish can be used as model organism to study human diseases affecting the cornea and iridocorneal angle.

Materials and methods

All experiments were approved by the IRB/Ethics committee and carried out according to approved guidelines for animal welfare, IACUC protocol 2021-3781 at the University of Kentucky assurance # D16-00,217 (A336-001) and in accordance to University of Kentucky biosafety regulations according to protocol B21-3711. All methods are reported in accordance with ARRIVE guidelines.

Zebrafish maintenance. Zebrafish were maintained and treated according to the IACUC standards of the University of Kentucky (protocol # 2021-3781). The embryos were initially raised in embryo media (E3) at 28 °C for 24 h and thereafter in E3 medium containing 1-phenyl 2-thiourea (PTU) to maintain embryo transparency. The transgenic line Tg[*foxc1b*:GFP] was a gift from Dr. Brian Link. As wildtype we used the AB line. All zebrafish methods reported are in accordance with ARRIVE guidelines.

Whole-Mount in situ hybridization (WISH). Whole mount in situ hybridizations were performed as previously described[38]. Briefly, larvae of different developmental stages (24 hpf, 72 hpf and 120 hpf) were fixed overnight in 4%PFA in PBS. At minimum 20 embryos were used for each timepoint. Antisense RNA probes were generated by means of polymerase chain reactions, including T7 promoters in the primers from a mix of 24 hpf, 48 hpf and 72 hpf cDNA and subsequent transcription with T7 Polymerase (Roche). Primer sequences used are found in Supplementary Table S1. Images of the embryos were taken with a Nikon Digital Sight DS-U3 camera and Elements software (Nikon). Image adjustments for brightness and contrast were done with Adobe Photoshop and figures were assembled using Adobe Illustrator. 120 hpf embryos, grown without PTU, were euthanized using MS222 and subsequently fixed in 4% paraformaldehyde for 24 h. Fixed embryos were treated in 10% and subsequently 30% sucrose for 24 h each. Post sucrose washes, the embryos were mounted in OCT (optimal cutting temperature) embedding medium (Sakura) and frozen in cryosection molds. Cryosections

were collected at 10 µm thickness using a Leica cryotome. Images were collected using a Nikon Ti2 compound microscope with a 20X NA 0.95 objective. Subsequent processing was performed using Adobe Photoshop and Illustrator.

Single cell transcriptome analysis. Single cell transcriptomes were generated as described previously[39]. In brief, embryos from the transgenic line Tg[*foxc1b*:GFP] were collected every 24 h between 48 and 144 hpf. Approximately 100 embryos were anesthetized with 3-amino benzoic acid ethyl ester (Tricaine). Then, their eyes were dissected and subsequently collected on ice, before being incubated for 8 min in 0.25% Trypsin with EDTA at 37 °C. After filtering, the dissociated single cells were sorted for GFP+ at the University of Kentucky Flow Cytometry and Immune Monitoring Core at the Markey Cancer Center. Thereafter, approximately 2000 cells per time point were loaded onto Chromium $10 \times$ V3 chips (10× genomics) and processed in the University of Kentucky Department of Biology Imaging Core to generate single cell barcoded cDNA. For sequencing, samples were sent to and processed at the University of Illinois at Urbana-Champaign Roy J. Carver Biotechnology Center. The resulting sequences were aggregated (incorporating mapped normalization), using the Cell Ranger 6.0 software (10× genomics). Further, we used the cell ranger pipeline to align our data to the GRCz11 assembly (NCBI Genome accession number: GCA_000002035.4).

Monocle 3 software was used according to the recommended settings to identify cell clusters, construct single-cell trajectories, and identify differentially expressed genes[40,41]. In brief, for every timepoint analyzed, cell types were clustered using principle component analysis (PCA) (n = 100 dimensions), we reduced the dimensions using uniform manifold approximation and projection (UMAP) and removed batch effects according to Haghverdi et al.[42]. After clustering, we identified the top 50 over-represented genes per cluster and manually annotated the represented cell types, based on the findings from our in situ hybridizations and public databases. To perform pseudotime analysis, we aggregated our five timepoints, performed preprocessing clustering as described above for individual timepoints, and then manually identified roots using the convention that cell types from younger individuals should represent earlier stages in their developmental trajectory. Top 500 genes from each cluster at each time point are found in Supplementary Table S2.

Fluorescent two color in situ hybridization (FWISH). Two color fluorescent in situ wholemount hybridization (FWISH) was performed as previously described[39]. In short, FITC and DIG labelled probes were co-hybridized, detected individually using FastBlue or FastRed (Sigma), and subsequently imaged using the C2+ Nikon confocal microscope using a 20X 0.95NA objective. DAPI was added to stain nuclei. 3D stacks were collected with 3 µm steps and displayed as volume projections. At minimum 20 embryos were collected and subjected to the FWISH protocol for each probe set combination.

Results

Assembly of a developmental AS transcriptomic atlas. To determine transcriptional changes occurring during early AS development we sought to isolate and analyze single cell transcriptomes of AS progenitors, the ASM. Owing to the lack of a universal genetic marker for ASM cells, analysis of whole eye tissue would limit our ability to isolate and specifically analyze ASM cells. As such, in order to facilitate direct sampling of ASM cells we employed the transgenic reporter line: Tg[*foxc1b*:GFP]. We have recently shown that ASM cells labelled in Tg[*foxc1b*:GFP] migrate into the AS space and contain transcriptomic profiles consistent with being AS progenitors[39]. The *foxc1* transcription factor is also a well-known regulator and marker of the AS[39,43]. To confirm that Tg[*foxc1b*:GFP] GFP+ cells can serve as reliable sources of ASM cell thorough AS development we assayed their expression patterns over AS development from 24-144 hpf (Fig. 1A). 3D confocal images from whole mount Tg[*foxc1b*:GFP] embryos show a robust and highly localized pattern of GFP+ cells in the anterior most regions of the eye, or the AS, up to and including 144 hpf. Furthermore, cryosections confirm that Tg[*foxc1b*:GFP] GFP+ cells were found throughout the central and peripheral regions of the AS, including the cornea and AL at all the timepoints examined (Fig. 1B). Overall, our previous and current characterizations indicate that Tg[*foxc1b*:GFP] GFP+ cells offer a comprehensive representation of ASM during the timepoints examined, 48-144 hpf, and therefore provide an appropriate source for transcriptomic analysis of AS development. To generate ASM specific transcriptomes, we collected whole eyes of Tg[*foxc1b*:GFP] embryos at 48, 72, 96, 120 and 144 hpf, dissociated them into a single cell suspension. Because Tg[*foxc1b*:GFP] GFP+ cells are not exclusive to the AS, fluorescence activated cell sorting (FACS) was performed on cells from dissected eyes, therefore ensuring only ASM associated GFP+ cells were collected. Purified ASM GFP+ cells were processed using the 10X Genomics Chromium platform to generate scRNA libraries (Fig. 1C). Two independent biological replicates were collected for each time point. Sequencing results indicated an average of ~ 2400 cells total were captured per time point, with a total data set cell count of 12,234. Each of our samples had an average of 800+ genes identified from ~ 100,000 + reads/cell. Sequencing data was processed using CellRanger 6.0 (10× genomics) and subsequently analyzed for UMAP based cluster distribution and pseudotime using Monocle3[41].

Anterior segment tissue type classification. Prior to conducting cluster and pseudotime analysis of the ASM scRNA data, we sought to establish classification criteria for AS tissue types. Since whole eyes were dissociated to isolate GFP+ cells via FACS, we needed to establish criteria to identify which cell and tissue types are represented by the clusters. Although zebrafish are well established as a model for ocular development, the field lacks a comprehensive set of marker genes for developing AS tissues. To establish reliable genetic markers for corneal and AL tissues, we conducted extensive literature and databank (zfin.org) searches to classify genes for specific tissue types. Since our study is the first to characterize zebrafish developmental stages of AS formation, we relied on previous research that primarily focused on adult tissues from various species and extrapo-
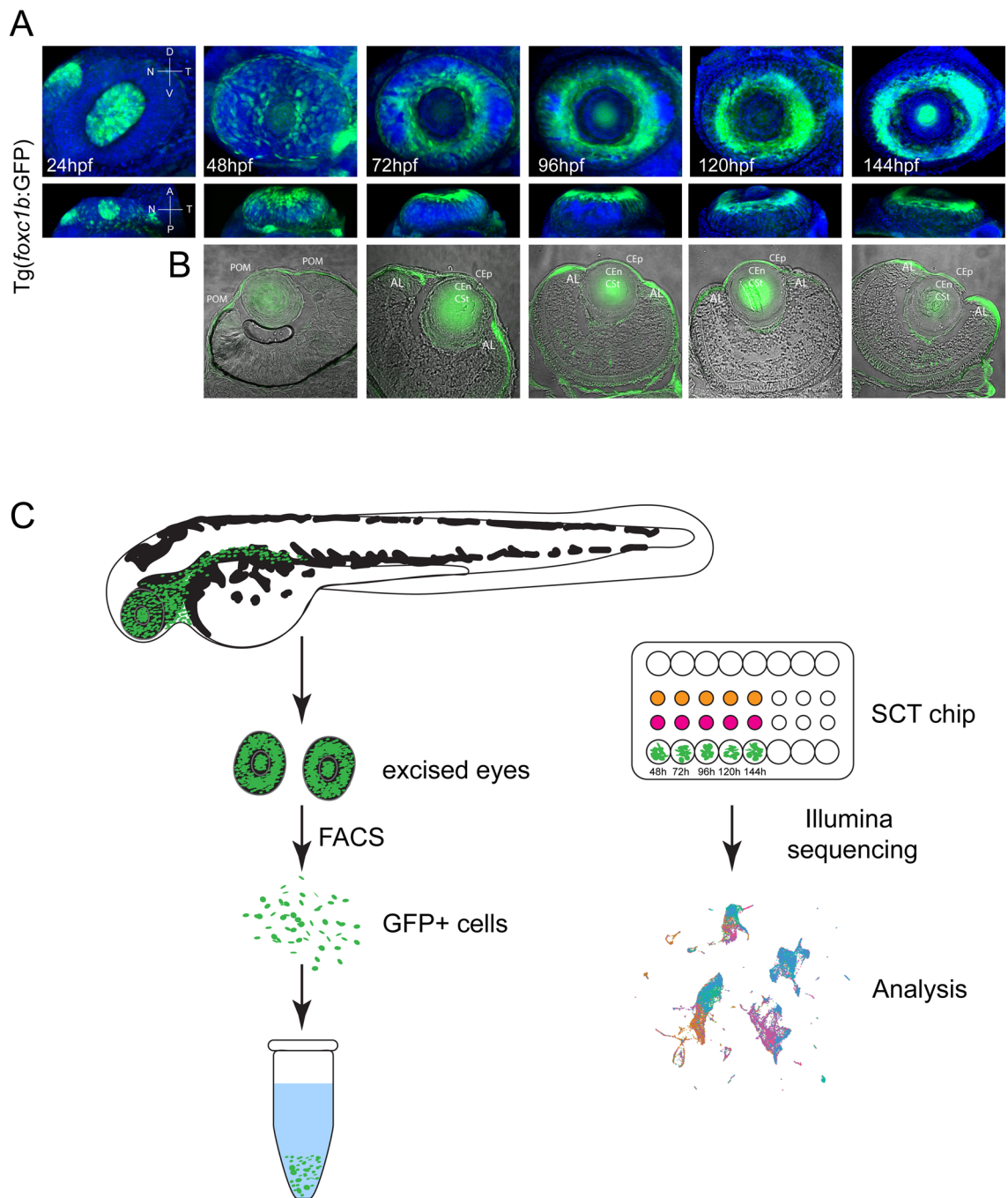
Figure 1. Strategy for generating ASM single cell transcriptomes over the course of early AS development. (**A**) Volume projections of ventral and dorsal views from 3D confocal stacks of whole mount Tg[*foxc1b*:GFP] embryos at 24, 48, 72, 96, 120 and 144 hpf. GFP+ cells (green) can be seen populating the AS regions by 48 hpf and assembling the typical doughnut like AS localization pattern from 72-144 hpf. DAPI (blue) was used to stain for nuclei. (**B**) Cryosections of Tg[*foxc1b*:GFP] embryos at 48, 72, 96, 120 and 144 hpf captured using confocal microscopy. GFP signal (green) is found in the AS regions in each of the time points. Transmitted light was used to capture tissue architecture. (**C**) Schematic for capture and subsequent assembly of single cell ASM transcriptomes. Dissected whole eyes were dissociated and GFP+ cells isolated using FACS. Subsequently single cells are captured using the 10X Genomics Chromium and the resulting single cell cDNA sequenced and analyzed.

lated from those findings[3,4,18,19,28–32]. As a primary source, we turned to recent single cell transcriptome analyses of individually dissected anterior segment tissues generated for various mammals, including human, different monkeys, and pigs[17–19,29–32,44]. Additionally, we also referenced analyses of the zebrafish AS performed using

