

An iPSC-based in vitro model recapitulates human thymic epithelial development and multi-lineage specification

Received: 21 October 2024

Accepted: 22 July 2025

Published online: 25 August 2025



Yann Preteemer¹, Yuxian Gao^{1,2}, Kaho Kanai^{1,3}, Takuya Yamamoto^{1,4,5}, Kohei Kometani¹, Manami Ozaki^{1,3}, Karin Nishigishi^{1,3}, Tadashi Ikeda⁶, Huaigeng Xu^{1,7}, Akitsu Hotta¹ & Yoko Hamazaki^{1,2} ✉

Thymic epithelial cells (TEC) are crucial in supporting T cell development, but their high heterogeneity and difficulty of isolation pose obstacles to their study in humans. Particularly, how diverse TEC lineages arise from a common progenitor remains poorly understood. To address this, here we establish a human iPSC-based model of thymus organogenesis capable of deriving these lineages in vitro. Through controlled retinoid signaling followed by self-directed differentiation, we obtain FOXN1⁺ TEC progenitor-like cells and diverse mature MHCII⁺ populations resembling cortical and medullary TECs, allowing us to infer their developmental trajectories. Upon thymocyte co-culture, induced TECs support the generation of naïve T cells with diverse TCR repertoires and further develop into *AIRE*⁺ and mimetic TEC subpopulations. Our system provides a fully in vitro model of human TEC differentiation from early fate specification to late-stage maturation, offering new insights into human thymus development and potential regenerative applications for congenital thymic disorders.

Thymic epithelial cells (TEC) constitute the major stromal cell compartment in the thymus and are crucial in supporting the differentiation and selection of T cells^{1–3}. They arise from the ventral side of the third pharyngeal pouch (3rd PP) endoderm, where the thymic anlage forms and starts to express the TEC master regulator *FOXN1* by 6 weeks of gestation in humans^{4–6}. As surrounding mesenchymal cells and lymphoid progenitors from the bone marrow start to populate the early thymic anlage, FOXN1⁺ TEC progenitors (TEP) give rise to cortical TECs (cTEC), which facilitate the lineage commitment and positive selection of thymocytes, and medullary TECs (mTEC), which are responsible for negative selection and regulatory T cell (Treg) generation to ensure self-tolerance^{1,4,7}. Recent murine studies have revealed a remarkable heterogeneity of TEC subpopulations crucial for

thymic function, including MHCII^{high} cells (mTEC^{high}) variably expressing Aire and peripheral tissue antigens, MHCII^{low} cells (mTEC^{low}) secreting Ccl21 and other cytokines for thymocyte chemoattraction, intertypical TECs with both cTEC and mTEC characteristics, and various mimetic mTECs imitating extrathymic cell types, each with distinct molecular characteristics^{8–15}. Defects in specific TEC subpopulations lead to impaired T cell development and autoimmunity^{16,17}.

However, how this stunning diversity of TECs arises during human development remains unclear. Particularly, upstream signals and molecular mechanisms of *FOXN1* leading to TEC fate specification in the 3rd PP are only poorly understood¹⁸, and the developmental processes underlying the establishment of specific lineage identities in

¹Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan. ²Laboratory of Immunobiology, Graduate School of Medicine, Kyoto University, Kyoto, Japan. ³Department of Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan. ⁴Institute for the Advanced Study of Human Biology, Kyoto University, Kyoto, Japan. ⁵RIKEN Center for Advanced Intelligence Project, Kyoto, Japan. ⁶Department of Cardiovascular Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan. ⁷Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California, San Francisco, San Francisco, CA, USA. ✉ e-mail: yoko.hamazaki@cira.kyoto-u.ac.jp

TEC subpopulations have not been clarified. Since abnormalities in TEC development cause immunodeficiencies and autoimmune diseases with high morbidity and mortality^{19–21}, a deeper understanding of these processes is of critical importance to improve patient outcomes and find new therapeutic strategies. Existing avenues for research include mouse models, which have hinted at complex mechanisms of shifting TEP identities and differentiation biases during development^{22–27}, but they are limited by their stark differences in gestation time and thymus morphology compared to humans, such as the near-absence of Hassall's corpuscles²⁸. Human primary TECs have also served as important platforms, allowing the identification of various TEC subsets through recent advances in scRNA-seq^{29–33}, but primary tissues are difficult to obtain from human embryos, contain fewer than 1% TECs, exhibit high sample-to-sample variability³⁰, and cannot be continuously observed during development. When placed into culture, primary TECs rapidly downregulate their *FOXNI* expression^{34,35}, preventing the study of their developmental dynamics. To advance our understanding of human thymus development, a model capable of recapitulating this process in vitro would be invaluable.

Human pluripotent stem cells (PSC) have served as a powerful tool to derive cells of all three germ layers and observe their development in vitro³⁶. In the past decade, several reports have shown successful induction of *FOXNI*⁺ TEP-like cells able to contribute to the reconstitution of thymic function in immunodeficient mice upon in vivo transplantation^{37–49}. However, in vitro, these cells remain characteristic of immature or intertypical TECs and cannot yet be differentiated into the diverse mature cTEC and mTEC lineages present in the human fetus, limiting their use as a model of thymus organogenesis. Faithful in vitro recapitulation of the full process of TEC development from PSCs would allow the analysis of fate decisions and trajectories that are difficult to observe in vivo, enable the interrogation of lineage-specific pathological mechanisms, and provide a potential source of specific types of TECs for regenerative applications.

Therefore, in this study, we establish a fully in vitro induction system capable of deriving diverse mature TEC-like lineages from human induced pluripotent stem cells (iPSC). By combining retinoic acid (RA)-based endodermal patterning with a period of self-directed differentiation, we obtain highly heterogeneous populations resembling cTECs, mTEC^{low}, mTEC^{high}, and mimetic mTECs. Using *FOXNI*^{mCherry} reporter lines, we observe the association of *FOXNI* with different TEC phenotypes over time and isolate *FOXNI*⁺ cells for functional evaluation through thymocyte co-culture, through which we further obtain *AIRE*⁺ and post-*AIRE* populations. With single cell profiling, we benchmark our induced cells against primary human TECs and determine their lineage trajectories and potential regulators. Together, our system provides a new window into the complex process of thymus organogenesis in an easily reproducible and accessible human model.

Results

Induction of TEP-like cells through RA-based endodermal patterning

To recapitulate the development of TECs in vitro, we established a chemically defined protocol that directs iPSCs through each intermediary stage in 2D culture by manipulating important signaling pathways governing lineage fate decisions. We started from a previously published protocol capable of efficiently producing anterior foregut endoderm (AFE)⁵⁰, an intermediary stage from which pharyngeal organs, as well as lungs, are known to develop⁵¹. Although WNT and BMP signaling have been implicated in thymus development^{52–54} and are abundantly used in TEP induction protocols^{37–49}, they are also known to posteriorize AFE into the lung lineage^{55–57}. To specifically induce pharyngeal endoderm (PE), we omitted both WNT and BMP and instead used RA, which directly promotes *HOXA3*, a transcription

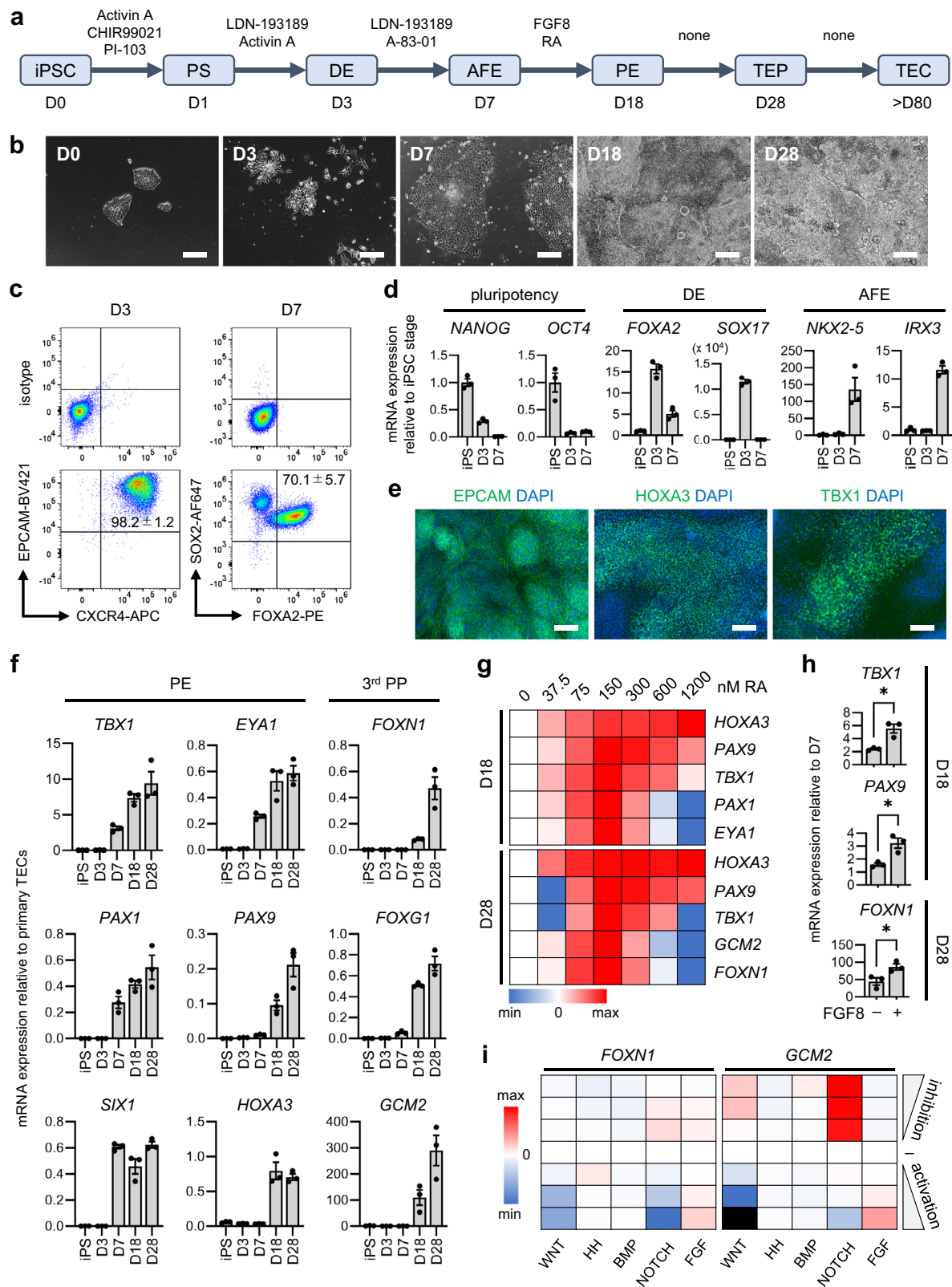
factor (TF) specifying the positional identity of the 3rd PP from which the thymus arises^{58–60} (Fig. 1a). We also included FGF8, which is required for the development of the 3rd PP derivatives^{61,62}.