7 day old and adult tissues[45]. Subsequently, we performed whole mount in situ hybridization (WISH) for chosen candidate genes at three different developmental stages, 24 hpf, 72 hpf and 120 hpf, including cryosections to validate expression patterns (Fig. 2). The chosen stages represent initiation of ASM colonization, completion of colonization and differentiation and finally rudimentary assembly of the AS respectively[9,39]. The two major tissue types belonging to the anterior segment we sought to classify were the cornea and the annular ligament (Fig. 3). Both are known to derive from POM and therefore the ASM. We subdivided the cornea into corneal epithelium (CEp), corneal endothelium (CEn) and corneal stroma (CSt). We based classification of **CEp** on expression of keratins *krt4, krt5* and *krt91* as well as *icn2*[46–48]. Among the studied tissues, the corneal epithelium is the most studied not only in zebrafish, but also mammals, likely due to its external position in the eye, which makes it easily accessible[3,31,32,44]. Also, as pointed out by Takamiya and co-workers (2015), penetrance of in situ probes can be more difficult in deeper tissues[45]. For all the CEp marker genes, *krt4, 5, 91* and *icn2*, we can clearly distinguish expression across the anterior most corneal region, the epithelium (Fig. 2A). Cryosections of 120 hpf larvae further confirm these expression patterns (Fig. 2B). Combinatorial expression analyses, using two color fluorescent WISH (FWISH) illustrates the high degree of co-expression of CEp markers at 120 hpf (Fig. 2C). We classified **CEn** identity based on the expression of *pmp22a* and *ctnnb2*[45,49], which was confirmed via WISH and cryosections (Fig. 2D–E). Expression of CEn markers differs from the CEp associated keratins with a significant reduction of expression in the most anterior corneal regions and higher expression in the dorsal and ventral periphery of the cornea. Maturation and differentiation of the corneal endothelium is a relatively slow process and full maturation is only reached after approximately 2 months[10]. A recent publication described several marker genes of corneal tissues, however only two genes were described in the corneal endothelium *pitx2* and *ctnnb2*[45]. Though, in the developing embryo the expression of *pitx2* is not limited to the endothelium[39] and the expression of *ctnnb2* was not known at stages younger than 7 days. Our in situ hybridizations indicated, that *ctnnb2* is indeed expressed in the developing cornea and might serve as a marker for the developing corneal endothelium at early time points. Furthermore, fluorescent double in situ hybridizations show, that *ctnnb2* is co-expressed with *pmp22a* in the corneal endothelium of zebrafish (Fig. 2F). Nonetheless, finding clear and specific marker genes for the corneal endothelium at these early developmental stages remains challenging due to ongoing differentiation. We classified **CSt** based on the expression of *dcn* and *lum*[11]. The development of the corneal stroma is also a slow process and during the timepoints examined in this study must be considered mostly rudimentary and still maturing. Morphological studies of the stroma have shown how different collagens begin to form layers around 5–7 days post fertilization, but they also describe the stroma itself as mostly absent of cells in these early stages[10,20]. WISH and cryosection images of *dcn* and *lum* confirm expression in the corneal region, primarily anterior of the lens (Fig. 2G–H), while co-expression can be detected in ventral and dorsal regions (Fig. 2I). As such, we determined these markers to be suitable for CSt identification (Fig. 2G–H). Finally, we chose to classify the **AL** by the expression of the *genes hmgn2, hmgb2b, scinla,* as well as *myoc*[18,19,50–53]. As previously mentioned, the molecular composition of AL has been seldom studied in zebrafish and as such, apart from *scinla* (previously described as gsln1), much of our identification of markers of this structure was based on assumptions of conservation with the mammalian ICA. Previous studies in zebrafish mainly focused on identifying a specific marker but did not explore the gene expression profile in detail[51–53]. Hence, due to the similarities in morphology and function, we used genes described in mammalian TM studies to identify basic markers. Hereby, we primarily relied on single cell transcriptome analyses that included isolated trabecular meshwork and Schlemm's canal cells as well as TM cell culture studies[18,19,50,54–57]. As was done for all the corneal tissues, AL marker gene expression patterns were also confirmed with WISH and cryosections (Fig. 2J–K). In general, we observed two distinct patterns of expression, one concentrated solely within the AL, *hmgb2b, hmgn2* and *myoc*, and the other exhibiting expression in the AL and in the CEn, *scinla* and *lxn* (homolog of rarres1). Since the AL is not fully developed until 17 days, at 120 hpf we refer to the cells in the AL region as annular ligament progenitors (ALP) (Fig. 2). FWISH analysis indicated that all four markergenes are found to co-express in the AL to varying degrees (Fig. 2L). Taken together, we felt confident to use the aforementioned tissue identifier genes as the criteria for assigning CEp, CEn, CSt or AL identity to cell clusters in our ASM scRNA dataset.
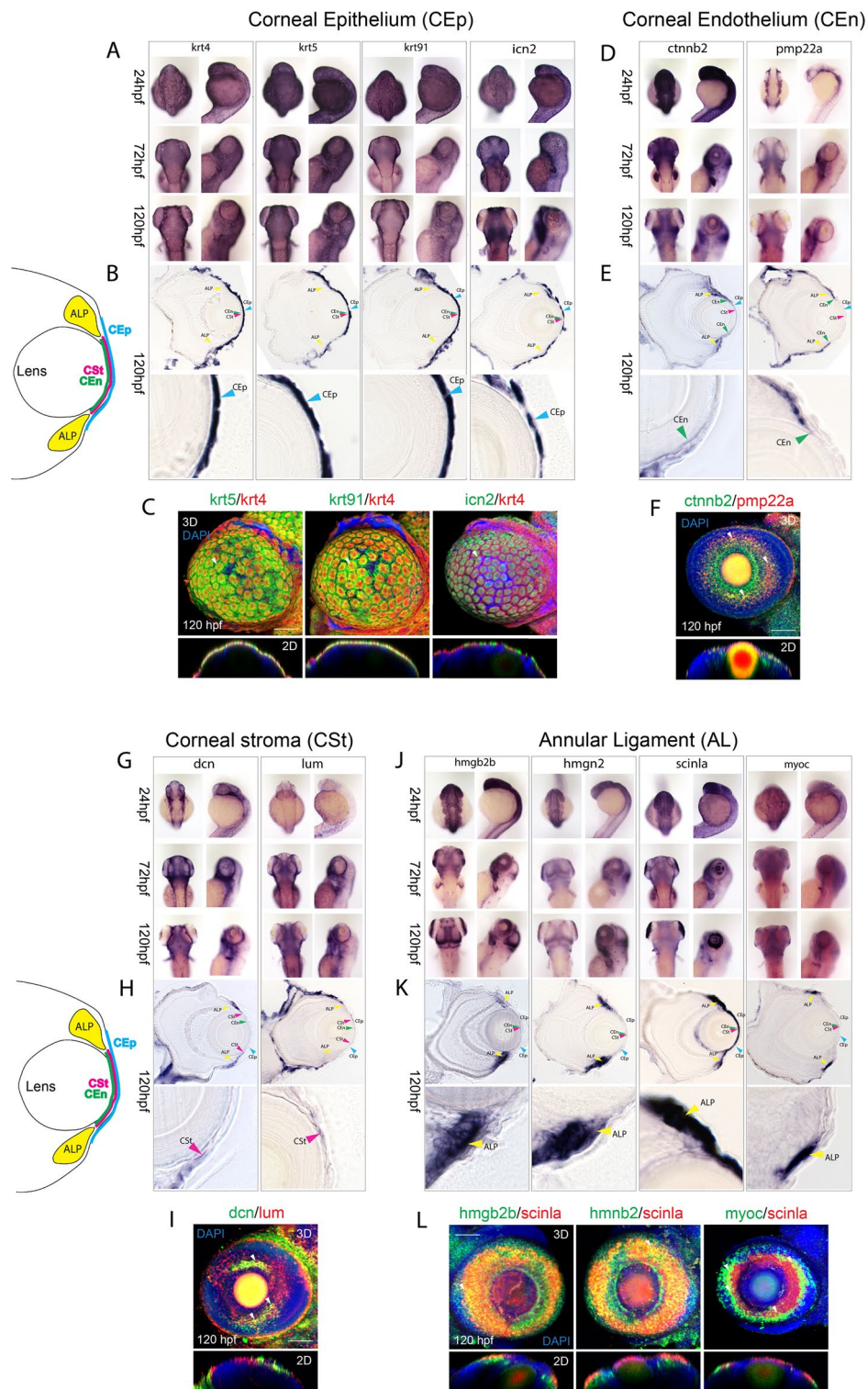
ASM scRNA cluster analysis throughout early AS development.

Having collected single cell sequencing data from various timepoints we first sought to analyze cellular distribution at each timepoint individually. We aggregated the biological replicates and performed UMAP clustering using Monocle 3[41]. We analyzed every timepoint, using principle component analysis (PCA) (n = 100 dimensions) and removed batch effects according to Haghverdi et al.[42]. Assignment of cluster identity was based on expression of the marker genes previously established (Fig. 2). As expected, data from early timepoints, 48 hpf and 72 hpf, depicted only a handful of clusters (Fig. 3A–B). At this early time of AS development, we expected the ASM to remain multipotent. However, already at 48 hpf we can identify clusters with AL and CEp signatures (Fig. 3A). The presence of CEp clusters was unexpected as corneal epithelium has been proposed to derive solely from the ectoderm (see discussion). Interestingly, CEp was bifurcated into CEp1, classified by elevated expression of *krt5* and *krt91*, and CEp2, classified by elevated expression of *krt4*, *cyt1* and *icn2* (Fig. S2). The AL signature was also found in two distinct populations termed AL1 and AL2 where AL1 is designated by elevated expression of *hmgn2, hmgb2b, mcm7* and *pcna* while AL2 is designated by elevated expression of *scinla* and *lxn* (Figs. 3A, S2–3). By 96 hpf, we observed a significant stratification of clusters, in particular for those associated with the annular ligament (Fig. 3C). Individual AL1 and AL2 clusters persist, but at 96 hpf we additionally detected clusters with signatures of the CEn (elevated *pmp22a* and *ctnnb2* expression) mixed with AL2 (AL2/CEn) (Figs. 3C, S4) and with CSt (CEn + CSt). This suggests that AL2, CEn and CSt cells may share a common progenitor and at 96 hpf still retain some levels of multipotency. 96 hpf is also when we detect the first clusters with signatures of solely CSt (elevated *dcn* and *lum* expression) mixed with CEn (Fig. S4). Again, this suggests that cells slated to become AL, CSt

Figure 2. Assembly of markers for developing zebrafish corneal epithelium, endothelium, stroma and annular ligament. (**A**) WISH analysis of corneal epithelium markers *krt4, ktr5, krt91* and *icn2* expression at 24, 72 and 120 hpf displayed in dorsal and lateral views. (**B**) 10 μm cryosections of 120 hpf embryo eyes after WISH. **C**) Two color FWISH depicting co-expression of *krt4*(red)/*krt5*(green), *krt91*(green)/*krt4*(red) and *icn2*(green)/*krt4*(red) at 120 hpf. DAPI (blue) was used to stain the nuclei. Volume projections in lateral view and single section dorsal views from 3D confocal stacks are displayed. White arrows indicate example regions of co-expression. Scale bar = 100 μm. Individual fluorescence channels are provided in Fig. S1A. (**D**) WISH analysis of corneal endothelium markers *ctnnb2* and *pmp22a* expression at 24, 72 and 120 hpf displayed in dorsal and lateral views. (**E**) 10 μm cryosections of 120 hpf embryo eyes after WISH. (**F**) Two color FWISH depicting co-expression of *pmp22a*(red)/*ctnnb2*(green) at 120 hpf. DAPI (blue) was used to stain the nuclei. Volume projections in lateral view and single section dorsal views from 3D confocal stacks are displayed. White arrows indicate example regions of co-expression. Scale bar = 100 μm. Individual fluorescence channels are provided in Fig. S1B. (**G**) WISH analysis of corneal stroma markers *dcn* and *lum* expression at 24, 72 and 120 hpf displayed in dorsal and lateral views. (**H**) 10 μm cryosections of 120 hpf embryo eyes after WISH. (**I**) Two color FWISH depicting co-expression of *lum*(red)/*dcn*(green) at 120 hpf. DAPI (blue) was used to stain the nuclei. Volume projections in lateral view and single section dorsal views from 3D confocal stacks are displayed. White arrows indicate example regions of co-expression. Scale bar = 100 μm. Individual fluorescence channels are provided in Fig. S1C. (**J**) WISH analysis of annular ligament markers *hmgb2b, hmgn2, scinla* and *myoc* expression at 24, 72 and 120 hpf displayed in dorsal and lateral views. **K**) 10 μm cryosections of 120 hpf embryo eyes after WISH. (**L**) Two color FWISH depicting co-expression of *hmgb2b*(green)/*scinla*(red), *hmnb2*(green)/*scinla*(red) and *myoc*(green)/*scinla*(red) at 120 hpf. DAPI (blue) was used to stain the nuclei. Volume projections in lateral view and single section dorsal views from 3D confocal stacks are displayed. White arrows indicate example regions of co-expression. *ALP* Annular ligament progenitor (yellow arrow), *CEp* Corneal epithelium (blue arrow), *CEn* Corneal endothelium (green arrow), *CSt* Corneal stroma (green arrow). Scale bar = 100 μm. Individual fluorescence channels are provided in Fig. S1D.

and CEn share a common developmental pathway. Finally, we also continue to detect distinct CEp1 and CEp2 clusters at 96 hpf. We never detect any association between CEp and AL, CEn or CSt clusters at any timepoints. As development proceeds, throughout 120 and 144 hpf, we continued to detect all the aforementioned corneal and annular ligament clusters in addition to detecting the first sole CEn cluster at 144 hpf (Figs. 3D–E, S5–6). We never found a sole CSt cluster, likely stemming from the slow immigration of cells into these tissues which concludes later than 144 hpf. To further validate the choice of genetic markers for assigning our cluster identities, we also examined relative expression of said markers at each timepoint (Fig. S7). Most of the assigned markers display a steady state expression throughout early development further confirming their suitability for tissue identification. Taken together, our data suggest that the early stratifications, between 72 and 96 hpf, become established developmental pathways for these cell types. These conclusions are supported by morphological characterizations of the timing of AS development in zebrafish[8,9]. First cells within the iridocorneal angle can be identified as early as 72 hpf, which grow in number by 120 hpf. However, it was also described that a full histological identification of the annular ligament needed until 17 days post fertilization and hence, function of these cells at early timepoints was questioned[9].

In addition to AS associated cell types, we also detected reoccurring clusters containing non-AS associated cells (Fig. 3). This likely represents the fact that while the majority of Tg[*foxc1b*:GFP] GFP+ cells was isolated from whole eyes represent ASM cells, a small subset must also involve non-ASM cells (Fig. 1B). The most abundant non-ASM associated cells included cells of the cardiovascular system (BV), expressing the markers *ednraa, rasl12* and *tie1*[58,59]. They also included pigment cells (P), expressing genes such as *tyrosinase* and *gch2*[60,61]. The pigment cell clusters were found at 72 and 96 hpf, and subsequently likely transition into retinal pigment epithelium (RPE) clusters throughout later timepoints (Fig. 3). Pigment cells, as iris stroma, are known to include a NCC developmental pathway and therefore it is not surprising to find them in our data set. Expression of select genes verified to associate with pigment cell fates are summarized in supplemental Fig. 8. In addition to vasculature and pigment cells, we also identified retinal-like clusters which were found to contain different retinal cell types, including rods (marker: *rho*), cones (marker: *gnat2*) and retinal ganglion cells. In total we found 21 clusters in our combined dataset. However, it must also be stressed that most non-ASM associated clusters contained very few cells compared to those of corneal or annular ligament lineages.