By day 3 of the induction using the standard 201B7 iPSC line⁶³, cells adopted a characteristic endodermal morphology and expressed markers of definitive endoderm (DE), with nearly all cells being EPCAM⁺CXCR4⁺ (Fig. 1b–d). Markers of pluripotency were downregulated by day 7, when AFE markers were highly expressed and 70% of cells were SOX2⁺FOXA2⁺ (Fig. 1c, d). By day 18, cells were overconfluent and started forming dense patches, with PE markers TBX1 and HOXA3 being widely expressed (Fig. 1e). High expression of TFs crucial for thymus organogenesis^{64–68} was seen by day 18, leading to the expression of *FOXNI* at around half of the level of purified human primary TECs from pediatric donors by day 28 (Fig. 1f). Strikingly, it is only a narrow range of RA concentrations from day 7 to 18 that eventually led to the expression of PE and 3rd PP markers, indicating that the precise level of *HOXA3*, which was dose-dependent on RA, is crucial in correctly specifying the 3rd PP and allowing the development of its derivatives (Fig. 1g). This is supported by the concurrent expression of *GCM2*, a marker of the parathyroid, which also arises from the 3rd PP, but on the dorsal side⁶⁹. RA alone was necessary and sufficient for this process, although the addition of FGF8 further increased PE markers and *FOXNI* (Fig. 1h).

Overall, similar results in marker expression and morphology were seen in two more commonly used iPSC lines, 409B2⁷⁰ and 1383D6^{71,72} (Supplementary Fig. 1a, b). Markers of other endodermal lineages, including thyroid and pancreas, were also detected at low levels on day 28 (Supplementary Fig. 1c). When RA or FGF8 supplementation was extended until day 28, no further increase in *FOXNI* expression was seen compared to no supplementation after day 18, indicating that their effect is temporally restricted to the PE stage (Supplementary Fig. 1d). Similarly, when adding agonists or antagonists of major developmental signaling pathways from day 18 to 28, we found that neither BMP nor WNT stimulation increased *FOXNI* expression (Fig. 1i). On the contrary, high WNT activation abolished *FOXNI* and *GCM2* expression, reflecting the disruption of thymus and parathyroid development in mice with elevated WNT signaling in TECs⁷³. NOTCH inhibition increased *GCM2* but not *FOXNI*, suggesting a possible role of this pathway in promoting the parathyroid fate over the thymic fate. Together, these results indicate that endodermal patterning based on RA modulating *HOXA3* is sufficient to drive *FOXNI* expression and induce TEP-like cells after the AFE stage, while the modulation of other signaling is largely dispensable. Indeed, signaling molecules of pathways involved in both early and later stages of TEC development, such as NOTCH^{74,75} and NF-κB^{76,77}, automatically increased over time without further signaling manipulation, suggesting possible self-directed differentiation after 3rd PP specification (Supplementary Fig. 1e).

Self-directed differentiation of cTEC- and mTEC-like populations

Since *FOXNI*⁺ TEP-like cells autonomously emerged after the PE stage, we next opted to maintain the induced cells in long-term culture to allow potential differentiation into all TEC lineages. To identify TEC-like cells and monitor their developmental changes over time, we constructed *FOXNI*^{mCherry} reporters from all three iPSC lines (Supplementary Fig. 2a). All *FOXNI*^{mCherry} reporter lines had heterozygous insertion of the T2A-mCherry construct immediately after the last codon of the *FOXNI* gene, showing no indels in the non-targeted allele and maintaining the normal karyotype of their parent lines (Supplementary Fig. 2b, d, e). Cell morphology was normal at the iPSC stage, and the expression of PE markers and *FOXNI* was similar to the parent lines on day 28 (Supplementary Fig. 2c, f). Consistent with the mRNA expression results (Fig. 1f and Supplementary Fig. 2f), mCherry fluorescence was visible by day 28, followed by an increase in brightness over time (Fig. 2a). Notably, while only a single mCherry⁺ population



was present on day 28, distinct mCherry^{high} and mCherry^{low} populations were evident by day 57, both of which were still maintained on day 133, indicating self-directed differentiation into stable, heterogeneous TEC-like lineages (Fig. 2b–d). By day 133, 3D-like nodules up to 100 μ m in height were observed within mCherry⁺ colonies (Supplementary Video 1).

When isolating the mCherry^{high}, mCherry^{low}, and mCherry⁺ populations, we found that cTEC-specific genes such as *PSMB11* and *PRSS16*, involved in self-peptide generation^{16,78,79}, were exclusively expressed by mCherry^{high} cells at levels matching or exceeding human primary TECs by day 39 (Fig. 2e). In contrast, expression of mTEC-specific genes such as *CCL19* and *CCL21*, crucial for thymocyte chemoattraction

Fig. 1 | TEC induction from human iPSCs in a 2D in vitro system. **a** Schematic of the TEC induction protocol. D, day; iPSC, induced pluripotent stem cell; PS, primitive streak; DE, definitive endoderm; AFE, anterior foregut endoderm; PE, pharyngeal endoderm; TEP, thymic epithelial progenitor; TEC, thymic epithelial cell. **b** Phase-contrast images of induced cells at each time point. Scale bars, 200 μ m. **c** Representative flow cytometry results (bottom) with the respective isotype controls (top) on D3 and D7. **d** mRNA expression of pluripotency, DE, and AFE markers over time. **e** Representative images of immunostaining on D18. Scale bars, 100 μ m. **f** mRNA expression of PE and third pharyngeal pouch (3rd PP) markers over time. **g** Heatmap of the mean log₂ fold change in mRNA expression in response to retinoic acid (RA) addition from D7 to D18 compared to DMSO only. **h** mRNA expression with or without addition of 50 ng/ml FGF8 from D7 to D18. **i** Heatmap of the mean log₂ fold change in mRNA expression in response to addition of inhibitors or activators of specific signaling pathways from D18 to D28 compared to untreated

(-) controls on D28. Black fields represent cases where expression was undetected in all experiments. The signaling modulators and concentrations are as follows: WNT inhibition, 0.2 μ M, 1 μ M, 5 μ M XAV939; WNT activation, 0.4 μ M, 2 μ M, 10 μ M CHIR99021; HH inhibition, 0.2 μ M, 1 μ M, 5 μ M Cyclopamine; HH activation, 20 nM, 100 nM, 500 nM SAG; BMP inhibition, 40 nM, 200 nM, 1 μ M LDN-193189; BMP activation, 4 ng/ml, 20 ng/ml, 100 ng/ml BMP4; NOTCH inhibition, 0.1 μ M, 0.5 μ M, 2.5 μ M RO4929097; NOTCH activation, 0.2 mM, 1 mM, 5 mM Valproic acid; FGF inhibition, 40 nM, 200 nM, 1 μ M BGJ398; FGF activation, 2 ng/ml, 10 ng/ml, 50 ng/ml FGF8. All results are from $n = 3$ independent experiments using 201B7 and values indicate the mean \pm SEM (standard error of the mean). Statistical significance was determined using unpaired two-sided t -tests (n.s. no significant difference, $*p < 0.05$). Exact p -values in (**h**) are as follows: *TBX1*, 0.0106; *PAX9*, 0.0146; *FOXP1*, 0.0374.

to the medulla^{80,81}, was found preferentially in mCherry^{low} cells or dispersed throughout all populations at levels similar to or lower than primary TECs by day 133. Only *AIRE* was not expressed, probably due to the lack of thymocyte-derived signals^{82,83}. These results suggest a preponderance of cTEC-like phenotypes within mCherry^{high} and early mTEC-like phenotypes within mCherry^{low}, consistent with previous reports showing higher *FOXP1* expression in cTECs compared to mTECs^{35,84}. Similar results were observed at the protein level on day 80, with cTEC marker PSMB11 being seen exclusively within bright mCherry⁺ areas, likely corresponding to the mCherry^{high} population in flow cytometry, and mTEC marker KRT5 being seen preferentially in areas with little to no mCherry fluorescence (Fig. 2f). KRT8 and CLDN4, used in murine studies to identify cTECs⁸⁵ and mTECs⁸⁶ respectively, were seen throughout the culture similar to epithelial marker EPCAM, suggesting that they may already be expressed in mCherry⁺ progenitors or possibly in cells of other epithelial lineages (Supplementary Fig. 3a). By day 133, PSMB11⁺ cTEC-like areas and KRT5⁺ mTEC-like areas were largely spatially separated with some overlapping junctional areas where 3D-like nodules formed, implying the presence of distinct niches similar to the thymus^{33,76,87} (Fig. 2g and Supplementary Video 2).