Molecular characterization of AS clusters during early development.

To showcase the unique molecular differences associated with each of the identified AS lineages (AL1, AL2/CEn, CEn + CSt, CEp1 and CEp2) we examined the top upregulated genes from each cluster at each timepoint (Fig. 4). At 48 hpf, as expected, the AL associated clusters exhibit high levels of expression for *hmgb2b* and *hmgn2*, but also *hmgb2a, h3f3a, fabp7a, cfl1, ccnd1* and *tmsb4x* (Fig. 4A). Uniquely for AL2 we also observe *pmp22a, tgfbi* and *tpm4a* upregulation. CEp1 and 2 are both defined by very high levels of *krt4* expression (Fig. 4A). CEp1 also exhibits higher levels of collagens *5a3b* and *1a1b* as well as *pnf1* compared to CEp2, while CEp2 exhibits high levels of *cyt1, cyt1l, anxa1c, agr1* and *icn2* (Fig. 4A). At 72 hpf, the primary difference between the AL and CEp multipotent clusters is high expression of *krt4, epcam, cyt1* and *tmsb1* in CEp and *hmgn2, cd81a, mdka, si:ch211-133n4.4* and *pmp22a* in AL (Fig. 4B). At 96 hpf, the time of significant stratification of the clusters, we see a much more defined molecular signature for all the different tissue types. CEp1 is defined by high expression of *krt4, krt5, krt91, pfn1, cldni* and *col11a1a* while CEp2 exhibits unique upregulation of *malb, cldnb, cldne, krt1-c5, cldnh* and *tnks1bp1* (Fig. 4C). AL1 is defined by high expression of *mdka, ccnd1, cadm3, hmgn2, fabp7a, mcm7, cad* and

Figure panels A–L showing gene expression in zebrafish corneal tissues. Panel headings: Corneal Epithelium (CEp), Corneal Endothelium (CEn), Corneal stroma (CSt), Annular Ligament (AL).
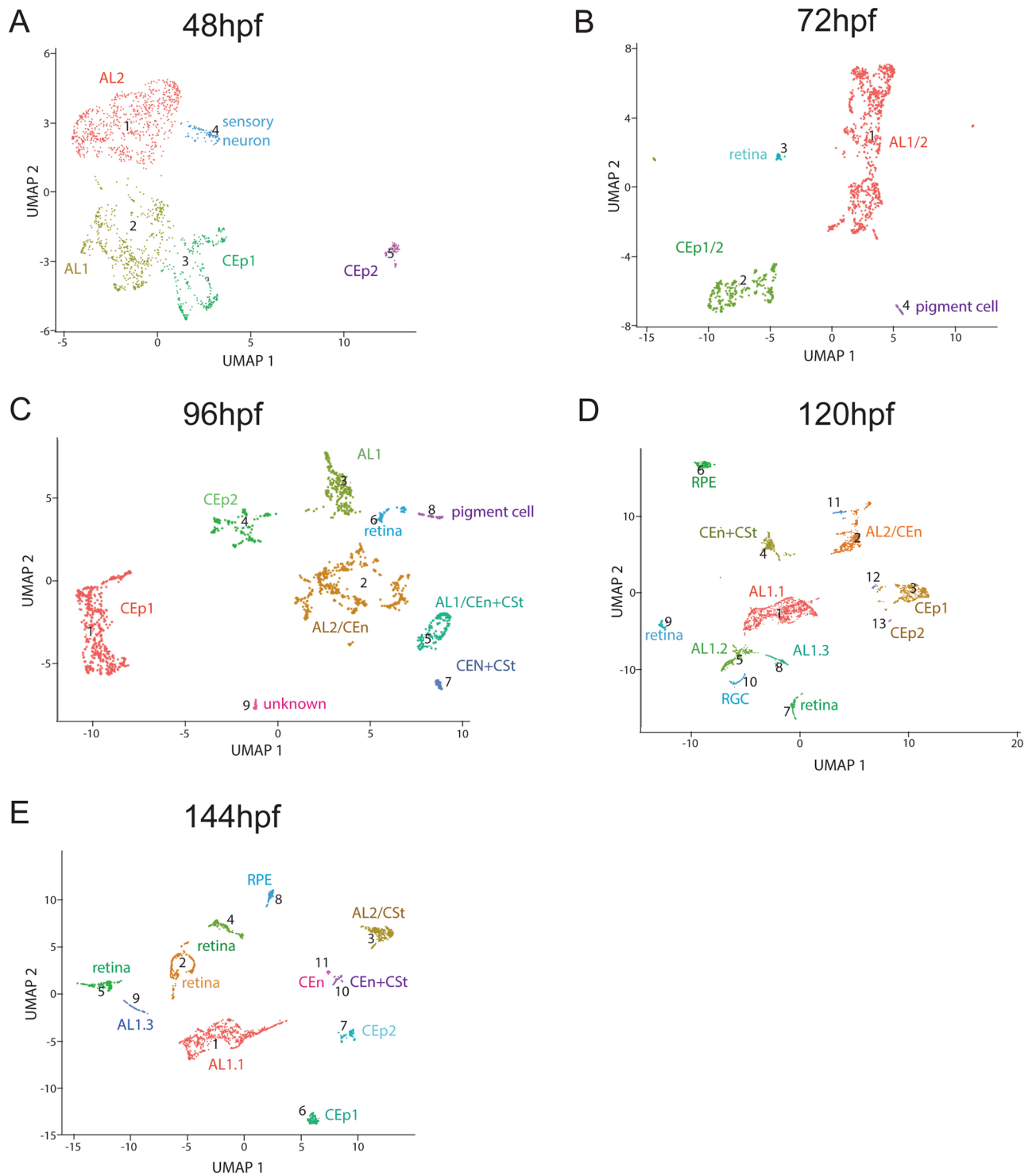
Figure 3. Clustering of purified ASM throughout early AS development. Cluster numbers are indicated in brackets. (**A**) Cluster distribution at 48 hpf. The five clusters observed were identified as: CEp1(2), CEp2(5), AL1(1), AL2(3) and sensory neurons(4). (**B**) Cluster distribution at 72 hpf. The four clusters observed were identified as: AL, CEp, retina and pigment cells. (**C**) Cluster distribution at 96 hpf. The nine clusters observed were identified as: CEp1(1), AL2/CEn(2), AL1(3), CEp2(4), AL1/CEn + CSt(5), retina(6), AL1/CEn + CSt(7), pigment cell(8), unknown(9). (**D**) Cluster distribution at 120 hpf. The fifteen clusters observed were identified as: AL1.1(1), AL2/CSt(2), CEp1(3), AL1/CEn + CSt(4), AL1.2(5), unknown(6), retina(7), AL1.3(8), retina(9), RGC(10), unknown(11). unknown(12), CEp2(13), unknown(14), unknown(15). (**E**) Cluster distribution at 144 hpf. The twelve clusters observed were identified as: AL1.1(1), retina(2), AL2/CEn(3), retina(4), retina(5), CEp1(6), CEp2(7), RPE(8), AL1.3(9), AL1/CEn + CSt(10), CEn(11), unknown(12).
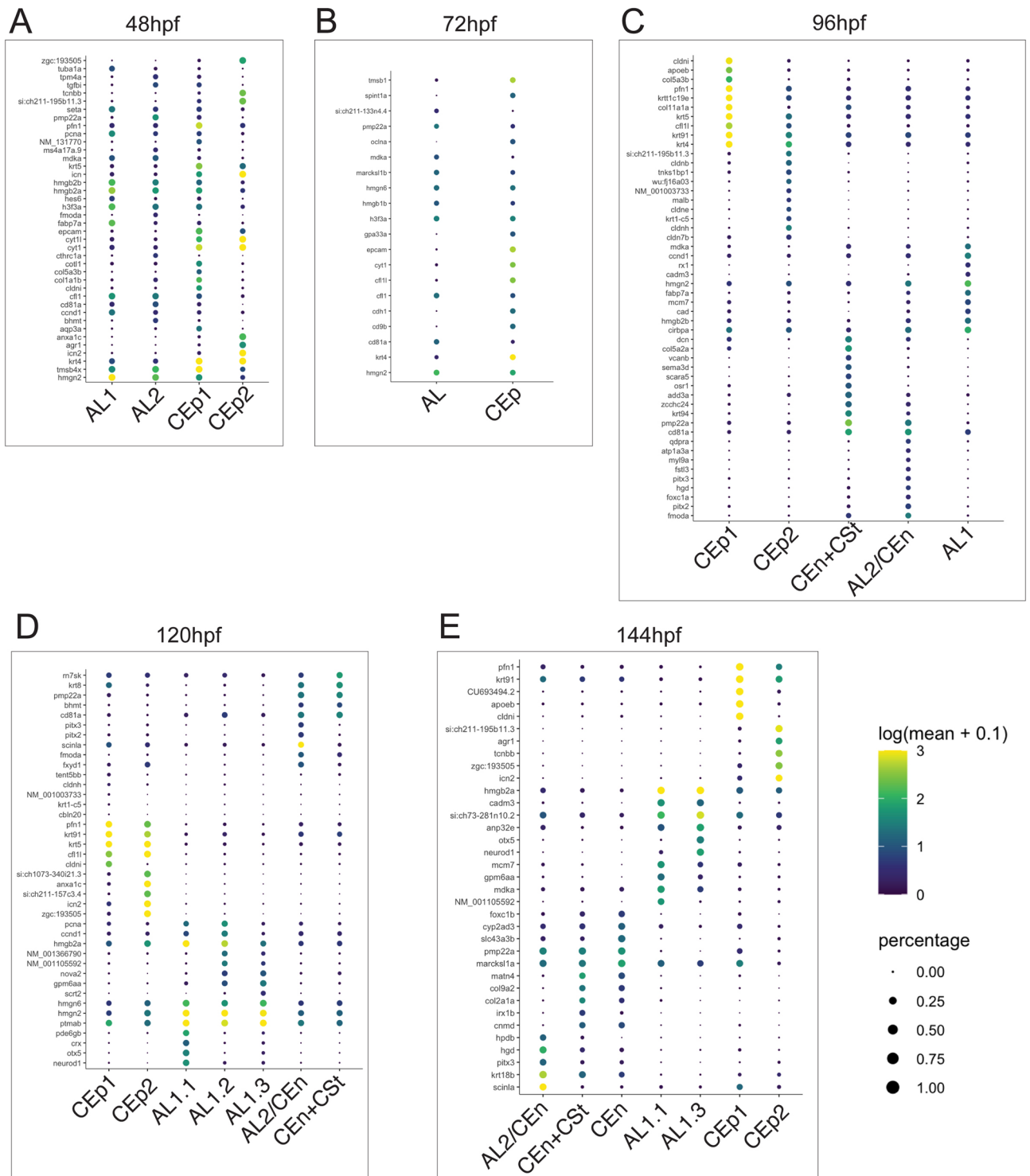
Figure 4. Top 10 genes for AS related cell types at each timepoint examined. (**A**) Top10 genes from AL1, AL2, CEp1 and CEp2 at 48 hpf. (**B**) Top10 genes from AL and CEp clusters at 72 hpf. (**C**) Top10 genes from AL1, AL1/CEn + Cst, AL2/CEn, CEp1 and CEp2 clusters at 96 hpf. (**D**) Top10 genes from AL1.1, AL1.2, AL1.3, AL1/CEn + CSt, AL2/CEn, CEp1 and CEp2 at 120 hpf. (**E**) Top10 genes from AL1.2, AL1.3, AL1/CEn + CSt, AL2/CEn, CEn, CEp1 and CEp2 at 144 hpf. Scale of degree of expression and percentage is indicated.

hmgb2b. CEn + CSt on the other hand exhibits high *dcn, col5a2a, cvanb, sema3d, scara5, osr1, add3a, pmp22a* and *cd81a*, while AL2/CEn cells display upregulated expression of *myl9a, fstl3, pitx3, hgd, foxc1a, pitx2* and *fmoda* (Fig. 4C). At 120 hpf, CEp1 continues to express *krt4, krt5, krt91, pfn1, cldni* at high levels, while also increasing

expression of *krt8*. CEp2 also exhibits high expression of *krt4, krt5, krt91,* and *pnf1,* in addition to *icn2, anxa1c* and several unknown genes including *zgc:193,505, si:ch1073-340i21.3* and *si:ch211-157c3.4* (Fig. 4D). By 120 hpf AL1, comprised of 3 individual clusters, AL1.1, AL1.2 and AL1.3. All three are defined by continued high expression of *hmgn2, hmgn6* and *hmgb2a* while upregulation of *nova2, gpm6aa* and *ccnd1* defines AL1.2 and AL1.3. CEn + CSt displays upregulation of *rn7sk, pmp22a, cd81a* and *krt8*. Distinct upregulated expression in AL2/CEn includes *pitx2, pitx3, scinla, fmoda* and *fxyd1* (Fig. 4D). By 144 hpf, we continue to see similar top upregulated genes in CEp1 while CEp2 exhibits upregulation of *agr1, tcnbb, icn2* and unknown genes *si:ch211-195b11.3* and *zgc:193,505* (Fig. 4E). AL1.1 and AL1.2 show similar expression patterns with upregulated *hmgb2a, mcm7,* but also now upregulate *cadm3, anp32e* and *mdka*. CEn + CSt continues to express high levels of *pmp22a,* while also increasing expression of *matn4, col9a2, col2a1a, cnmd* and *irx1b*. AL2/CEn also continues to express *pmp22a* at high levels, in addition to *hpdb, pitx3, hgd* and *scinla* (Fig. 4E). The first independent CEn cluster is defined by upregulated expression of *marcks1a, pmp22a, scl43a3b, cyp1ad3* and *foxc1b* (Fig. 4E). Based on these expression patterns we observe that many of the genes used in our assay as determinants of cell fate are constantly expressed at high levels throughout early development while other genes, many previously not associated with AS development, display various patterns of expression.

Developmental dynamics and trajectory analysis of AS development.

Having classified AS cell types over the course of early AS development, we next sought to track trajectories and relationships in time. To establish and illustrate the developmental relationship between clusters over time, we compiled a timeline (Fig. 5). Trajectories were determined by manually examining the top50 upregulated genes in each cluster and comparing patterns of expression in subsequent time points to generate most likely trajectories over time. Starting at 48 hpf with the AL clusters, AL1 and AL2, we track their convergence into one multipotent cluster at 72 hpf. By 96 hpf we track the AL cell type to stratify into AL2/CEn, AL1 and CEn + CSt. The main AL1 lineage further stratifies into three sub clusters at 120 hpf, AL1.1, 1.2 and 1.3, and settles as two clusters, AL1.1 and AL1.3 by 144 hpf. The CEn + CSt lineage also appears to be specified by 96 hpf and persists up to 144 hpf. However, at 144 hpf we also see a portion of CEn + CSt diverge into a possibly committed CEn lineage. These observations indicate that the AL1 lineage exhibits multipotency including potential for CEn and CSt lineages early in development. Conversely, starting at 96 hpf, the AL2/CEn cluster persists up to 144 hpf and is therefore likely already specified by 96 hpf. This suggests a very close relationship between AL and CEn cells during early development. When tracking CEp development, starting with CEp1 and CEp2, we observe the two clusters converge at 72 hpf. However, this is likely based on the fact that CEp2 generally occurs in smaller numbers which merged with CEp1 at this timepoint. Subsequently, they re-diverge into CEp1 and CEp2 populations by 96 hpf. Both CEp1 and CEp2 trajectories remain independent up to 144 hpf. This pattern was also observed for the AL clusters at 72 hpf. Overall, during early AS development the number of clusters representing the corneal epithelium remains at two, indicating the slow developmental pace after their rapid initiation. This finding is supported by the fact that zebrafish corneal epithelium remains 2-layered for the first 4 weeks of development[10].

To complement our cluster-based trajectories over developmental time we also performed a pseudotime analysis of our combined dataset using Monocle 3[41]. The major difference from our cluster analysis is that pseudotime analysis focuses on the trajectory of cells instead of whole clusters and helps identify how gene expression changes within a cell trajectory. Hence, this analysis allowed us to characterize if particular marker genes are strongly expressed at specific points in the trajectory. To perform pseudotime analysis all our timepoints were aggregated into a single data set and then subsequently analyzed using the Monocle 3 pseudotime algorithm (Figs. 6A, S9). We summarize our results based on the AS tissue types being tracked.

Corneal epithelium. When tracking markers of the corneal epithelium it becomes apparent, that marker genes representing CEp1 and CEp2 are expressed in this trajectory, indicating a close relation between the clusters. Many, of the characterized marker genes such as *krt4, krt5, krt91* or *icn2* are strongly expressed throughout the identified trajectory, validating their general importance for the cells (Fig. 6B). This appears to confirm the relatively conservative use of this core group of genes as markers for this tissue throughout early AS development. Other CEp associated genes display varying pattens of expression throughout the trajectory. For instance, *sparc* and *lgals1l1,* are strongly upregulated in the early parts of the trajectory but mostly absent from the later points (Fig. S9A). Interestingly, in contrast to most keratins that are upregulated throughout the trajectory, collagens (*col1a1b, col5a3b, col17a1a*) are upregulated early, but then decrease over time (Fig. S9B). Other genes, such as *icn2, anxa1c* and *pnp5a* show a reduced or completely absent expression in the early time points but are highly upregulated at the later points of the trajectory (Fig. S9C).

Corneal endothelium and stroma. Genes marking the corneal endothelium were found in the same trajectory as the AL markers, confirming our cluster analysis results (Fig. 6B). By analyzing the distribution of the genes in the trajectory, we found that *pmp22a* is strongly expressed throughout the entire trajectory. *Ctnnb2* and *dcn* on the other hand, share a sparse expression and a distribution that is relatively even throughout, but severely reduced toward the tail end of the trajectory (Fig. 6B).

Annular ligament. When comparing the pseudotime distributions of AL marker genes, we found that both tissue types, AL1 and AL2 share the same trajectory, further confirming the close relation between the two (Fig. 6B). AL marker genes *hmgb2b* and *hmgn2* are broadly distributed throughout the AL trajectory. However, upon closer inspection it appears that *hmgb2b* is present in early sections of the trajectory while mostly absent from later branches (Fig. 6B). *Scinla* and *fabp11a* (also called *fabp4a*) on the other hand, are mostly absent from the early branches of the trajectory but are strongly upregulated in later branches (Fig. 6B, S9D). Two previously
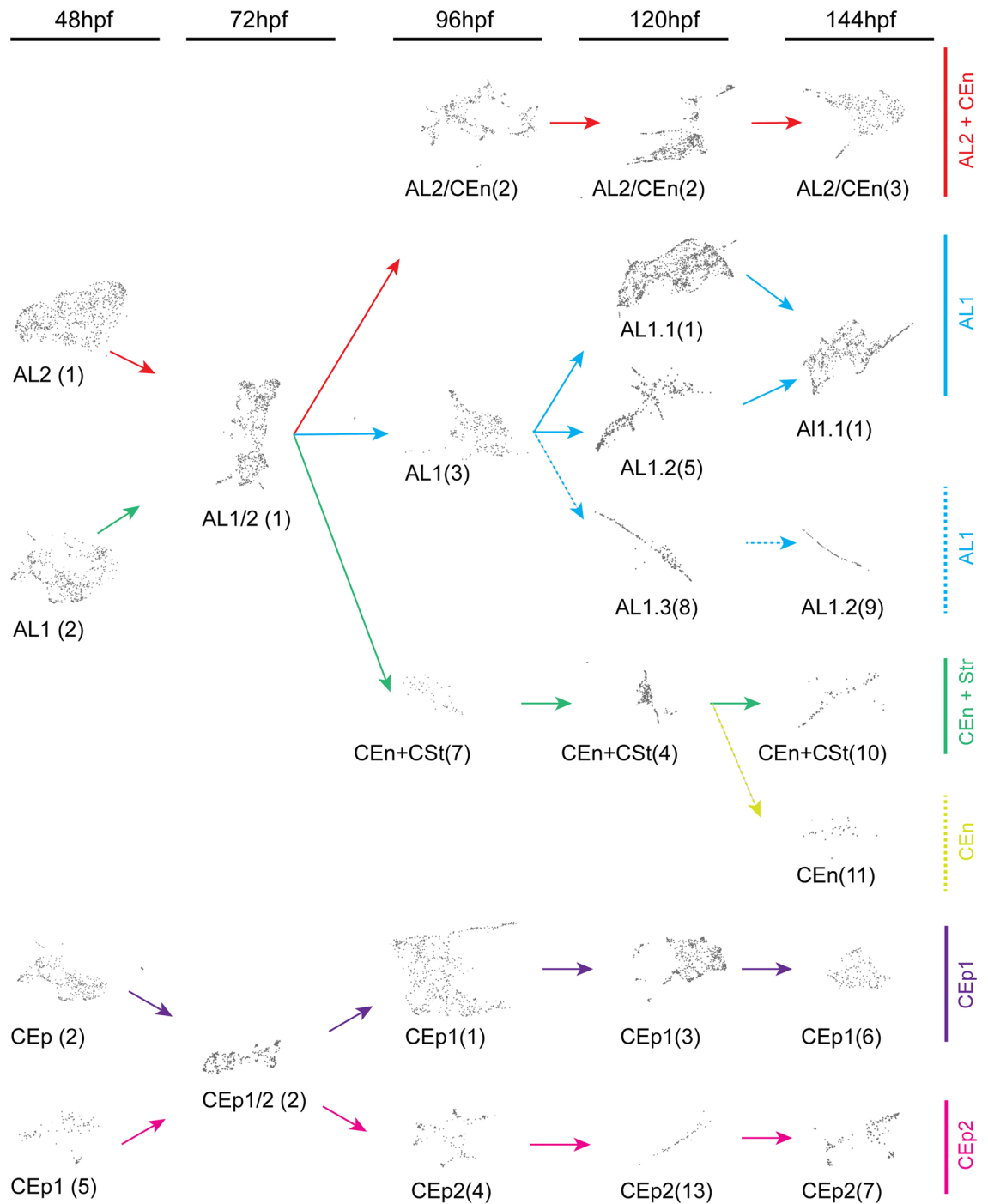
Figure 5. Cluster trajectories during early development of corneal epithelium, endothelium, stroma, and annular ligament. Developmental trajectories of individual clusters from each timepoint, 48–144 hpf, displaying temporal relationships. Numbers in brackets correspond to timepoint specific clusters (Fig. 3). Arrows indicate lineage relationships of the AL1 (blue), AL2/CEn (magenta), AL1/CEn + CSt (blue), CEn (light blue), CEp1 (purple) and CEp2 (pink) cell types. CEp lineages show early and dedicated commitment while AL lineages display multipotency for AL, CEn or CSt cell types.

unknown AS associated genes, *hgd* and *si:ch211-251b21.1* as well as *vim* and *hpdb* can be found expressed early in the trajectory and then again towards the end of the trajectory (Fig. S9E), while *fmoda* can be found throughout all branches of the trajectory (Fig. S9F). As such, our analysis of AL development, the first to be conducted in zebrafish, clearly defines this tissue as having arisen from multipotent progenitors which over time specify in AL, CEn and CSt dedicated populations. Overall, pseudotime analysis supports our cluster-based trajectory tracking and establishes developmental trajectories for zebrafish AS tissues.
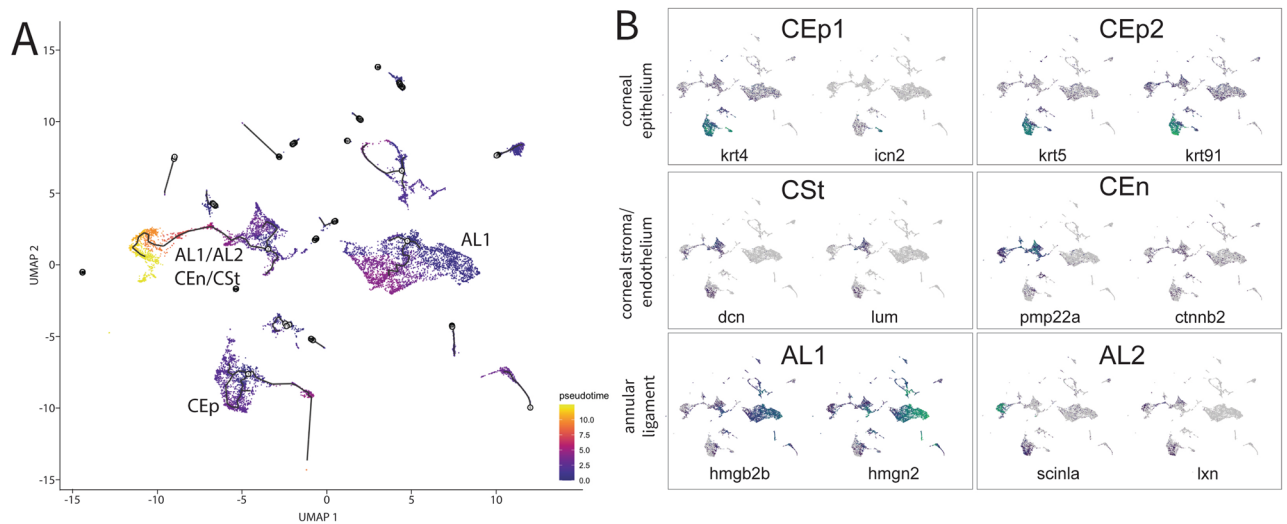
Figure 6. Pseudotime analysis of purified ASM cells during early AS development. (**A**) The trajectory analysis showing three major trajectories including CEp, AL + CEn/CSt, AL1. (**B**) Marker genes for the different cornea tissues and annular ligament projected onto the trajectory graph. Markers for CEp1: krt4 and icn2, for CEp2: krt5, krt91, for CSt: dcn and lum, for CEn pmp22a and ctnnb2, for AL1: hmgb2b and hmgn2 and for Al2: scinla and lxn.

Conservation of early zebrafish AS development.

Considering our data set being the first to examine expression over the course of AS development we wanted to assess the utility of our data set. To do so, we referenced genes in our data set to those that have been identified in mammalian, and particularly human AS tissues. Recent work examining adult cornea and the entire AS, also at single cell levels, was used to cross reference with our own[16–19,29–32,49]. In doing so, we identified significant overlap in conservation of expression for CEp, CEn as well as the AL, compared to the cornea and trabecular meshwork, respectively. Using WISH, cryosections and two color FWISH we confirmed CEp expression patterns for 14 genes found within the CEp1 and CEp2 clusters (Fig. 7A–B, S12A) that were also identified in human corneal data sets. We show clear co-expression of *zgc92380, sparc, pfn1, epcam* and *col1a1b* with our established CEp marker *krt4* (Fig. 7C). This outcome thus suggests a high degree of conservation between zebrafish and human corneal epithelial development. In addition, we characterized expression of eight genes previously not associated with CEp development including: *cavin2a, col4a5, col5a3b, pnp5a, soul2, si211:133,157, zgc:158,463* and *zgc:175,088* (Fig. S12B). All of which display characteristic expression patterns for corneal epithelium. Furthermore, thanks to the multi timepoint nature of our data, we also analyzed total expression changes for each of these genes during the course of development by comparing relative expression within clusters across developmental time (Fig. 7D). When examining *zgc92380, sparc, pfn1, epcam* and *col1a1b* relative expression levels in CEp1 we see a relatively steady state of expression throughout development with only *epcam* displaying a significant drop in expression at 96 hpf (Fig. 7D). In CEp2 we see a significant spike in expression of all five genes at 72 hpf and a subsequent decline. For CEn associated genes, we confirmed the expression of *matn4*, a known component of the human cornea, and identified *cthrc1a* as a potential novel regulator of CEn development and or function (Fig. 7E–F)[62,63]. Both display spatial expression patterns like those of CEn markers *pmp22a* and *ctnnb2*. When tracking relative expression levels over time, we note that in case of the CEn markers, *pmp22a* is initially relatively strongly expressed, but its expression decreases a 120 hpf and slightly rebounds at 144 hpf. Whereas the expression of *ctnnb2* is relatively low and further decreases at 96 hpf (Fig. 7G). *Matn4* expression declines by 96 hpf with a slight rebound by 144 hpf. *Cthrc1a* follows a similar pattern with a gradual decline, most evident at 96 hpf, and a rebound by 144 hpf (Fig. 7G).