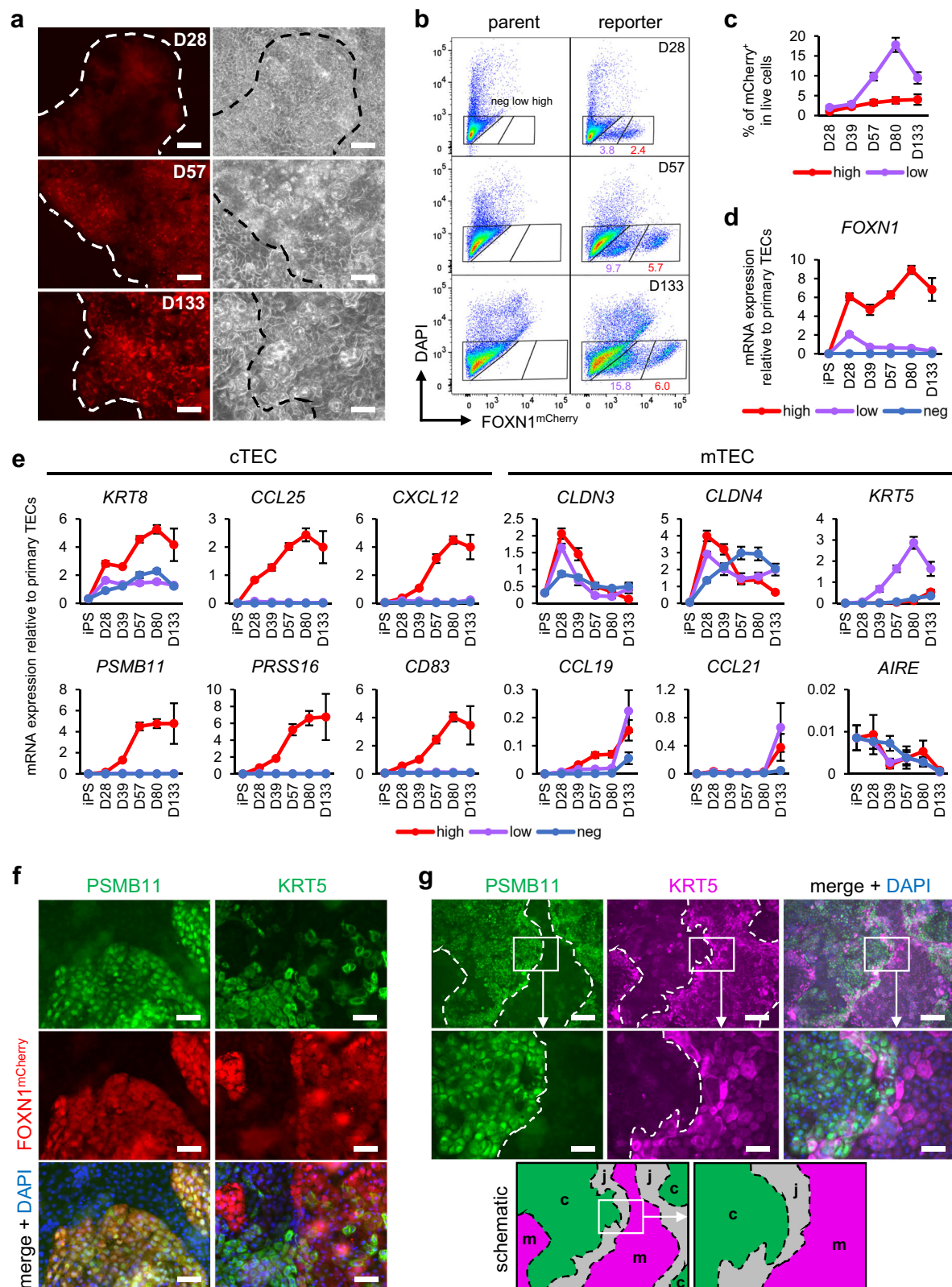
Over time, expression of several PE markers remained high in the mCherry^{high} population but decreased to levels similar to primary TECs in the mCherry^{low} population (Supplementary Fig. 3b). Of note, after a period of co-expression with *FOXP1* on day 28 to 39, *GCM2* and *TBX1* starkly decreased in mCherry^{high} and mCherry^{low} cells (Supplementary Fig. 3b). Since *GCM2* and *TBX1* are essential for parathyroid organogenesis⁸⁸, this suggests the presence of a transcriptional program suppressing the parathyroid fate once *FOXP1* is expressed. Both NOTCH and NF- κ B signaling genes remained highly expressed after day 28 in all populations, with NF- κ B being especially prominent in the mCherry^{high} population, and postnatal TEC progenitor marker^{29,30} *DLK2* being expressed by mCherry^{high} and mCherry^{low} cells in the later stages (Supplementary Fig. 3c). Ki-67⁺ cells were detected even on day 133, suggesting the development and long-term maintenance of a progenitor-like population in our system (Supplementary Fig. 3d).

Induced TECs are mature and capable of positive selection in vitro

Importantly, after day 28, our induced TECs (iTTC) started expressing functional markers, such as *IL7* and *DLL4*, crucial for the development of T cells in the thymus^{89,90} (Fig. 3a). MHC class II molecule HLA-DR, required for the selection of CD4⁺ T cells, was seen especially in clusters in the central parts of mCherry⁺ colonies, indicating the emergence of functionally mature iTTCs from these areas (Fig. 3b). HLA-DR and HLA-DQ were seen predominantly in the mCherry^{high} population, where the majority of the cells also expressed CD205, a marker of cTECs⁹¹, although some expression also appeared in mCherry^{low} and mCherry⁺ cells by day 133 (Fig. 3c–e). These results suggest that most of the mature iTTCs in our system are characteristic of cTECs, while

mTEC-like cells take longer to develop and mature, consistent with previous research indicating cTECs maturing earlier than mTECs in the human fetus³². MHC class I molecules HLA-A, -B, and -C, required for the selection of CD8⁺ T cells, were maintained in all populations. CD90, a mesenchymal marker recently shown to also be expressed in TECs, possibly contributing to hybrid epithelial-mesenchymal characteristics³⁵, was also detected in most mCherry⁺ cells. Overall, similar results were seen in another 201B7 *FOXP1*^{mCherry} reporter clone, as well as *FOXP1*^{mCherry} reporters established from 409B2 and 1383D6 iPSC lines (Supplementary Fig. 4a–g and 5a–e).

We then proceeded to assess the functionality of our iTTCs. Up to now, various systems for in vitro T cell production have been developed, including DLL1-expressing stromal cell lines and feeder-free DLL4 microbead systems, but while they are highly adept at generating CD4⁺CD8⁺ double-positive (DP) T lineage cells, efficient induction of mature single-positive (SP) cells, especially CD4⁺ SP cells, remains challenging^{41,92–94}. Therefore, we tested the capability of our iTTCs to engage in positive selection, a crucial event in the development of SP cells orchestrated by cTECs in the thymus^{1,2,7}. We sorted day 37 mCherry⁺ iTTCs and co-cultured them with pre-selection (CD69⁺) DP thymocytes from pediatric human thymi in 3D organoids at the air-liquid interface, similarly to a previous report using the murine stromal MS5-hDLL1 cell line⁹², with the same cell line used as control (Fig. 4a, b and Supplementary Fig. 6a, b). After 2 weeks of organoid co-culture, MS5-hDLL1 still had 91% of the CD3⁺TCR $\alpha\beta$ ⁺ population remaining as CD4⁺CD8⁺ DP, consistent with the previously reported long co-culture periods required to produce mature SP cells⁹², while iTTCs led to CD4⁺ and CD8⁺ SP phenotypes in respectively 47% and 21% of CD3⁺TCR $\alpha\beta$ ⁺ (Fig. 4c, d). Within both CD4⁺ and CD8⁺ SP cells, the percentage of CD62L⁺CD45RA⁺CCR7⁺ was significantly higher for iTTCs compared to MS5-hDLL1, consistent with the acquisition of a mature naïve T cell phenotype. Furthermore, 55% of the total CD8⁺ SP cells resulting from iTTC co-culture were CD8 $\alpha\beta$ ⁺, while only 22% were CD8 $\alpha\beta$ ⁺ in the case of MS5-hDLL1, indicating that agonist selection resulting in innate-like CD8 $\alpha\alpha$ phenotypes⁹⁵ was dominant in MS5-hDLL1, while high rates of positive selection were achieved by iTTCs. Unlike MS5-hDLL1, iTTCs were also capable of generating 8% TCR $\gamma\delta$ ⁺ cells, similarly to the human thymus where some TCR $\gamma\delta$ ⁺ cells develop along the DP pathway⁹⁶. To confirm the positive selection of CD4⁺ and CD8⁺ SP cells with diverse TCR rearrangements by iTTCs, we performed TCR repertoire analysis after the 2 weeks of co-culture. TCR β repertoires of both CD4⁺ and CD8⁺ SP cells generated through iTTCs showed high diversity similar to that of pediatric donor-derived CD4⁺ and CD8⁺ SP thymocytes, which had comparable gene usage frequencies and Shannon index⁹⁷ values (Fig. 4e and Supplementary Fig. 6c, d). When co-culturing mCherry⁺ iTTCs with earlier-stage CD4⁺CD8⁺ double-negative (DN) thymocytes, we found a greater propensity toward innate-like TCR $\gamma\delta$ and CD8 $\alpha\alpha$ phenotypes, likely in part due to the inherent heterogeneity of DN cells⁹⁸, but conventional CD3⁺TCR $\alpha\beta$ ⁺



thymocytes including 45% DP, 8% CD4⁺ SP, and 42% CD8⁺ SP cells were also generated (Supplementary Fig. 7a–d). In contrast, mCherry⁺ iTECs could neither induce SP thymocytes nor maintain DP or DN thymocytes in co-culture (Supplementary Fig. 6e, 7e). Together, these results demonstrate the capability of our iTECs, specifically the mCherry⁺ population, to support both early T cell development and later positive selection of SP cells with diverse TCR repertoires.

iTECs consist of heterogeneous populations resembling primary TECs

Next, to characterize the different induced lineages and trace their development, we performed scRNA-seq at multiple time points during the induction. The resulting dataset contained a major *EPCAM*-expressing epithelial partition, as well as minor non-epithelial clusters including neurons, muscle, and mesenchyme, with similar cell