Very few genes were previously known to be expressed in the AL of zebrafish[51–53]. Of those known, such as *scinla*, these genes were not specific markers of the annular ligament, as they are also found expressed in adjacent tissues, such as the corneal endothelium. Hence, this study provides the first extensive gene expression analysis of the zebrafish annular ligament and establishes a set of marker genes more specific to this tissue. Using WISH, we confirmed expression of 31 genes found in AL1 and AL2 clusters that are also known to be expressed in human TM (Figs. 8, S13A). Of those 31, ten were further characterized using cryosection as well as co-expression analysis with *scinla* using two color FWISH. Five of the genes characterized were associated with AL1, which included *nusap1, frzb, cxcr4b, rrm2* and *cdh5* (Fig. 8A–B), while the other five, *mcm7, vim, lxn, fmoda* and *fabp11a* were associated with AL2 (Fig. 8E–F). Cryosection images confirm the distinction as AL2 genes tend to display both AL and corneal endothelium expression patterns (Fig. 8B, F). Furthermore, 3D and 2D confocal images clearly confirm co-expression of these genes with *scinla* in AL tissues (Fig. 8C,G). Of the genes examined, we observed that *frzb* expression was the only one displaying dorsally restriction, which suggests it may have a specialized function related to AL physiology (Fig. 8B–C). When tracking mean expression within AL clusters over time, we observe steady state levels of expression for our canonical markers *hmgn2, hmgb2b, scinla* and *myoc* with a slight increase at 72 hpf and 144 hpf (Fig. 8D). In case of *rrm2*, the expression increases constantly, but slightly drops at 144 hpf. *Nusap1* on the other hand, is initially relatively weakly expressed, but strongly increases after
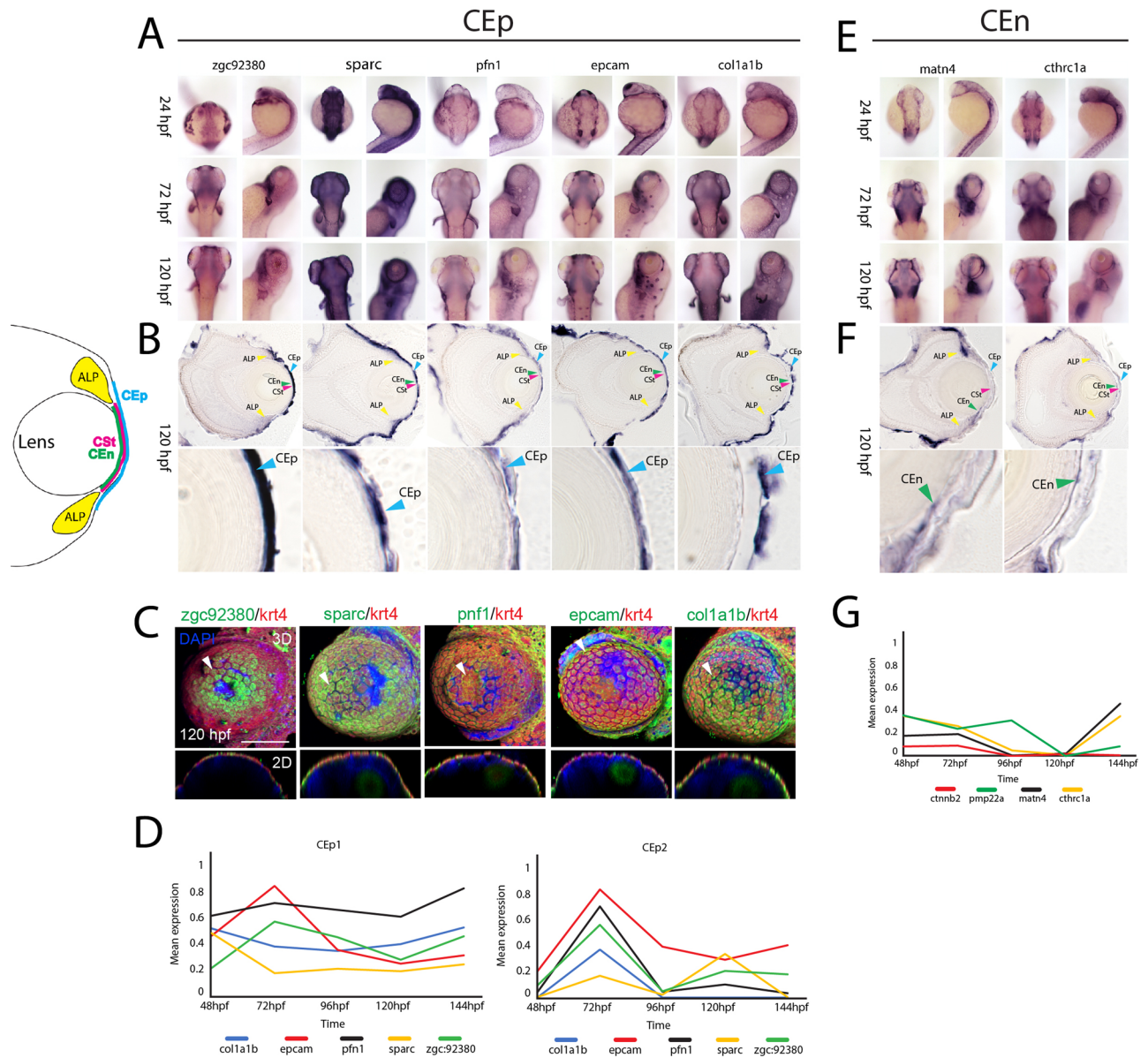
Figure 7. Spatiotemporal expression patterns of conserved corneal epithelium and endothelium associated genes. (**A**) WISH analysis of corneal epithelium associated genes *zgc:92,380, sparc, pfn1, epcam* and *col1a1b* expression at 24, 72 and 120 hpf displayed in dorsal and lateral views. (**B**) 10 μm cryosections of 120 hpf embryo eyes after WISH. (**C**) Two color FWISH depicting co-expression of *krt4*(red)/*zgc:92,380*(green), sparc(green)/*krt4*(red), *pfn1*(green)/*krt4*(red), *epcam*(green)/*krt4*(red) and *col1a1b*(green)/*krt4*(red) at 120 hpf. DAPI (blue) was used to stain the nuclei. Volume projections in lateral view and single section dorsal views from 3D confocal stacks are displayed. White arrows indicate example regions of co-expression. Scale bar = 100 μm. Individual fluorescence channels are provided in Fig. S10. (**D**) Average mean expression measurements from CEp1 and CEp2 clusters for marker genes *krt4, krt5, krt91* and *icn2* as well as *col1a1b, epcam, pfn1, sparc* and *zgc92380* at each timepoint examined, 48–144 hpf. (**E**) WISH analysis of corneal endothelium associated genes *matn4* and *cthrc1a* expression at 24, 72 and 120 hpf displayed in dorsal and lateral views. (**F**) 10 μm cryosections of 120 hpf embryo eyes after WISH. (**G**) Average mean expression measurements from CEn clusters for *pmp22a, ctnnb2, matn4* and *cthrc1a* at each timepoint examined, 48–144 hpf. *ALP* Annular ligament progenitor (yellow arrow), *CEp* Corneal epithelium (blue arrow), *CEn* Corneal endothelium (green arrow), *CSt* Corneal stroma (green arrow).

96 hpf (Fig. 8D). The expression of both, *mcm7* and *fabp11a* increases until 96 hpf, but while mcm7 initially drops at 120 hpf and then rises again, *fabp11a* remains at a constant level (Fig. 8H). *Fmoda* and *rn7sk* are both expressed up to 72 hpf, but their expression decreases in the following time points (Fig. 8H). In addition to AL genes conserved in TM tissues, we also identified 13 potentially novel components of the AL (Fig. 9, S13B). These included *stmn1a, si:ch211-251b21.1, hgd, phghd* and *cndp1* (Fig. 9A). All five of these genes display AL1 characteristic expression, strong staining in the AL regions and absence from the cornea and co-expressing with
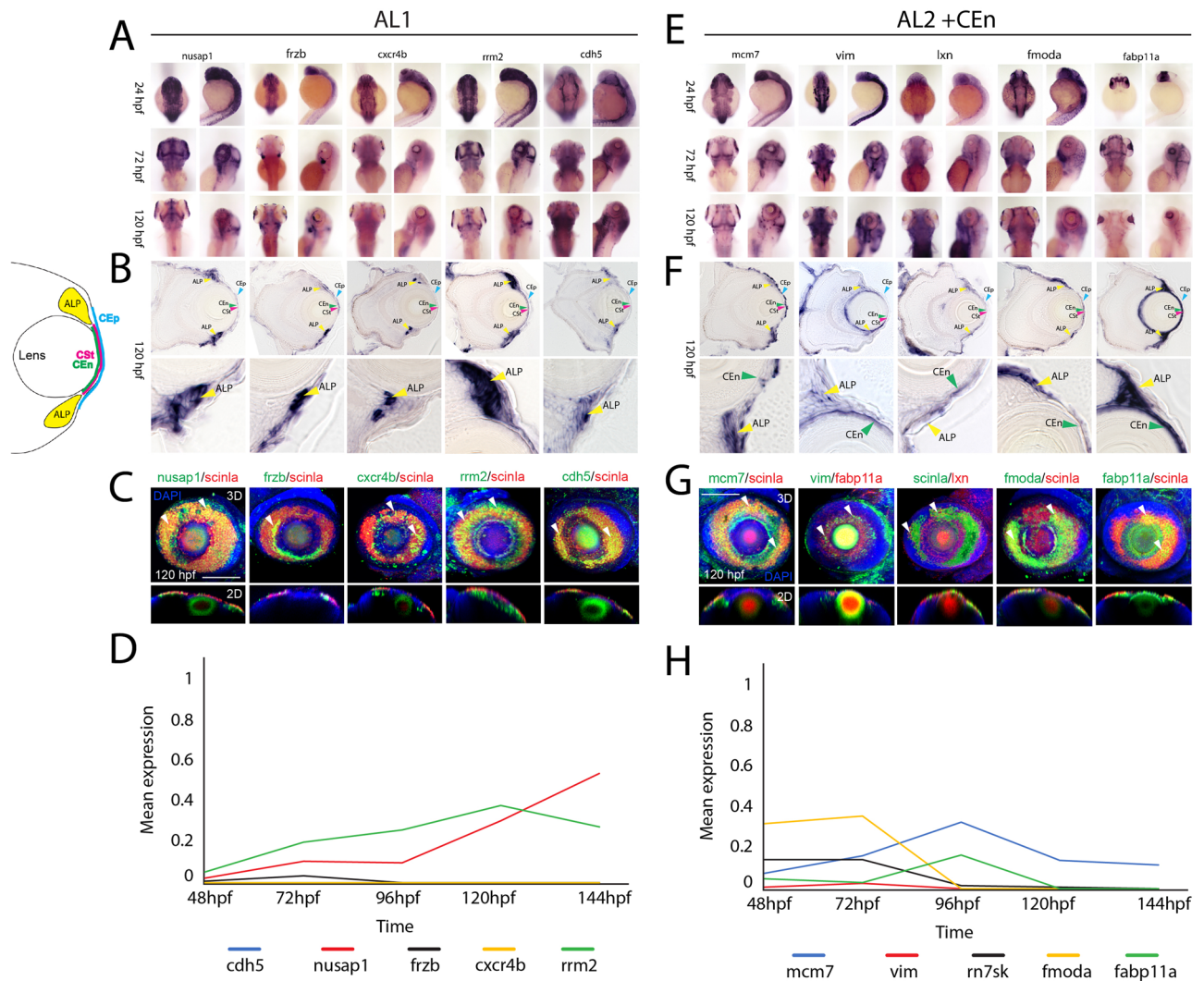
Figure 8. Spatiotemporal expression patterns of conserved annular ligament associated genes. (**A**) WISH expression analysis of annular ligament 1 (AL1) associated genes *nusap1, frzb, cxcr4b, rrm2* and *cdh5* at 24, 72 and 120 hpf displayed in dorsal and lateral views. (**B**) 10 µm cryosections of 120 hpf embryo eyes after WISH. (**C**) Two color FWISH depicting co-expression of *nusap1*(green)/*scinla*(red), *frzb*(green)/*scinla*(red), *cxcr4b*(green)/*scinla*, *rrm2*(green)/*scinla*(red) and *cdh5*(green)/*scinla*(red) at 120 hpf. DAPI (blue) was used to stain the nuclei. Volume projections in lateral view and single section dorsal views from 3D confocal stacks are displayed. White arrows indicate example regions of co-expression. Scale bar = 100 µm. Individual fluorescence channels are provided in Fig. S11A. (**D**) Average mean expression measurements from AL1 clusters for genes *nusap1, frzb, cxcr4b rrm2*, and *cdh5* at each timepoint examined, 48–144 hpf. (**E**) WISH expression analysis of annular ligament 2 (AL2) associated genes *mcm7, vim, lxn, fmoda* and *fabp11a* at 24, 72 and 120 hpf displayed in dorsal and lateral views. (**F**) 10 µm cryosections of 120 hpf embryo eyes after WISH. (**G**) Two color FWISH depicting co-expression of *mcm7*(green)/*scinla*(red), *vim*(green)/*scinla*(red), *lxn*(green)/*scinla*, *fmoda*(green)/*scinla*(red) and *fabp11a*(green)/*scinla*(red) at 120 hpf. DAPI (blue) was used to stain the nuclei. Volume projections in lateral view and single section dorsal views from 3D confocal stacks are displayed. White arrows indicate example regions of co-expression. Scale bar = 100 µm. Individual fluorescence channels are provided in Fig. S11B. (**H**) Average mean expression measurements from AL2 clusters for genes *mcm7, vim, lxn, fmoda* and *fabp11a* over developmental time. *ALP* Annular ligament progenitor (yellow arrow), *CEp* corneal epithelium (blue arrow), *CEn* Corneal endothelium (green arrow), *CSt* Corneal stroma (green arrow).

scinla as shown using FWISH (Fig. 9A–C). When tracking their mean expression patterns over time we note that *cndp1, phgdh* and *hgd* display steady state expression, *stmn1a* expression gradually increases over time while *si211-251b21.1* peaks at 72 hpf and gradually declines thereafter (Fig. 9D). Taken together, we not only confirm significant conservation of gene expression between human TM and developing zebrafish AL, but we also identify several new target genes as potential constituents of the AL. While we examined a large number of genes from our data set, many more remain to be validated using WISH and additional techniques.
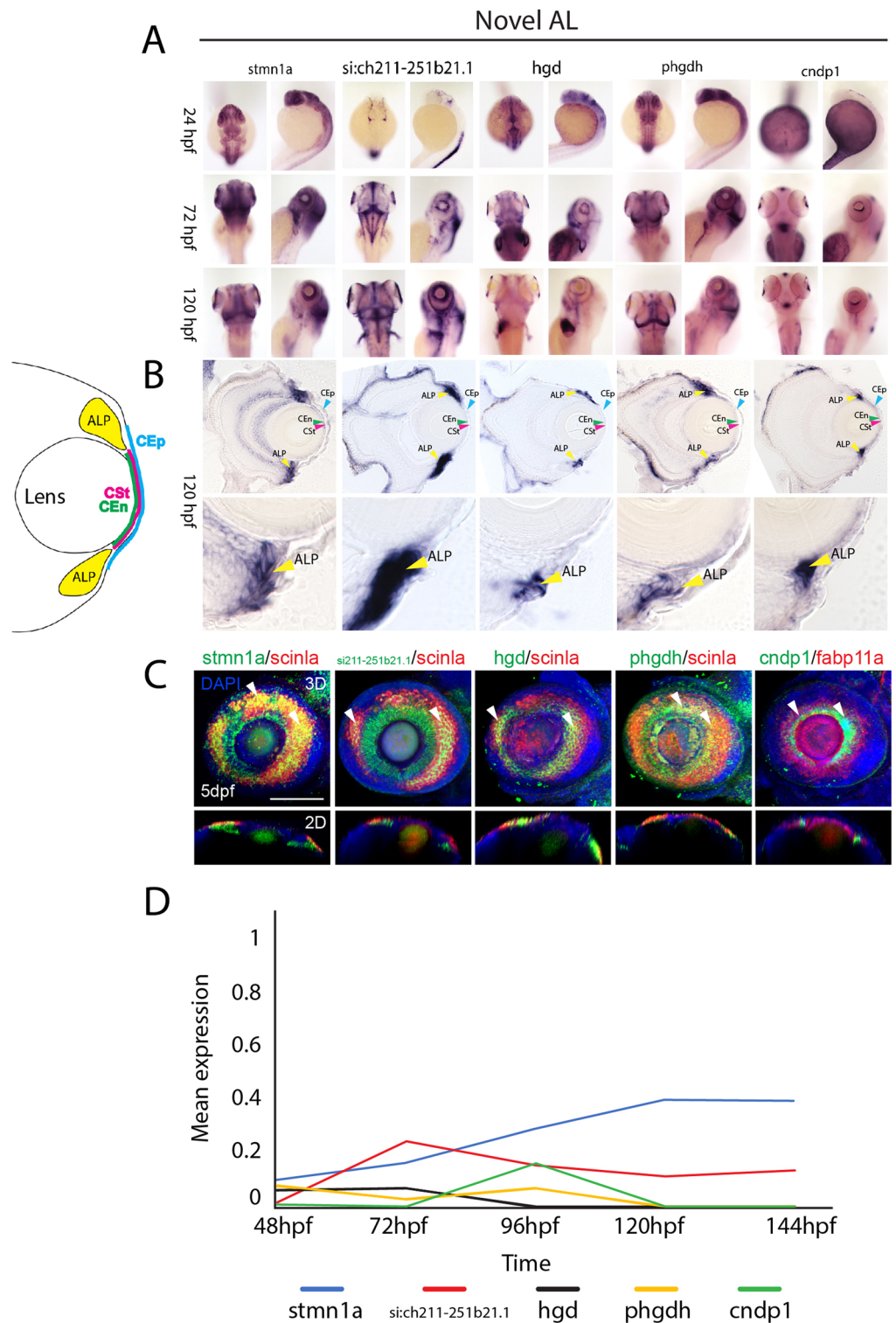
Figure 9. Expression patterns of novel annular ligament associated genes. (**A**) WISH expression analysis of annular ligament associated genes *stmn1a, si:211-251b21.1, hgd, phgdh* and *cndp1* at 24, 72 and 120 hpf displayed in dorsal and lateral views. (**B**) 10 µm cryosections of 120 hpf embryo eyes after WISH. (**C**) Two color FWISH depicting co-expression of *stmn1a*(green)/*scinla*(red), *si:ch211-252b21.1*(green)/*scinla*(red), *hgd*(green)/*scinla, phgdh*(green)/*scinla*(red) and *cndp1*(green)/*scinla*(red) at 120 hpf. DAPI (blue) was used to stain the nuclei. Volume projections in lateral view and single section dorsal views from 3D confocal stacks are displayed. White arrows indicate example regions of co-expression. Scale bar = 100 µm. Individual fluorescence channels are provided in Fig. S11C. (**D**) Average mean expression measurements over developmental time. *ALP* Annular ligament progenitor (yellow arrow), *CEp* Corneal epithelium (blue arrow), *CEn* Corneal endothelium (green arrow), *CSt* Corneal stroma (green arrow).

Discussion

In this study, we present a spatiotemporal analysis of 12,234 single cell transcriptomes of anterior segment tissues of zebrafish isolated during initial AS formation. While recent single cell transcriptomic studies focused on the analysis of mature tissues, our analysis focuses on early development. This gives us the unique opportunity to characterize and follow gene expression changes throughout multiple timepoints of early AS development. The key points of our analysis are the establishment of a defined set of marker genes for developing anterior segment tissues in zebrafish, the identification of multiple cell types in the cornea and annular ligament of zebrafish, establishment of developmental trajectories for corneal and annular ligament lineages and validating conservation of zebrafish AS gene expression with mammalian and human AS tissues.

Establishing a set of marker genes for anterior segment structures in zebrafish. The focus of our analysis were the cornea and the annular ligament. Since developmentally the molecular composition of both tissues in zebrafish, especially the annular ligament, was limited to a handful of associated genes, we first had to establish a solid base of tissue marker genes. For the corneal epithelium we found that the keratins *krt4,5, 91* and also *icn2* to be excellent markers, which for the most part have also described in adult tissue of several other species[46–49,64,65]. For the corneal endothelium we found that *ctnnb2* as well as *pmp22a* are already expressed in the early corneal tissues, confirming the finding of *ctnnb2* in zebrafish at later stages[45,49]. Additionally, *lum* and *dcn* are expressed in the developing corneal stroma, as reported in other organisms[11]. Regarding the annular ligament, we identified *hmgb2b, hmgn2, mcm7, pcna* as well as *scinla* and *lxn* to be suitable expressed markers, of which only *scinla* was previously known to be expressed in the zebrafish tissue[51–53]. Establishment of genetic markers for developing zebrafish AS tissues not only helps with their identification in transcriptomic data sets, but also enables comparisons to other model organisms. It also serves as a resource for performing subsequent unbiased analysis on cells isolated from whole eye tissues.

Identification of different cell types in the cornea and annular ligament. Analysis of our combined dataset revealed a total of 21 clusters, indicating the presence of multiple cell types. AS related cell types included the annular ligament, corneal endothelium, corneal stroma and also corneal epithelium. The finding of corneal epithelium cells in our dataset comes as a surprise, considering that corneal epithelial tissues are derived from ectoderm and hence, epidermal tissues. Since our data is based on transcriptomes from GFP+ cells isolated from eyes of Tg[foxc1b:GFP] embryos, and Foxc1b is a well-established marker of the periocular mesenchyme and ASM, we did not expect to detect corneal epithelium cells[24,39,66]. POM and ASM cells are thought to be involved in forming anterior tissues such as the corneal endothelium, corneal stroma, the trabecular meshwork, ciliary body and iris stroma, but not the corneal epithelium. There are two possible explanations for this contradictory finding. Either, a subpopulation of corneal epithelial cells could indeed originate from neural crest cells, or this is an artefact of our transgenic line, which marks at least some corneal epithelial cells. Since ours is the first comprehensive molecular examination of POM/ASM lineage tracing, perhaps a previously unidentified population of POM does differentiate and incorporate into corneal epithelium. Support for this hypothesis stems from the fact that even at 48 hpf, we detect a significant number of epithelial like cells. However, we consider the latter possibility of CEp cells originating as an artifact of the transgenic line more likely. Perhaps there is leaky expression from the transgene in corneal epithelium, as we do observe GFP+ cells, particularly at older stages, anterior of the lens (Fig. 1A). Yet, at this point we cannot confidently rule out either possibility. Of the cells we identify as corneal epithelium they are represented by two clusters CEp1 and CEp2. Furthermore, determining if these clusters in fact represent two cell types is difficult. On the one hand, several marker genes such as *krt5* can be found in both clusters at the early time points but on the other, there appear to be functional differences between CEp1 and CEp2. While CEp1 mostly expresses keratins and collagens, as would be expected for this tissue, we found several markers of the immune system and stress response in CEp2, such as *anxa1a, b, c* and *serp1*[29,67]. Accordingly, CEp2 might represent a corneal immune cell type, as has been described in other organisms[68].

Corneal endothelium cells were observed to associate with the corneal stroma since both cell types closely intertwined and expressing very similar genes during early stages. In fact, we were unable to identify an independent cluster of corneal stroma cells in any of the analyzed time points. This comes as no surprise, since the stroma is known to be populated with a second wave of POM cells, after an initial wave forms the foundation of the endothelium[10]. We found one cluster representing corneal endothelium and stroma, though it diverged into two clusters at 144 hpf. Additionally, we found another cluster of the corneal endothelium in combination with the annular ligament cluster AL2. The major difference between the two is that the former strongly expresses markers such as *pmp22a* and *dcn*, the latter was found to also upregulate the expression of annular ligament markers such as *scinla* and *hgd*. *Pmp22* is critically involved in the myelination of neurons and mutations can cause Charcot-Marie-Tooth disease[69,70]. This may result in ocular manifestations such as optic atrophy but can also affect the cornea[71]. *Dcn* on the other hand, is a multifunctional part of the extracellular matrix with a variety of binding partners that contributes to general stability, but is also involved in receptor signaling[72,73]. In the cornea dcn regulates transparency[74]. Scinla is a member of the gelsolin family and is described as a crystallin in zebrafish[51,52,75]. A knockdown with morpholinos in zebrafish resulted in misregulated dorsalization of the embryos and a functional involvement with actin filament regulation has been reported[53,75]. Hgd, or homogentisate 1,2-dioxygenase, is an important part of the tyrosine metabolism and mutations can cause a disease called alkaptoneuria, due to accumulation of homogentisate[76,77]. A clear separation of annular ligament and endothelial markers at these early stages is challenging. Particularly, the annular ligament makers are broadly expressed and due to the proximity of these tissues, a clear-cut separation may not be possible at the early developmental stages. In addition to AL2 we also found another group of annular ligament cells named AL1, that separates into

multiple clusters at later time points. Stratification of AL1 may represent differentiation of drainage structures analogous to the Schlemm's canal. While the markers for both, AL1 and AL2 often overlap in expression, there are key differences in gene expression between the two cell types. For instance, the known TM and glaucoma marker *myoc* can only be found in AL2[28,78], just like the marker gene *lxn*, a homolog of *rarres1*, which is associated with the regulation of the fatty acid metabolism[18,79]. AL1 on the other hand strongly expresses genes involved in general development like *mcm7* and *pcna*[80–82]. However, during this early time of development, analysis of functional differences remains highly speculative since the cells are not yet fully differentiated. Nonetheless, finding potentially different cell types within the annular at the early stages is exciting, since this provides a first insight into the mechanisms governing specification and assembly of this tissue.