Fig. 2 | Differentiation of cTEC- and mTEC-like cells over time. a $FOXN1^{mCherry}$ fluorescence (left) and phase-contrast (right) images over time. Scale bars, 50 μm . **b** Representative flow cytometry results of the reporter compared to the parent line at each time point. **c** Quantification of the percentage of $mCherry^{\text{high}}$ and $mCherry^{\text{low}}$ cells using flow cytometry. All values indicate the mean \pm SEM from $n = 6$ (D0 to D80) or $n = 4$ (D133) independent experiments. **d, e** mRNA expression of thymic epithelial cell (TEC) markers in the $mCherry^{\text{high}}$, $mCherry^{\text{low}}$, and $mCherry^{\text{neg}}$ populations over time. All values indicate the mean \pm SEM from $n = 6$ (D0 to

D80) or $n = 3$ (D133) independent experiments. cTEC, cortical TEC; mTEC, medullary TEC. **f** Representative images of immunostaining on D80 from $n = 6$ independent experiments. Scale bars, 50 μm . **g** Representative images of immunostaining on D133 from $n = 3$ independent experiments. Scale bars, 200 μm (top row) and 50 μm (middle row). The bottom row schematically depicts the $PSMB11^+$ area (green), the $KRT5^+$ area (magenta), and the overlapping junctional area (gray) as seen in the immunostaining. c, cTEC-like; m, mTEC-like; j, junction. All results were obtained using the 201B7 $FOXN1^{mCherry}$ reporter iPSC line.

distributions in two separate inductions (Fig. 5a–c and Supplementary Fig. 8a, b). Epithelial cells, starting from a common AFE progenitor expressing *OTX2* on day 7, moved through the PE stage and diverged into several clusters with varying levels of *FOXN1* expression (Fig. 5a–c). From these terminal clusters, we identified one cTEC-like cluster and five mTEC-like clusters by scoring each cell according to its expression of the top human primary cTEC and mTEC markers defined by Park et al.³⁰ (Fig. 5d). As expected, *FOXN1* expression was highest in the cTEC cluster, which also exclusively showed high expression of cTEC markers including *PSMB11*, *LY75*, *CCL25*, *DLL4*, etc. (Fig. 5e, f). In contrast, *FOXN1* expression was generally low in the five mTEC clusters, which showed heterogeneous expression of diverse mTEC markers including *CLDN4*, *CCL19*, *KRT5*, *CD24*, etc. Mature TEC markers *HLA-DRA* and *CD74* were observed in the cTEC, mTEC I, and mTEC II clusters. Interestingly, a small cluster scoring high for mTEC marker expression while simultaneously expressing *KRT1* and *KRT10*, characteristic of keratinized cells (KC) present in Hassall's corpuscles^{99–102}, was also detected (Fig. 5d, e). The top differentially expressed genes in each of these terminal clusters constitute both novel and previously defined^{30,80,86,103–107} markers of TECs, including putative markers from recent scRNA-seq studies^{29,30,108,109} such as *IGFBP6*, *KRT15*, *KRT17*, *GAS6*, *AQP3*, *GNG11*, *PLTP*, *ELF3*, and *ANXA1*, confirming the identity of our iTECs and their use in finding and validating new markers (Fig. 5g).

To ascertain that the findings of our iTECs are biologically relevant and reflective of actual TECs in the human body, we isolated thymic stroma from three pediatric donors and performed scRNA-seq for comparison. Apart from mesenchyme, endothelium, and other non-epithelial cells, five TEC clusters expressing *EPCAM* were identified (Fig. 6a, b, and Supplementary Fig. 8c). Similarly to iTECs, primary cTECs uniformly expressed high levels of *FOXN1*, *PSMB11*, and *CCL25*, while primary mTECs were highly diverse and variably expressed *CCL19*, *CLDN4*, and *CD24* (Fig. 6b, c). The mTEC cluster labeled mTEC-CCL19/21 was marked by high expression of cytokines *CCL19* and *CCL21* and keratins *KRT5* and *KRT15*, with lower expression of *HLA-DRA*, similar to previously defined mTEC^{low} populations^{8,9,29}. The mTEC-CLDN3/4 cluster showed high expression of different claudins, such as *CLDN3* and *CLDN4*, and *HLA-DRA*, characteristic of previously defined mTEC^{high} populations^{8,29,86}. This cluster also contained the *AIRE*⁺ subpopulation, as well as cells expressing KC marker *KRT10* or secretory cell marker *SLPI*, indicating the presence of mimetic mTECs¹⁵. However, most mimetic mTECs were found in the cluster labeled mTEC-mim, which included diverse populations expressing neuronal marker *NEUROD1*, myocyte marker *MYOG*, and many others. In addition to these diverse mTECs, TECs with an intertypical phenotype labeled mcTEC were also detected, and these cells expressed various keratins and other markers of progenitor-like cells^{30,31}.

To directly compare the terminal clusters of iTECs to primary TECs, we integrated the samples of the primary thymic stroma with the iTEC samples from day 133 (Fig. 6d). Cells from the induced cTEC cluster co-clustered with primary cTECs and bridged the gap between cTECs and mcTECs, indicating they consist of a mixture of cTEC-fated progenitor-like cells and mature cTECs indistinguishable from their primary counterparts. Induced mTEC I cells co-clustered predominantly with the primary mTEC-CCL19/21 cluster, indicating an

mTEC^{low}-like phenotype, with some cells also overlapping with mcTECs. The induced mTEC II cluster showed substantial overlap with primary mTEC-CLDN3/4, suggesting an mTEC^{high}-like phenotype, while also extending toward mTEC-mim. Both induced mTEC III and KC clusters co-clustered with mTEC-CLDN3/4, with the former showing near-complete overlap with the subpopulation expressing secretory cell markers, and the latter co-clustering closely with the keratinized subpopulation (Fig. 6e, f). Multiple markers of secretory cells were indeed highly expressed in the induced mTEC III population, including *SLPI*, *CXCL17*, and various mucins and uroplakins reminiscent of secretory mimetic mTECs, indicating the possibility of mimetic mTEC-like cells being induced in our system despite the lack of *AIRE* expression (Fig. 6g). Recent studies have shown that mimetic mTECs, being fewer but not absent in *Aire*^{-/-} mice, require specific lineage-defining TFs in their development^{10,11}. To examine whether such TF activity is recapitulated in iTECs, we performed gene regulatory network (GRN) analysis and identified regulons—groups of genes that consist of a TF and its direct-binding targets—that were highly active in each cluster. Strikingly, among the top regulons in induced mTEC clusters, particularly mTEC III, were *ELF3*, *GRHL1*, and *GRHL3*, which have been implicated in the development of mimetic mTECs and the regulation of *Foxn1* in murine studies^{18,108,109} (Fig. 6h). Consistent with this, expression of *ELF3* and *EHF*, both putative regulators of late mTEC development¹⁰⁹, was especially high in induced mTEC III, suggesting crucial aspects of *AIRE*-independent mimetic mTEC differentiation were recapitulated in our system (Supplementary Fig. 8d). Some induced cells classified as neurons, myocytes, and ciliated cells also co-clustered with the primary mTEC-mim and mTEC-CLDN3/4 clusters, implying they could be mimetic mTEC-like cells, although only the Cil cluster showed high expression of *ELF3* and *EHF* (Supplementary Fig. 8d–f). In contrast, the induced mTEC IV and mTEC V clusters did not co-cluster with primary TECs (Fig. 6d). This is despite their expression of genes indicative of mcTECs or early mTECs, including *BCAM*, *DLK2*, *IGFBP6*, *ITGA6*, *CXCL14*, *WNT4*, *LAMB3*, and various keratins^{29–31,110,111}, but no thyroid, pancreas, lung or other epithelial lineage markers, indicating they could be either a transient or aberrant population not seen in vivo (Supplementary Fig. 8g).

Interestingly, the induced Mes cluster partially co-clustered with primary thymic mesenchyme (Fig. 6d), raising the question of whether these cells could affect iTECs similarly to how mesenchyme shapes the thymic microenvironment by secreting growth factors and extracellular matrix (ECM) components¹¹². Therefore, we performed ligand-receptor interaction analysis based on the expression of potential ligands on the Mes cluster and potential receptors on AFE-derived terminal clusters on day 133 (Supplementary Fig. 9a). This revealed high expression of BMPs, FGFs, and ECM components such as collagens and laminins in the Mes cluster, with appropriate receptors being heterogeneously expressed in iTEC clusters, implying the recapitulation of a thymus-like microenvironment with epithelial-mesenchymal interactions that may contribute to the self-directed differentiation in our system. These mesenchyme-like cells, marked by VIM, are present by day 28 and closely associate with $mCherry^+$ colonies by day 133, with localized accumulations of these cells below the $mCherry^+$ layer being observed in 3D-like nodules (Supplementary

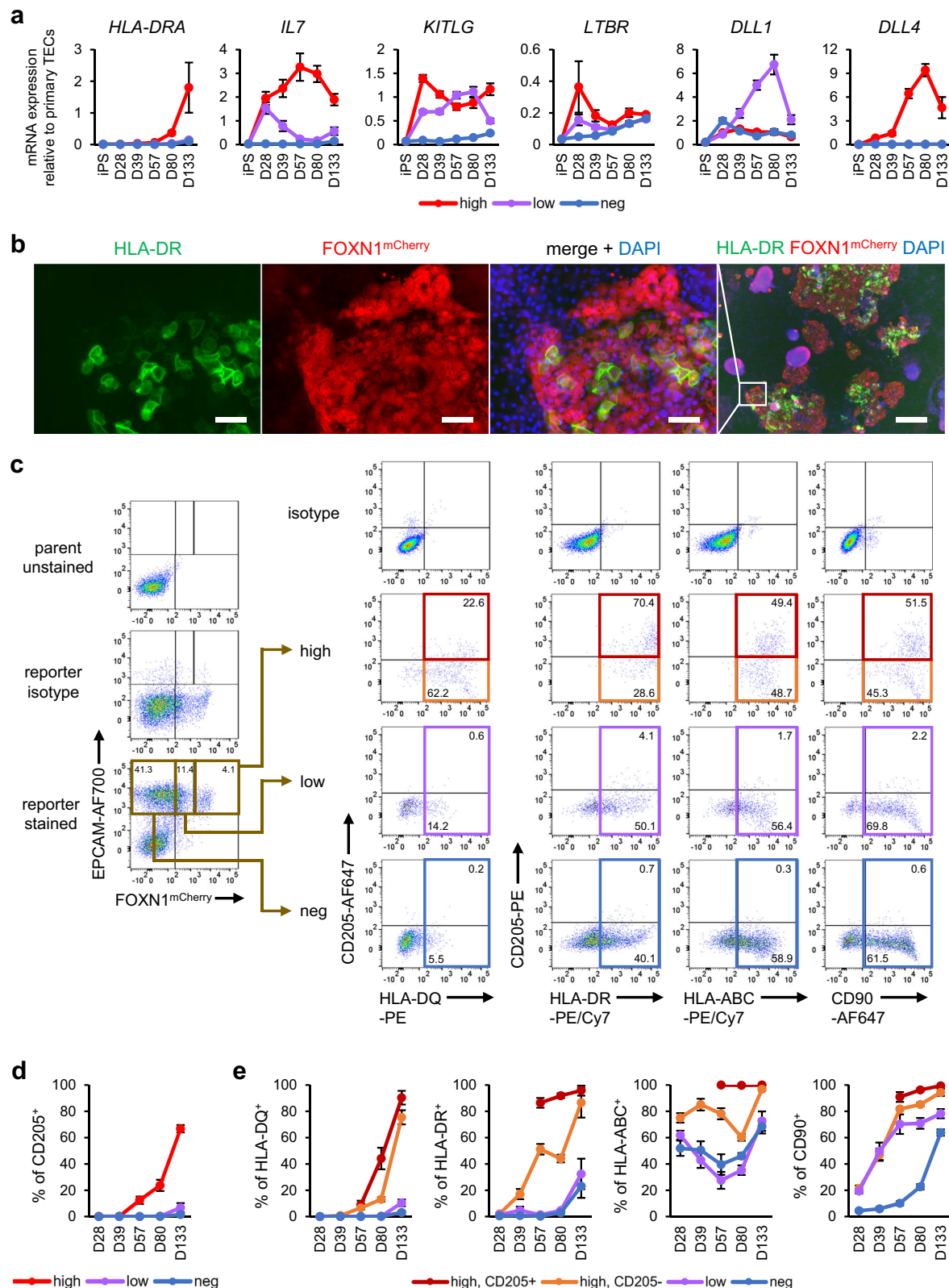


Fig. 3 | iTECs express functional and mature TEC markers. **a** mRNA expression in the mCherry^{high}, mCherry^{low}, and mCherry⁻ (neg) populations over time, D, day. **b** Representative images of immunostaining on D80 from $n = 6$ independent experiments. Scale bars, 50 μ m (left three images) and 500 μ m (right image). **c** –**e** Representative flow cytometry results on D133 (**c**) with quantification of CD205⁺ cells in each of the EPCAM⁺-gated mCherry^{high}, mCherry^{low}, and mCherry⁻

populations (**d**) and quantification of HLA-DQ⁺, HLA-DR⁺, HLA-ABC⁺, and CD90⁺ cells in the mCherry^{high}CD205⁺, mCherry^{high}CD205⁻, mCherry^{low}, and mCherry⁻ populations (**e**). In (**c**), colored gates correspond to the populations used for quantification in (**e**). All results were obtained using the 201B7 *FOXP1*^{mCherry} reporter iPSC line. All values indicate the mean \pm SEM from $n = 6$ (D0 to D80) or $n = 3$ (D133) independent experiments. TEC, thymic epithelial cell; iTEC, induced TEC.

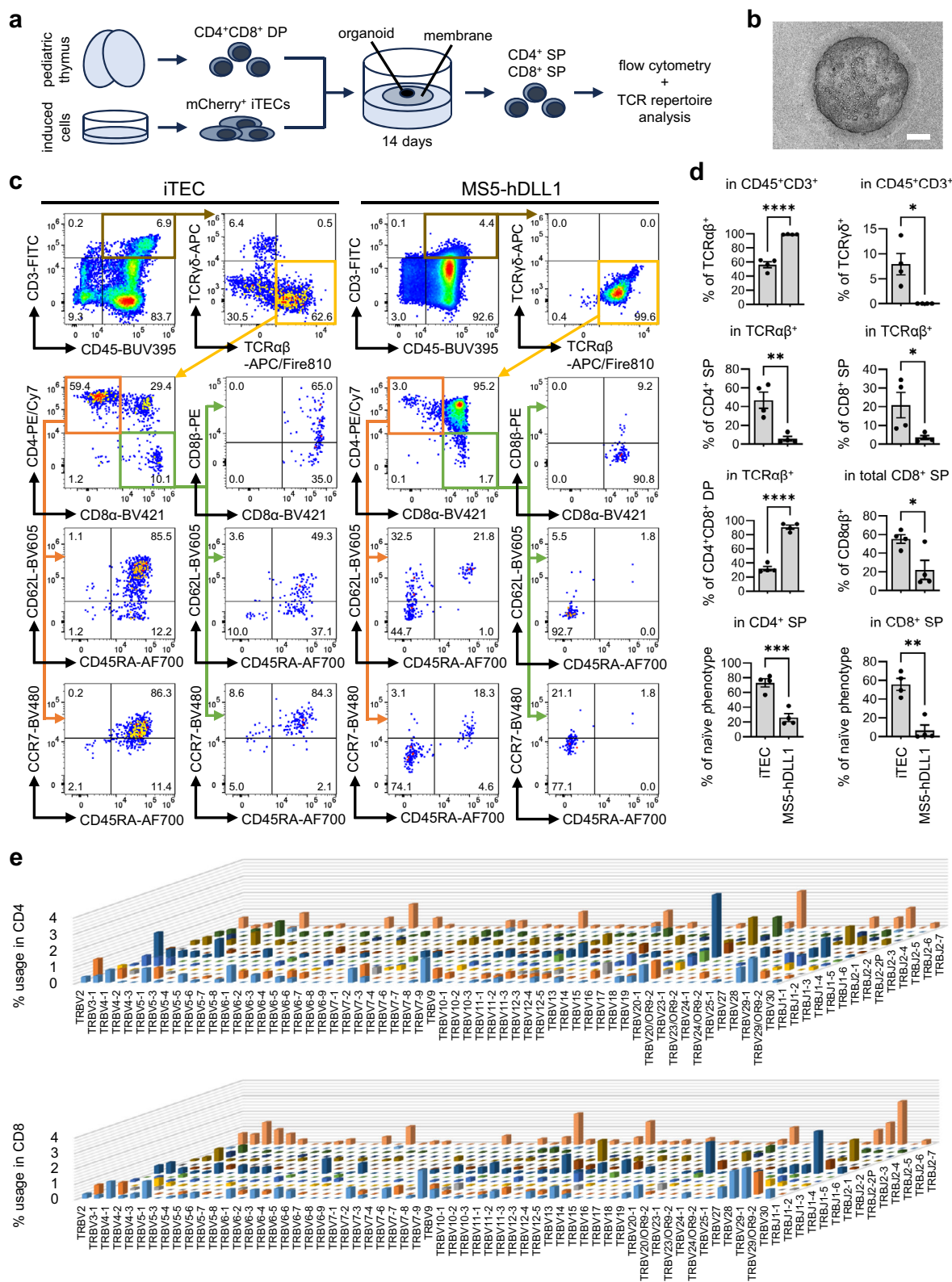


Fig. 9b, c and Supplementary Video 3). We are unable to determine the origin of the Mes cluster due to its lack of connections to other clusters, but iTECs themselves variably express low levels of mesenchymal markers, including *THY1*, *COL1A1*, *VIM*, *FN1*, and *CD44* (Supplementary Fig. 9d), similarly to primary TECs³⁵, possibly allowing epithelial-to-mesenchymal transition (EMT) to occur^{110,113}.

mTEC-like lineages diverge at varying stages of cTEC development

Aiming to shed light on the origins of TEC heterogeneity, we used pseudotime and RNA velocity analysis to determine how the different TEC-like lineages arise from a common progenitor in our system. We find the earliest lineage bifurcation at the PE stage, with one path

Fig. 4 | Efficient in vitro production of CD4⁺ and CD8⁺ T cells with diverse TCR repertoires through iTEC co-culture. **a** Schematic of induced thymic epithelial cell (iTEC) co-culture with double-positive (DP) thymocytes. SP, single-positive. **b** Phase-contrast image of an iTEC/thymocyte organoid during co-culture, representative of $n = 4$ independent experiments. Scale bar, 500 μm . **c** Representative flow cytometry results after 14 days of mCherry⁺ iTEC or MSS-hDLL1 organoid co-culture with DP thymocytes. **d** Quantification of co-culture results from (c), with values indicating the mean \pm SEM from $n = 4$ independent experiments. Naïve phenotype is defined as CD62L⁺CD45RA⁺CCR7⁺ triple-positive. Statistical

significance was determined using unpaired two-sided t -tests (n.s. no significant difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). Exact p -values are as follows: TCR $\alpha\beta$ ⁺, 0.0000557; TCR $\gamma\delta$ ⁺, 0.0102; CD4⁺ SP, 0.00378; CD8⁺ SP, 0.0447; CD4⁺CD8⁺ DP, 0.00000877; CD8 $\alpha\beta$ ⁺, 0.0257; naïve phenotype (in CD4⁺ SP), 0.000944; naïve phenotype (in CD8⁺ SP), 0.00123. **e** Gene usage of *TRBV* and *TRBJ* in CD4⁺ (top) and CD8⁺ (bottom) SP thymocytes obtained after 14 days co-culture of mCherry⁺ iTECs and DP thymocytes in $n = 1$ independent experiment. All iTEC results were obtained using the 201B7 *FOXP1*^{mCherry} reporter iPSC line.

leading to cTEC, mTEC I, and mTEC II clusters, and the other to mTEC III, mTEC IV, mTEC V, and KC clusters, with some possible plasticity between mTEC II and mTEC III (Fig. 7a, b). The TEP cluster, through which the cTEC, mTEC I, and mTEC II trajectories pass, appears by day 28, greatly shifts over time, and disappears by day 133 (Supplementary Fig. 10a). It can be divided into several subclusters, with clusters representing days 28 and 38 already expressing *FOXP1* and some cTEC markers, and clusters dominant at later time points co-expressing cTEC and mTEC markers to varying degrees (Supplementary Fig. 10b, c). Similar shifts in progenitor identity have also been recently reported in mice²⁷. Although it did not survive beyond day 49, a small parathyroid cluster was also detected on day 28, indicating that this time point likely corresponds to the point of divergence of the thymic and parathyroid primordia around week 6 of gestation in humans^{4,114}.