Trajectory of corneal and annular ligament cells throughout development. By collecting our data at multiple time points throughout early development, we were able to track cellular trajectories over time. To do so, we first manually compared gene expression of all clusters to determine if a potential relation between individual clusters within or among different time points exists. Subsequently, we performed pseudotime analysis of all combined time points, and analyzed variation in gene expression throughout the established trajectories. Manual comparison revealed that both clusters of the corneal epithelium not only share a common origin, but also are distinct from all remaining clusters, including the corneal endothelium and stroma. We found a close relation between the annular ligament and the corneal endothelium and corneal stroma clusters, indicating a common developmental origin of these tissues. When we compare our results to what is known about the physiology of AS development, they can be explained as follows. The corneal epithelium originates from epidermal cells, whereas corneal endothelium, stroma and annular ligament originate from POM cells, or more precisely ASM cells[39,83,84]. The corneal epithelium begins development as epidermal cells overlaying the lens, starting at approximately 16 hpf and is involved in initiating lens development. As mentioned previously, it remains peculiar that our data set identified corneal epithelium signatures since it is solely ASM derived. POM cells are well known for their migratory behavior and populate the anterior segment as ASM in different waves in chick, mouse and humans[37,85,86]. The first wave of POM cells arriving at the anterior segment, migrates into the space between lens and corneal epithelium to form the endothelium. After this initial formation of the endothelium, a second wave arrives to migrate into the stroma. It is not clear at this point if the multi-wave migration of POM and ASM is recapitulated in zebrafish. Furthermore, based on our data it may be that zebrafish derive all of the aforementioned tissues from a common progenitor pool that arrives early in development. Interestingly, recent work in our lab has observed the earliest ASM cells arriving as early as 16 hpf, while the bulk of ASM are observed to arrive approximately 26–28 hpf[39,87]. The results of our current study suggest that annular ligament, as well as corneal endothelium and the cells in the corneal stroma all share a common origin, a subgroup of ASM that are at part of the foxc1b derived lineage. We have previously shown that ASM cells are a heterogenous population during AS colonization, so the multipotent aspect of f*oxc1b* derived ASM may be lineage specific rather than universal amongst all ASM. As such, in zebrafish we can conclude that corneal stroma, endothelium and annular ligament all develop from a common progenitor. Physiologically this also makes sense as stroma, endothelium and the AL share close physical association in context of AS organization. Our pseudotime analysis corroborates our manual cluster trajectories as it shows three major trajectories within our dataset (Fig. 6A). The first trajectory entails corneal epithelial cells and further indicates the common origin of CEp1 and CEp2 (Fig. 6B). The second trajectory contains mainly AL1 cells, whereas the third includes both AL1 and AL2, as well as CEn and CSt, further validating a relation between these tissues. Markers for AL2 such as *scinla, lxn* or *pdgfrl* are mostly found in the third pseudotime trajectory, but not the second (Figs. 6B, S8). *Lxn*, a homolog of *rarres1*, is a carboxypeptidase inhibitor and plays multiple roles during development, including posttranslational modifications of alpha-tubulin, regulation of the mitochondrial membrane potential and control of the fatty acid metabolism[17,88]. *Pdgfrl* is mostly known to be a tumor suppressor gene and plays different roles during development depending on the cell type[89,90]. Since we can find AL1 markers such as *mcm7* and *hmgb2b* in both trajectories (Figs. 6B, S8) it seems possible that there is also a relation between these groups. Both, *mcm7* and *hmgb2b* are involved in cell proliferation and genome stability. While the former is particularly known to be associated with DNA replication, the latter is a highly conserved nuclear protein that acts as a chromatin binding factor[91–94]. Migrating POM cells receive developmental signals early[85,95] and it appears possible that if even younger stages would have been included in this study, the separate trajectories could have joined into one. Accordingly, these results provide support for the hypothesis, that there is a common origin of the corneal endothelium/ stroma and at least parts of the annular ligament.

Overall, our findings corroborate previous studies that pioneered lineage tracing of POM cells in different model organisms[96–100]. These studies found that migratory neural crest cells derived from the hindbrain, are critically important to form structures of the anterior segment. For instance, transplantation experiments in quail and chicken have shown the migration of NCCs into the corneal endothelium and stroma, as well as the iridocorneal angle[96,98–100].

Conservation of genetic identity between the annular ligament and trabecular meshwork. As stated above, prior to this study, our knowledge about the gene expression profile of the zebrafish anterior segment and particularly the annular ligament was very limited. As such, our study provides the first and detailed transcriptomic atlas of the developing zebrafish AS. One of the important aspects of such data is its use for comparative studies. Zebrafish, particularly their embryos, offer a highly versatile model system for modeling human disease and high throughput for functional assays related to developing treatments or interventions. While already accepted as a suitable model for ASD, zebrafish had been seldom used to assess TM biology and therefore predisposition to glaucoma. Having generated said atlas, we were able to not only uncover genes likely

involved in AL development, but also compare their expression to mammalian and importantly human TM tissues. In doing so we confirmed over 31 genes to be expressed in both the TM and AL. This is not an exhaustive list, in fact many more genes should be examined for spatiotemporal expression in zebrafish and expression in TM tissues. However, clearly there is a high degree of molecular similarity between the AL and TM therefore supporting the idea of modeling TM related ASD and glaucoma pathologies. In addition to genes previously shown to associate with TM, we also identified several novel genes that are either AL specific or are not expressed in fully developed TM. Here we want to point out five of those genes previously not associated with the AS, that stood out due to their strong gene expression in our in situ hybridization experiments. *Hgd* and *phgdh* both are part of the pathway of tyrosine and phenylanaline degradation and we found them both to be strongly expressed in the AL and CEn tissues (see Fig. 9). Mutations of *hgd* have already been proven to be the cause of a disease named Alkaptonuria, which can also affect the eye[101,102]. However, their function in the anterior segment is, to our knowledge, not known. *Stmn1a* is involved in microtubule organization and it's homologs have previously been associated with macular degeneration of the retina[103]. However, we found it to be strongly expressed in AL and CEn tissues (Fig. 9). Also, when projected onto publicly available single cell transcriptomes of human tissues of the anterior segment, we found it to be widely expressed, pointing to a potentially broader function also in these tissues. *Cndp1*, a metalloproteinase family member, that has been associated with diabetic nephropathy shows a very specific staining in our study in the annular ligament (Fig. 9)[104]. Organization and likely remodeling of the extracellular matrix are important for TM function; hence, it may be worthwhile to determine the role that *cndp1* plays in these processes. Finally, *si:211-251b21.1* is a predicted kainate receptor. Kainate receptors are known to play important roles in different tissues, including the retina and the cornea[105,106]. Due to its highly enriched expression in the AL (see Fig. 9A), we assume that this particular gene might play an important function for the AL. Future functional studies will need to be carried out to assess the functional contribution of these genes to AL development as well as association with ASD or TM deficiencies.

Zebrafish as a bona fide model for anterior segment disease. The use of zebrafish as a model for different developmental processes is well established. However, some previous studies have postulated that zebrafish should only be used with care, when applying it as a model for ocular diseases, affecting the TM or cornea due to differences in structure and molecular makeup[10,26,107]. On the contrary, other studies suggested that zebrafish could indeed be used as a great model for certain ocular diseases and in fact started using it in this respect[8,20–25]. With this study we aim to provide a base for a better understanding of the anterior segment with focus on the annular ligament and highlight the molecular connection it shares with the TM at early developmental time points. There are undoubtedly differences in morphology and molecular makeup of the zebrafish tissue, compared to the human TM[10,23,26,107]. However, having found that many TM markers described in human tissue are expressed in the AL, we strongly believe that both tissues share developmental and molecular pathways. This hopefully encourages the use of zebrafish as a functional model for disease related to TM function, such as glaucoma[4,7].

Data availability

References

1. Cvekl, A. & Tamm, E. R. Anterior eye development and ocular mesenchyme: New insights from mouse models and human diseases. *BioEssays* **26**(4), 374–386 (2004).
2. Reneker, L. W., Silversides, D. W., Xu, L. & Overbeek, P. A. Formation of corneal endothelium is essential for anterior segment development: A transgenic mouse model of anterior segment dysgenesis. *Development* **127**(3), 533–542 (2000).
3. Swamynathan, S. K. Ocular surface development and gene expression. *J. Ophthalmol.* **2013**, 103947 (2013).
4. Weinreb, R. N., Aung, T. & Medeiros, F. A. The pathophysiology and treatment of glaucoma a review. *Jama-J. Am. Med. Assoc.* **311**(18), 1901–1911 (2014).
5. Wulle, K. G. Electron-microscopy of fetal development of corneal endothelium and descemets membrane of human eye. *Invest. Ophth. Visual.* **11**(11), 897–900 (1972).
6. McMahon, C., Gestri, G., Wilson, S. W. & Link, B. A. Lmx1b is essential for survival of periocular mesenchymal cells and influences Fgf-mediated retinal patterning in zebrafish. *Dev. Biol.* **332**(2), 287–298 (2009).
7. Kang, J. M. & Tanna, A. P. Glaucoma. *Med. Clin. North Am.* **105**(3), 493–510 (2021).
8. Gestri, G., Link, B. A. & Neuhauss, S. C. The visual system of zebrafish and its use to model human ocular diseases. *Dev. Neurobiol.* **72**(3), 302–327 (2012).
9. Soules, K. A. & Link, B. A. Morphogenesis of the anterior segment in the zebrafish eye. *BMC Dev. Biol.* **5**, 12 (2005).
10. Zhao, X. P. C. *et al.* The zebrafish cornea: Structure and development. *Invest. Ophth. Vis. Sci.* **47**(10), 4341–4348 (2006).
11. Espana, E. M. & Birk, D. E. Composition, structure and function of the corneal stroma. *Exp. Eye Res.* **198**, 108137 (2020).
12. DelMonte, D. W. & Kim, T. Anatomy and physiology of the cornea. *J. Cataract Refract. Surg.* **37**(3), 588–598 (2011).
13. Buffault, J., Labbé, A., Hamard, P., Brignole-Baudouin, F. & Baudouin, C. The trabecular meshwork: Structure, function and clinical implications. A review of the literature Le trabéculum: Structure, fonction et implications cliniques Une revue de la littérature. *J. Français d'Ophtalmol.* **43**(7), e217–e230 (2020).
14. Dautriche, C. N., Xie, Y. & Sharfstein, S. T. Walking through trabecular meshwork biology: Toward engineering design of outflow physiology. *Biotechnol. Adv.* **32**(5), 971–983 (2014).
15. Lutjen-Drecoll, E. Functional morphology of the trabecular meshwork in primate eyes. *Prog. Retin. Eye Res.* **18**(1), 91–119 (1999).

16. Callaghan, B. *et al.* Genome-wide transcriptome profiling of human trabecular meshwork cells treated with TGF-beta2. *Sci. Rep.* **12**(1), 9564 (2022).
17. Patel, G. *et al.* Molecular taxonomy of human ocular outflow tissues defined by single-cell transcriptomics. *Proc. Natl. Acad. Sci. U S A.* **117**(23), 12856–12867 (2020).
18. van Zyl, T. *et al.* Cell atlas of aqueous humor outflow pathways in eyes of humans and four model species provides insight into glaucoma pathogenesis. *Proc. Natl. Acad. Sci. U S A.* **117**(19), 10339–10349 (2020).
19. van Zyl, T. *et al.* Cell atlas of the human ocular anterior segment: Tissue-specific and shared cell types. *Proc. Natl. Acad. Sci. U S A.* **119**(29), e2200914119 (2022).
20. Akhtar, S., Schonthaler, H. B., Bron, A. J. & Dahm, R. Formation of stromal collagen fibrils and proteoglycans in the developing zebrafish cornea. *Acta Ophthalmol.* **86**(6), 655–665 (2008).
21. Bibliowicz, J., Tittle, R. K. & Gross, J. M. Toward a better understanding of human eye disease insights from the zebrafish, Danio rerio. *Prog. Mol. Biol. Transl. Sci.* **100**, 287–330 (2011).
22. Blanco-Sanchez, B., Clement, A., Phillips, J. B. & Westerfield, M. Zebrafish models of human eye and inner ear diseases. *Method Cell Biol.* **138**, 415–467 (2017).
23. Link, B. A. & Collery, R. F. Zebrafish models of retinal disease. *Annu. Rev. Vis. Sci.* **1**, 125–153 (2015).
24. Skarie, J. M. & Link, B. A. FoxC1 is essential for vascular basement membrane integrity and hyaloid vessel morphogenesis. *Invest. Ophthalmol. Vis. Sci.* **50**(11), 5026–5034 (2009).
25. Veth, K. N. *et al.* Mutations in zebrafish lrp2 result in adult-onset ocular pathogenesis that models myopia and other risk factors for glaucoma. *PLoS Genet.* **7**(2), e1001310 (2011).
26. Chen, C. C. *et al.* Morphological differences between the trabecular meshworks of zebrafish and mammals. *Curr. Eye Res.* **33**(1), 59–72 (2008).
27. Gray, M. P., Smith, R. S., Soules, K. A., John, S. W. & Link, B. A. The aqueous humor outflow pathway of zebrafish. *Invest. Ophthalmol. Vis. Sci.* **50**(4), 1515–1521 (2009).
28. Atienzar-Aroca, R. *et al.* Transgenic overexpression of myocilin leads to variable ocular anterior segment and retinal alterations associated with extracellular matrix abnormalities in adult zebrafish. *Int. J. Mol. Sci.* **23**(17), 9989 (2022).
29. Gautam, P. *et al.* Multi-species single-cell transcriptomic analysis of ocular compartment regulons. *Nat. Commun.* **12**(1), 5675 (2021).
30. Kaplan, N. *et al.* Single-cell RNA transcriptome helps define the limbal/corneal epithelial stem/early transit amplifying cells and how autophagy affects this population. *Invest. Ophthalmol. Vis. Sci.* **60**(10), 3570–3583 (2019).
31. Li, D. Q. *et al.* Single-cell transcriptomics identifies limbal stem cell population and cell types mapping its differentiation trajectory in limbal basal epithelium of human cornea. *Ocul. Surf.* **20**, 20–32 (2021).
32. Li, J. M. *et al.* Single-cell transcriptomics identifies a unique entity and signature markers of transit-amplifying cells in human corneal limbus. *Invest. Ophthalmol. Vis. Sci.* **62**(9), 36 (2021).
33. Lwigale, P. Y. Corneal development: Different cells from a common progenitor. *Prog. Mol. Biol. Transl. Sci.* **134**, 43–59 (2015).
34. Lwigale, P. Y. & Bronner-Fraser, M. Lens-derived Semaphorin3A regulates sensory innervation of the cornea. *Dev. Biol.* **306**(2), 750–759 (2007).
35. Lwigale, P. Y. & Bronner-Fraser, M. Semaphorin3A/neuropilin-1 signaling acts as a molecular switch regulating neural crest migration during cornea development. *Dev. Biol.* **336**(2), 257–265 (2009).
36. Takamiya, M. *et al.* Pax6 organizes the anterior eye segment by guiding two distinct neural crest waves. *PLoS Genet.* **16**(6), e1008774 (2020).
37. Williams, A. L. & Bohnsack, B. L. The ocular neural crest: Specification, migration, and then what?. *Front. Cell Dev. Biol.* **8**, 595896 (2020).
38. Carrara, N., Weaver, M., Piedade, W. P., Vocking, O. & Famulski, J. K. Temporal characterization of optic fissure basement membrane composition suggests nidogen may be an initial target of remodeling. *Dev. Biol.* **452**(1), 43–54 (2019).
39. Van Der Meulen, K. L., Vocking, O., Weaver, M. L., Meshram, N. N. & Famulski, J. K. Spatiotemporal characterization of anterior segment mesenchyme heterogeneity during zebrafish ocular anterior segment development. *Front. Cell Dev. Biol.* **8**, 379 (2020).
40. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**(7745), 496–502 (2019).
41. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods.* **14**(10), 979–982 (2017).
42. Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**(5), 421–427 (2018).
43. Vocking, O., Van Der Meulen, K., Patel, M. K. & Famulski, J. K. Zebrafish anterior segment mesenchyme progenitors are defined by function of tfap2a but not sox10. *Differentiation* **130**, 32–42 (2022).
44. Sonam, S. *et al.* Cellular and molecular profiles of larval and adult Xenopus corneal epithelia resolved at the single-cell level. *Dev. Biol.* **491**, 13–30 (2022).
45. Takamiya, M. *et al.* Molecular description of eye defects in the zebrafish Pax6b mutant, sunrise, reveals a Pax6b-dependent genetic network in the developing anterior chamber. *PLoS ONE* **10**(2), e0117645 (2015).
46. Kurpakus, M. A., Maniaci, M. T. & Esco, M. Expression of keratins K12, K4 and K14 during development of ocular surface epithelium. *Curr. Eye Res.* **13**(11), 805–814 (1994).
47. Ligocki, A. J. *et al.* Molecular characteristics and spatial distribution of adult human corneal cell subtypes. *Sci. Rep.* **11**(1), 16323 (2021).
48. Tanifuji-Terai, N., Terai, K., Hayashi, Y., Chikama, T. & Kao, W. W. Expression of keratin 12 and maturation of corneal epithelium during development and postnatal growth. *Invest. Ophthalmol. Vis. Sci.* **47**(2), 545–551 (2006).
49. Collin, J. *et al.* A single cell atlas of human cornea that defines its development, limbal progenitor cells and their interactions with the immune cells. *Ocul. Surf.* **21**, 279–298 (2021).
50. Borras, T. *et al.* Mechanisms of AAV transduction in glaucoma-associated human trabecular meshwork cells. *J. Gene Med.* **8**(5), 589–602 (2006).
51. Jia, S., Nakaya, N. & Piatigorsky, J. Differential expression patterns and developmental roles of duplicated scinderin-like genes in zebrafish. *Dev. Dyn.* **238**(10), 2633–2640 (2009).
52. Jia, S. *et al.* Duplicated gelsolin family genes in zebrafish: A novel scinderin-like gene (scinla) encodes the major corneal crystallin. *FASEB J.* **21**(12), 3318–3328 (2007).
53. Yoshikawa, S., Norcom, E., Nakamura, H., Yee, R. W. & Zhao, X. P. C. Transgenic analysis of the anterior eye-specific enhancers of the zebrafish gelsolinlike 1 (gsnl1) gene. *Dev. Dynam.* **236**(7), 1929–1938 (2007).
54. Borras, T. Gene expression in the trabecular meshwork and the influence of intraocular pressure. *Prog. Retin. Eye Res.* **22**(4), 435–463 (2003).
55. Borras, T. The effects of myocilin expression on functionally relevant trabecular meshwork genes: A mini-review. *J. Ocul. Pharmacol. Ther.* **30**(2–3), 202–212 (2014).
56. Tabak, S., Schreiber-Avissar, S. & Beit-Yannai, E. Trabecular meshwork's collagen network formation is inhibited by non-pigmented ciliary epithelium-derived extracellular vesicles. *J. Cell Mol. Med.* **25**(7), 3339–3347 (2021).
57. Xue, W., Wallin, R., Olmsted-Davis, E. A. & Borras, T. Matrix GLA protein function in human trabecular meshwork cells: Inhibition of BMP2-induced calcification process. *Invest. Ophth. Vis. Sci.* **47**(3), 997–1007 (2006).

58. Korhonen, E. A. *et al.* Tie1 controls angiopoietin function in vascular remodeling and inflammation. *J. Clin. Invest.* **126**(9), 3495–3510 (2016).

59. Yasuno, K. *et al.* Common variant near the endothelin receptor type A (EDNRA) gene is associated with intracranial aneurysm risk. *P. Natl. Acad. Sci. USA* **108**(49), 19707–19712 (2011).

60. Iozumi, K., Hoganson, G. E., Pennella, R., Everett, M. A. & Fuller, B. B. Role of tyrosinase as the determinant of pigmentation in cultured human melanocytes. *J. Invest. Dermatol.* **100**(6), 806–811 (1993).

61. Lister, J. A. Larval but not adult xanthophore pigmentation in zebrafish requires GTP cyclohydrolase 2 (gch2) function. *Pigment Cell Melanoma Res.* **32**(5), 724–727 (2019).

62. Klatt, A. R. *et al.* Molecular structure, processing, and tissue distribution of matrilin-4. *J. Biol. Chem.* **276**(20), 17267–17275 (2001).

63. Szalai, E. *et al.* Fibrillin-2, tenascin-C, matrilin-2, and matrilin-4 are strongly expressed in the epithelium of human granular and lattice type I corneal dystrophies. *Mol. Vis.* **18**, 1927–1936 (2012).

64. Hayashi, Y. *et al.* Monoallelic expression of Krt12 gene during corneal-type epithelium differentiation of limbal stem cells. *Invest. Ophthalmol. Vis. Sci.* **51**(9), 4562–4568 (2010).

65. Irvine, A. D. *et al.* A novel mutation in KRT12 associated with Meesmann.s epithelial corneal dystrophy. *Brit. J. Ophthalmol.* **86**(7), 729–732 (2002).

66. Weigele, J. & Bohnsack, B. L. Genetics underlying the interactions between neural crest cells and eye development. *J. Dev. Biol.* **8**(4), 26 (2020).

67. Gavins, F. N. & Hickey, M. J. Annexin A1 and the regulation of innate and adaptive immunity. *Front. Immunol.* **3**, 354 (2012).

68. Liu, J. & Li, Z. Resident innate immune cells in the cornea. *Front Immunol.* **12**, 620284 (2021).

69. Li, J., Parker, B., Martyn, C., Natarajan, C. & Guo, J. The PMP22 gene and its related diseases. *Mol. Neurobiol.* **47**(2), 673–698 (2013).

70. Naef, R. & Suter, U. Many facets of the peripheral myelin protein PMP22 in myelination and disease. *Microsc. Res. Tech.* **41**(5), 359–371 (1998).

71. Moshirfar, M., Tukan, A. N., Bundogji, N. & Ronquillo, Y. C. Charcot-marie-tooth disease and implications on corneal refractive surgery. *Ophthalmol. Ther.* **11**(4), 1301–1307 (2022).

72. Gubbiotti, M. A., Vallet, S. D., Ricard-Blum, S. & Iozzo, R. V. Decorin interacting network: A comprehensive analysis of decorin-binding partners and their versatile functions. *Matrix Biol.* **55**, 7–21 (2016).

73. Neill, T., Schaefer, L. & Iozzo, R. V. Decorin: A guardian from the matrix. *Am. J. Pathol.* **181**(2), 380–387 (2012).

74. Mohan, R. R., Tovey, J. C., Gupta, R., Sharma, A. & Tandon, A. Decorin biology, expression, function and therapy in the cornea. *Curr. Mol. Med.* **11**(2), 110–128 (2011).

75. Kanungo, J., Kozmik, Z., Swamynathan, S. K. & Piatigorsky, J. Gelsolin is a dorsalizing factor in zebrafish. *Proc. Natl. Acad. Sci. U S A* **100**(6), 3287–3292 (2003).

76. Felbor, U., Mutsch, Y., Grehn, F., Muller, C. R. & Kress, W. Ocular ochronosis in alkaptonuria patients carrying mutations in the homogentisate 1,2-dioxygenase gene. *Br. J. Ophthalmol.* **83**(6), 680–683 (1999).

77. Wilson, P. J. M., Ranganath, L. R., Bou-Gharios, G., Gallagher, J. A. & Hughes, J. H. Expression of tyrosine pathway enzymes in mice demonstrates that homogentisate 1,2-dioxygenase deficiency in the liver is responsible for homogentisic acid-derived ochronotic pigmentation. *JIMD Rep.* **58**(1), 52–60 (2021).

78. Wang, H. *et al.* Physiological function of myocilin and its role in the pathogenesis of glaucoma in the trabecular meshwork (Review). *Int. J. Mol. Med.* **43**(2), 671–681 (2019).

79. Maimouni, S. *et al.* Tumor suppressor RARRES1: A novel regulator of fatty acid metabolism in epithelial cells. *PLoS ONE* **13**(12), e0208756 (2018).

80. Fujita, M., Kiyono, T., Hayashi, Y. & Ishibashi, M. HCDC47, a human member of the MCM family. Dissociation of the nucleus-bound form during S phase. *J. Biol. Chem.* **271**(8), 4349–4354 (1996).

81. Kelman, Z. PCNA: Structure, functions and interactions. *Oncogene* **14**(6), 629–640 (1997).

82. Strzalka, W. & Ziemienowicz, A. Proliferating cell nuclear antigen (PCNA): a key factor in DNA replication and cell cycle regulation. *Ann. Bot.* **107**(7), 1127–1140 (2011).

83. Griswold, S. L. & Lwigale, P. Y. Analysis of neural crest migration and differentiation by cross-species transplantation. *J. Vis. Exp.* **60**, 3622 (2012).

84. Ma, J. & Lwigale, P. Transformation of the transcriptomic profile of mouse periocular mesenchyme during formation of the embryonic cornea. *Invest. Ophthalmol. Vis. Sci.* **60**(2), 661–676 (2019).

85. Akula, M., Park, J. W. & West-Mays, J. A. Relationship between neural crest cell specification and rare ocular diseases. *J. Neurosci. Res.* **97**(1), 7–15 (2019).

86. Quantock, A. J. & Young, R. D. Development of the corneal stroma, and the collagen-proteoglycan associations that help define its structure and function. *Dev. Dyn.* **237**(10), 2607–2621 (2008).

87. Vocking, O., Van Der Meulen, K., Patel, M. M., Famulski, J. Zebrafish anterior segment mesenchyme progenitors are defined by function of tfap2a but not sox10. bioRxiv. 2022:2022.10.20.513081.

88. Chen, A., Lee, K. & He, J. C. Autocrine and paracrine effects of a novel podocyte gene, RARRES1. *Kidney Int.* **100**(4), 745–747 (2021).

89. Guo, F. J. *et al.* Expression and functional characterization of platelet-derived growth factor receptor-like gene. *World J. Gastroenterol.* **16**(12), 1465–1472 (2010).

90. Kawata, K. *et al.* A tumor suppressor gene product, platelet-derived growth factor receptor-like protein controls chondrocyte proliferation and differentiation. *J. Cell Biochem.* **118**(11), 4033–4044 (2017).

91. Ha Thi, B. M. *et al.* Microarray analysis of cell cycle gene expression in adult human corneal endothelial cells. *PLoS ONE* **9**(4), e94349 (2014).

92. Hock, R., Furusawa, T., Ueda, T. & Bustin, M. HMG chromosomal proteins in development and disease. *Trends Cell Biol.* **17**(2), 72–79 (2007).

93. Shin, Y. J. *et al.* High-mobility group box 2 (HMGB2) modulates radioresponse and is downregulated by p53 in colorectal cancer cell. *Cancer Biol. Ther.* **14**(3), 213–221 (2013).

94. Wang, Q. *et al.* Heterogeneity of human corneal endothelium implicates lncRNA NEAT1 in Fuchs endothelial corneal dystrophy. *Mol. Ther. Nucleic Acids.* **27**, 880–893 (2022).

95. Prasad, M. S., Charney, R. M. & Garcia-Castro, M. I. Specification and formation of the neural crest: Perspectives on lineage segregation. *Genesis* **57**(1), e23276 (2019).

96. Creuzet, S., Couly, G. & Le Douarin, N. M. Patterning the neural crest derivatives during development of the vertebrate head: Insights from avian studies. *J. Anat.* **207**(5), 447–459 (2005).

97. Gould, D. B., Smith, R. S. & John, S. W. Anterior segment development relevant to glaucoma. *Int. J. Dev. Biol.* **48**(8–9), 1015–1029 (2004).

98. Johnston, M. C., Noden, D. M., Hazelton, R. D., Coulombre, J. L. & Coulombre, A. J. Origins of avian ocular and periocular tissues. *Exp. Eye Res.* **29**(1), 27–43 (1979).

99. Noden, D. M. An analysis of migratory behavior of avian cephalic neural crest cells. *Dev. Biol.* **42**(1), 106–130 (1975).

100. Smith, R. S., Zabaleta, A., Savinova, O. V. & John, S. W. The mouse anterior chamber angle and trabecular meshwork develop without cell death. *BMC Dev. Biol.* **1**, 3 (2001).
101. Moshirfar, M., Kuang, G. T., Ronquillo, Y. Ocular Manifestations of alkaptonuria. StatPearls. Treasure Island (FL) 2022.
102. Vilboux, T. *et al.* Mutation spectrum of homogentisic acid oxidase (HGD) in alkaptonuria. *Hum. Mutat.* **30**(12), 1611–1619 (2009).
103. Ethen, C. M., Reilly, C., Feng, X., Olsen, T. W. & Ferrington, D. A. The proteome of central and peripheral retina with progression of age-related macular degeneration. *Invest. Ophthalmol. Vis. Sci.* **47**(6), 2280–2290 (2006).
104. Boldyrev, A. A., Aldini, G. & Derave, W. Physiology and pathophysiology of carnosine. *Physiol. Rev.* **93**(4), 1803–1845 (2013).
105. Borghuis, B. G., Looger, L. L., Tomita, S. & Demb, J. B. Kainate receptors mediate signaling in both transient and sustained OFF bipolar cell pathways in mouse retina. *J. Neurosci.* **34**(18), 6128–6139 (2014).
106. Ibitokun, B. O. & Miller, K. E. CNQX blocks kainate-induced corneal irritation. *Invest. Ophth. Vis. Sci.* **53**(14), 1855 (2012).
107. Puzzolo, D. *et al.* Structural, ultrastructural, and morphometric study of the zebrafish ocular surface: a model for human corneal diseases?. *Curr. Eye Res.* **43**(2), 175–185 (2018).

Acknowledgements

Author contributions

O.V performed all the experiments. O.V and J.K.F prepared and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-32212-4.

Correspondence and requests for materials should be addressed to J.K.F.

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.