For the cTEC-fated cells, *FOXP1* expression increased and stayed high throughout the induction, whereas mTEC I- and mTEC II-fated cells downregulated their *FOXP1* expression at different points (Fig. 7c). cTEC cytokines *CCL25* and *CXCL12* were expressed by the mTEC I and mTEC II lineages during their progression through the TEP stage and later lost, reflecting murine studies showing at least some mTEC lineages being derived from progenitors with characteristics of early cTECs^{22–24}. mTEC I-fated cells also transiently expressed the cTEC-specific proteasome subunit *PSMB11*, while high expression of the cytokine *CCL21*, required for the mTEC function of attracting positively selected thymocytes to the medulla⁸¹, only arose once *PSMB11* expression was lost. Conversely, mTEC cytokine *CCL19* was also expressed to some degree in cTEC-fated cells before being downregulated. Surprisingly, many genes commonly used to identify mTECs, including *CLDN4*^{26,86,115,116} and *CD24*¹¹⁷, were already expressed by the AFE cluster on day 7, suggesting that they may already be present in progenitors before mTEC lineage commitment. In several mTEC-like lineages, expression of secretory or keratinized markers was preceded by the upregulation of *EHF*, *ELF3*, *ELF5*, and *GRHL3*, substantiating the possible role of these TFs in mTEC development^{108,109} (Supplementary Fig. 10d).

Genes that were most differentially expressed along pseudotime included many previously identified genes essential for TEC development, including *SIX1*⁶⁵ and *PAX1*⁶⁴ (Fig. 7d). Cells of the cTEC-like lineage also maintained high expression of *NFKB1*, as well as *GAS6*, which is mediated by NF- κ B signaling^{118,119}, both of which were only transiently seen in mTEC I. *PIK3R1*, *GOLM1*, and *CXXC5*, reported to be highly expressed in cTECs^{12,30}, were also exclusively expressed along the path leading to cTEC, mTEC I, and mTEC II. In contrast, on the path leading to mTEC III and mTEC IV, high transient expression of *SOX2*, *FOXEL*, and *FOXA1* was observed (Fig. 7d, e). *FOXA1* has been implicated in the development of secretory cells as well as mimetic mTECs^{11,120}, and a *Foxn1*^{Sox2}^{Foxe1} population has been reported in murine thymi at E11.5¹². Since expression of these genes decreased by day 133 and no cluster clearly co-expressing them was found in our pediatric thymic stroma, this may constitute an early, transient population fated toward a secretory phenotype (Fig. 7d, e, and Supplementary Fig. 10d, e). The transient expression of *FOXA1* and *FOXEL*, as well as the subsequent upregulation of secretory markers *PSCA* and *CXCL17*, was found almost

exclusively in the mCherry⁺ population, suggesting that this lineage may have the potential to develop independently of *FOXP1*, although this will require further confirmation by lineage tracing (Supplementary Fig. 10f). To examine the signaling that could potentially govern this lineage fate decision, we looked at regulon activity and analyzed differentially expressed genes along pseudotime in the mTEC III compared to the mTEC II lineage. *SOX2*, *FOXEL*, and *FOXA1* regulons were indeed highly active along the path to mTEC III, but not mTEC II, where regulons including *SIX1*, *GATA3*, and *FOXP1* were active (Supplementary Fig. 10g). *SHH*, reported to be upstream of *FOXP1*^{121,122}, and NOTCH inhibitor *DLK1* were transiently highly expressed in mTEC III-fated cells, whereas NOTCH target gene *HEY1*, as well as *BMP4* and *ID* genes downstream of BMP signaling^{123,124} were highly expressed in mTEC II-fated cells (Supplementary Fig. 10h). Furthermore, differential regulon activity between clusters was observed in RA receptors, indicating lineage-specific temporal changes in retinoid signaling (Supplementary Fig. 10i).

Thus, early differential signaling in PE may already determine the lineage fate of some types of mimetic mTECs, concurrently with or even before the establishment of the thymic and parathyroid primordia, while other heterogeneous mTEC populations diverge thereafter from *FOXP1*⁺ progenitors with characteristics of cTECs (Fig. 7f).

iTECs model *AIRE*⁺ mTEC differentiation in thymocyte co-culture

Since *AIRE* was not expressed at any time point during the induction, we next examined whether iTECs have the potential for *AIRE*⁺ mTEC differentiation by co-culturing them with thymocytes, known to promote mTEC development^{82,83}. After 2 weeks of organoid co-culture of day 37 mCherry⁺ iTECs with DP thymocytes, *AIRE* was significantly upregulated in the mCherry⁺ fraction, reaching 6% of the expression level of primary TECs (Fig. 8a, b). scRNA-seq of the organoids revealed diverse TEC-like populations similar to the non-co-cultured iTECs, with clusters of cTEC, mTEC I, mTEC II, and mTEC III that co-cluster closely with their primary counterparts (Fig. 8c, d, and Supplementary Fig. 11a, b). Mesenchyme-like clusters were also detected with partial co-clustering of primary mTECs and induced mTEC I from day 133, as well as their respective Mes clusters, implying that mCherry⁺ iTECs may partially undergo EMT in organoid co-culture, similarly to TECs acquiring EMT signatures during aging in vivo^{125,126} (Fig. 8c, d, and Supplementary Fig. 11a, c). In addition, small lymphocyte clusters including one containing *FOXP3*⁺*IL2RA*⁺ cells were present, indicating possible iTEC functionality of inducing Tregs from the DP stage (Fig. 8c, e and Supplementary Fig. 11c). *AIRE* expression was observed within a section of the mTEC II cluster directly connected to clusters co-expressing neuronal markers and *EPCAM*, reminiscent of neuronal mimetic mTECs (Fig. 8e, f). These clusters, assigned as mTEC-mim I-III, co-clustered closely with the mTEC-mim cluster in our dataset of human primary thymic stroma, as well as the TEC(neuro) cluster in the dataset of human thymic epithelium by Park et al.³⁰ (Fig. 8d and Supplementary Fig. 11b).

RNA velocity analysis, with the inferred latent time representing the cell's internal clock¹²⁷, indicates the likely origin of these neuronal

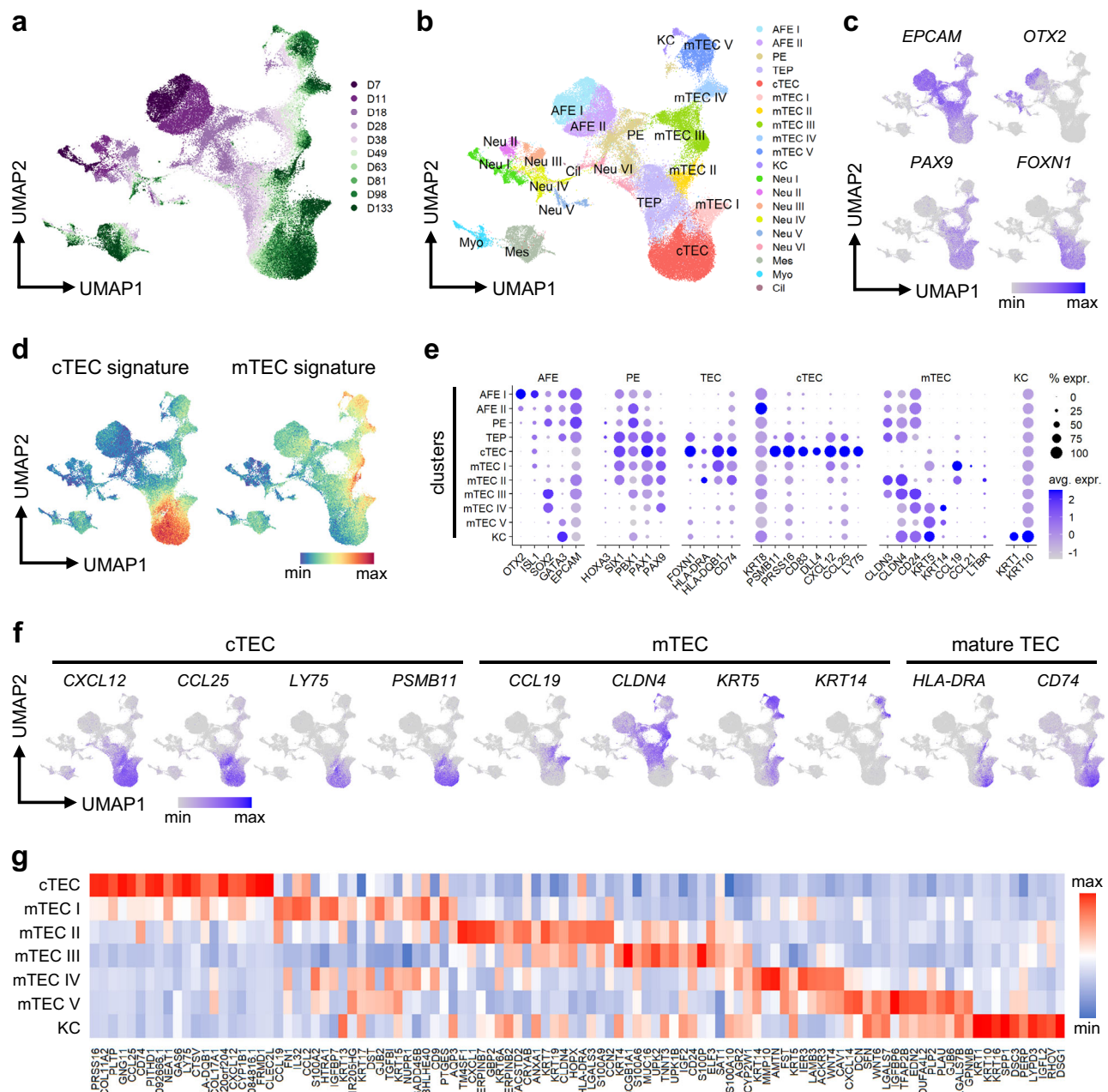


Fig. 5 | Single cell profiling of iTECs reveals heterogeneous TEC-like populations. a, b Induced cells from D7 to D133 according to time point (**a**) and unsupervised clustering (**b**) on UMAP. D, day; AFE, anterior foregut endoderm; PE, pharyngeal endoderm; TEP, thymic epithelial progenitor; TEC, thymic epithelial cell; cTEC, cortical TEC; mTEC, medullary TEC; KC, keratinized cell; Neu, neuron; Myo, myocyte; Mes, mesenchyme; Cil, ciliated. **c** Expression of AFE, PE, and TEC marker genes on UMAP. **d** Module scores of the top 20 cTEC and mTEC marker

genes as defined by Park et al.³⁰. **e** Dot plot of TEC and other marker gene expression in induced AFE-derived lineages. % expr., percentage of cells in which the gene is expressed; avg. expr., average expression. **f** Expression of cTEC, mTEC, and mature TEC marker genes on UMAP. **g** Heatmap of the average expression of the top 20 differentially expressed marker genes in each AFE-derived terminal cluster. All results were obtained using the 201B7 *FOXN1*^{mCherry} reporter iPSC line.

mimetic mTEC-like cells to be a subset within the mTEC II cluster expressing various keratins (Fig. 8g, h, and Supplementary Fig. 11d–f). The expression of keratins and *BCAM*, reflective of progenitors³¹, decreased across pseudotime, while recently identified^{30,128} markers of *AIRE*⁺ mTECs, such as *NTHL1*, *CALML3*, and *INSML1*, increased either transiently or permanently (Fig. 8i). *ELF3*, *EHF*, and *GRHL1* were expressed broadly within the mTEC II and the secretory mimetic mTEC III clusters independently of *AIRE* expression (Supplementary Fig. 11g). *AIRE*⁺ mTEC markers peaked nearly concurrently with those of early

neuronal differentiation, including *ASCL1* and *NEUROG1*, possibly driven by *INSML1*¹²⁹ (Fig. 8i). Thereafter, genes abundantly expressed in neurons, such as *STMN2* and *RTN1*, increased especially in mTEC-mim II, and markers more specific to differentiated neuronal subtypes, such as cerebellar Purkinje cells (*PCP4*), as well as glia (*METRN*) and cochlear hair cells (*ATOH1*), appeared in mTEC-mim III (Fig. 8i). Together, these results show that, upon co-culture with thymocytes, our iTECs are capable of recapitulating the dynamics of *AIRE*⁺ mTEC and post-*AIRE* neuronal mimetic mTEC development.

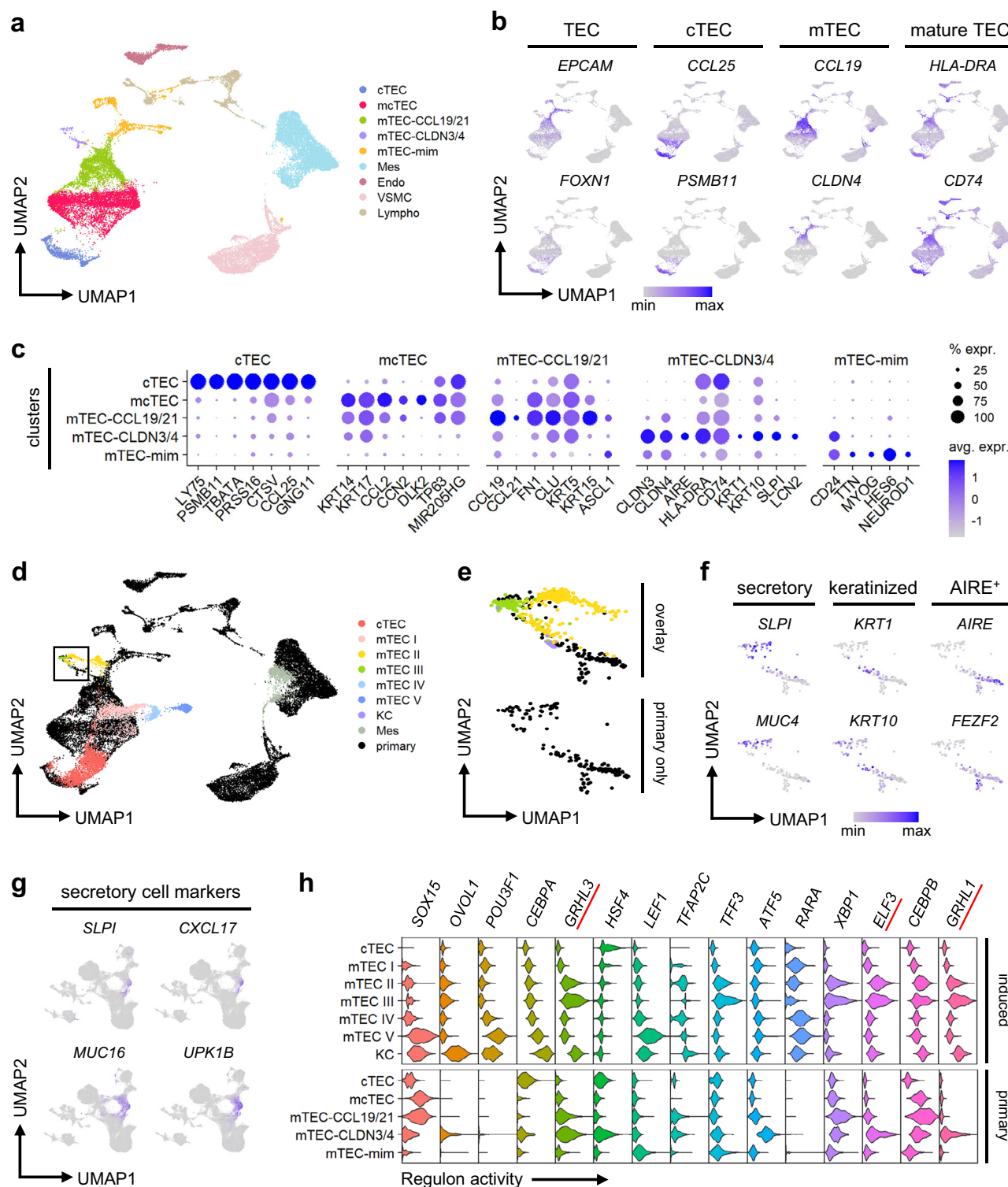


Fig. 6 | Induced cTEC- and mTEC-like cells resemble human primary TECs.

a UMAP of unsupervised clustering of primary thymic stroma. TEC, thymic epithelial cell; cTEC, cortical TEC; mTEC, medullary TEC; mTEC-mim, mimetic mTEC; mcTEC, intertypical TEC; Mes, mesenchyme; Endo, endothelial cell; VSMC, vascular smooth muscle cell; Lympho, lymphocyte. **b** Expression of TEC markers in primary thymic stroma on UMAP. **c** Dot plot of selected cluster and lineage marker gene expression in the TEC populations of primary thymic stroma. % expr., percentage of cells in which the gene is expressed; avg. expr., average expression. **d**, **e** UMAP of induced TEC, KC, and Mes clusters from day 133, with populations and colors

corresponding to Fig. 5b, integrated with primary thymic stroma (black). The primary mTEC-CLDN3/4 cluster and the co-clustering induced day 133 cells highlighted in (d) are shown separately in (e). **f** Expression of secretory, keratinized, and AIRE⁺ mTEC marker genes in the primary mTEC-CLDN3/4 cluster on UMAP. **g** Expression of secretory cell markers in induced cells on UMAP. **h** Violin plot of regulon activity in each cluster represented by scaled AUC (Area Under the Curve) scores from SCENIC, showing the top 15 induced mTEC marker regulons also found in primary mTECs. All results were obtained using primary thymic stroma from $n = 3$ pediatric donors and the 201B7 *FOXP1*^{mCherry} reporter iPSC line.

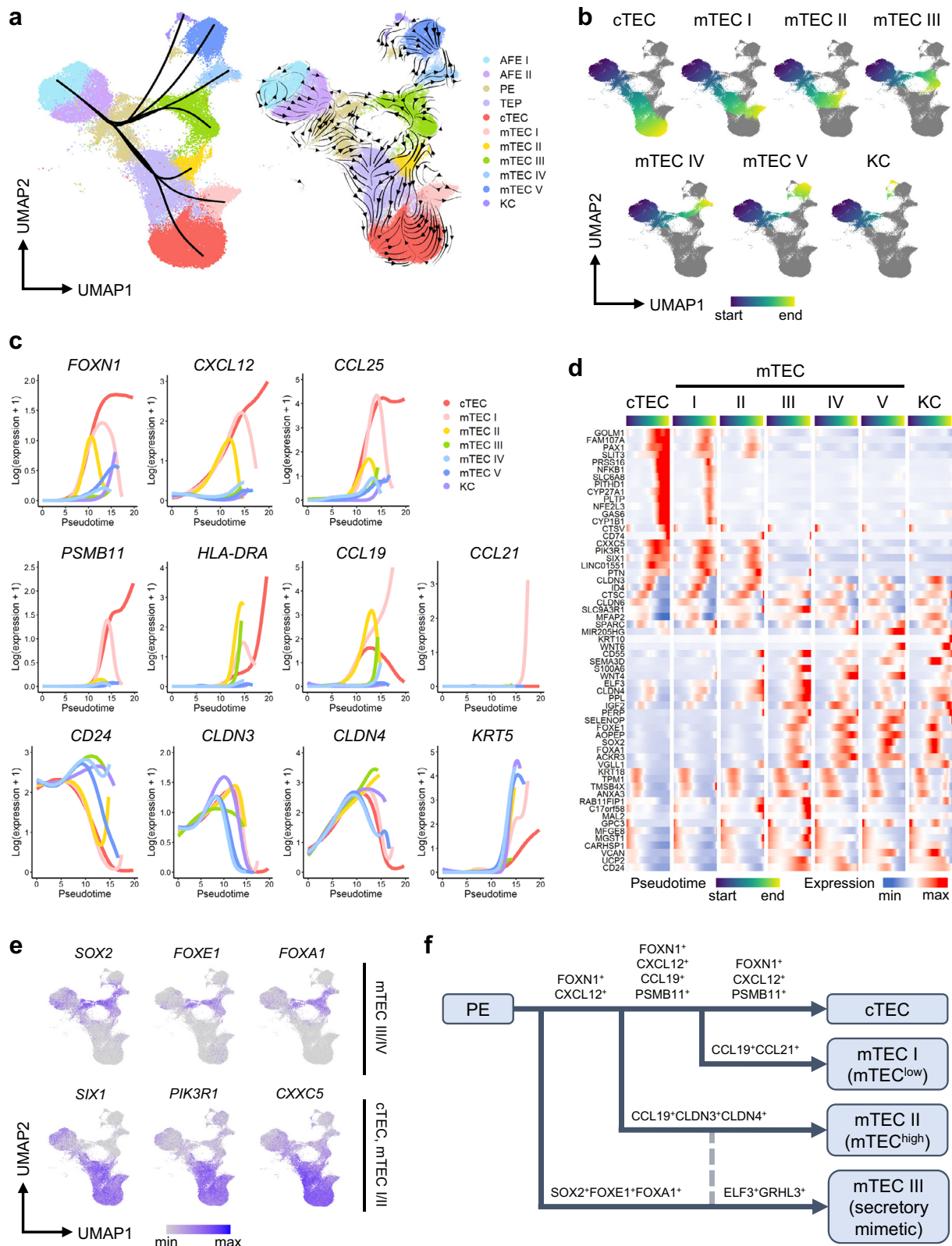
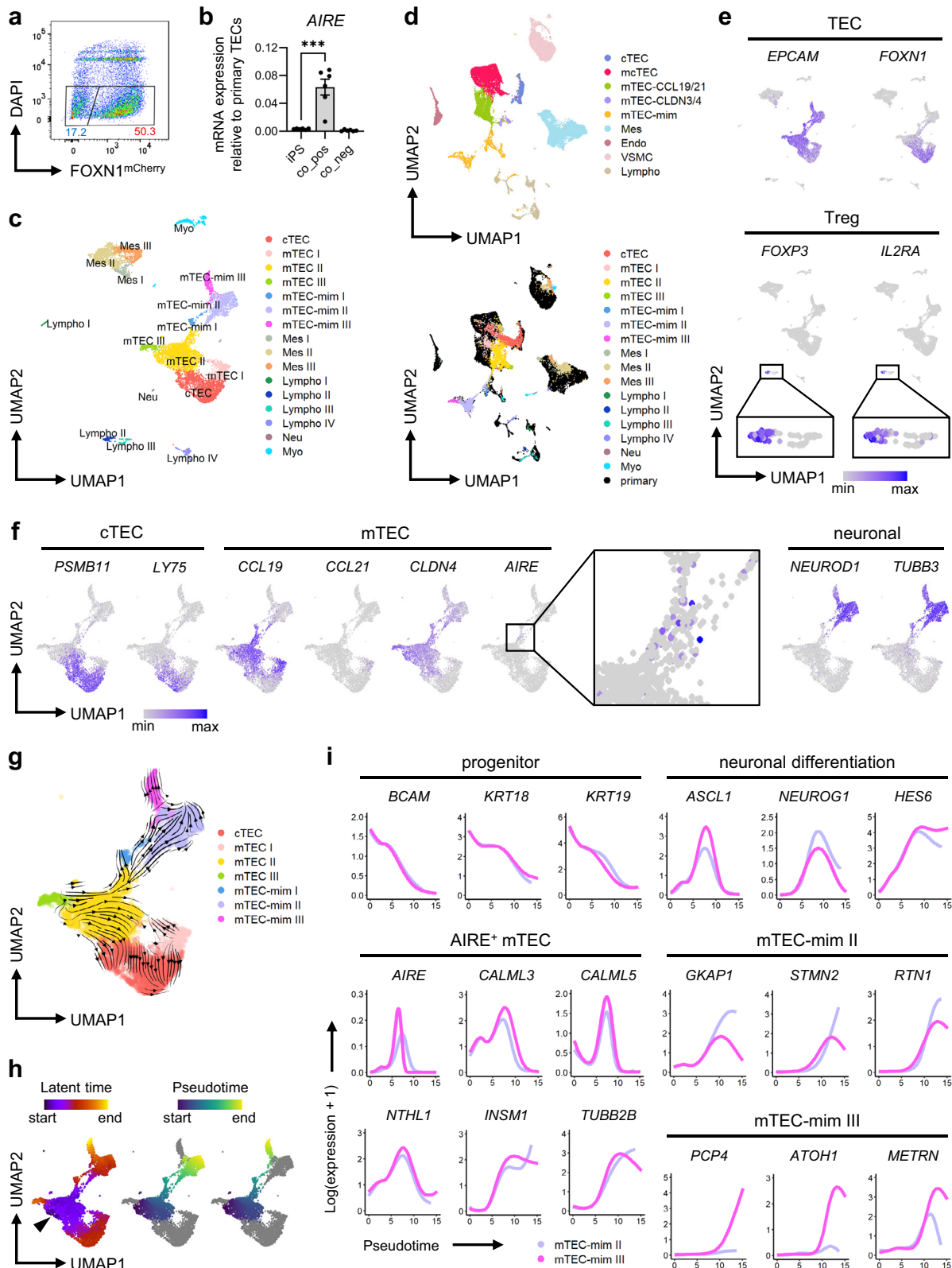


Fig. 7 | Single cell trajectory analysis of iTTECs models TEC development and lineage divergence. **a** Slingshot-identified lineage trajectories (left) and RNA velocity by scvelo (right) on UMAP. AFE, anterior foregut endoderm; PE, pharyngeal endoderm; TEP, thymic epithelial progenitor; TEC, thymic epithelial cell; cTEC, cortical TEC; mTEC, medullary TEC; KC, keratinized cell. **b** UMAP showing the trajectories of each lineage from Slingshot colored by pseudotime. **c** Smoothed gene expression of cTEC and mTEC markers along pseudotime in each lineage.

d Clustered heatmap of the scaled smoothed gene expression of the top 60 differentially expressed genes along pseudotime using the patternTest in tradeSeq. **e** Expression of selected genes from (d) showing the earliest lineage divergence from PE on UMAP. **f** Schematic of TEC lineage divergence inferred from pseudotime analysis of induced TECs (iTTEC), displaying the clusters with clear developmental paths and associated lineages in primary TECs. All results were obtained using the 201B7 *FOXP1*^{mCherry} reporter iPSC line.



Discussion

In this study, we successfully established a system to induce diverse, functional TEC-like cells from human iPSCs in vitro. While previous PSC-based protocols have required transplantation, complex 3D co-cultures, or decellularized scaffolds for TEC induction^{37–49}, our system can produce mature iTECs using simple, chemically defined 2D culture, allowing us to observe their development in a stable, easily

reproducible environment. In an important advance beyond these previous studies, where the induced cells were homogeneous with progenitor-like or intertypical phenotypes, our model recapitulates the divergence of heterogeneous cTEC and mTEC populations from a common progenitor. Through our approach combining RA-based endodermal patterning with a period of self-directed differentiation, we are able to produce highly diverse types of TEC-like cells, which

Fig. 8 | iTECs recapitulate *AIRE*⁺ and post-*AIRE* mTEC development in thymocyte co-culture. **a** Representative flow cytometry results of iTEC/thymocyte organoids after 2 weeks of co-culture from $n = 6$ independent experiments. TEC, thymic epithelial cell; iTEC, induced TEC. **b** mRNA expression of *AIRE* at the iPSC stage (iPS) and in the mCherry⁺ (co_pos) and mCherry⁻ (co_neg) populations after 2 weeks of iTEC/thymocyte co-culture. All values indicate the mean \pm SEM from $n = 6$ independent experiments. Statistical significance was determined using unpaired two-sided *t*-tests (n.s. no significant difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). The exact *p*-value is 0.000390. **c** UMAP of unsupervised clustering of cells after iTEC/thymocyte co-culture for 2 weeks from $n = 1$ independent experiment, consisting of 14 pooled organoids from one induction using four thymocyte donors with 3–4 organoids each. cTEC, cortical TEC; mTEC, medullary TEC; mTEC-mim, mimetic mTEC; Mes, mesenchyme; Lympho, lymphocyte; Neu, neuron; Myo, myocyte. **d** UMAP of the co-cultured cells from (c) integrated with the primary

thymic stroma from Fig. 6a, showing the clustering of the primary thymic stroma alone (top) and the overlay of the iTEC/thymocyte co-culture cells (bottom). **e, f** Expression of TEC and regulatory T cell (Treg) marker genes in all co-cultured cells (e) and cTEC, mTEC, and neuronal marker genes in the *EPCAM*⁺ clusters (f) on UMAP. Smaller areas containing *FOXP3*⁺*IL2RA*⁺ cells in (e) and the majority of the *AIRE*⁺ cells in (f) are also shown magnified. **g** RNA velocity of the *EPCAM*⁺ clusters of co-cultured cells by scvelo on UMAP. **h** UMAP showing the latent time inferred from RNA velocity (left) and the pseudotime of the mTEC-mim II and III lineages derived by Slingshot (middle and right). The black arrowhead indicates the area with the earliest latent time. **i** Smoothed gene expression of epithelial progenitor, *AIRE*⁺ mTEC, and early neuronal differentiation markers, as well as selected mTEC-mim II and mTEC-mim III cluster markers, along pseudotime in each lineage. All results were obtained using the 201B7 *FOXP3*^{mCherry} reporter iPSC line co-cultured with primary double-positive (DP) thymocytes.

share expression profiles and functional features with their primary counterparts in the pediatric thymus. Although it has long been assumed that signals necessary for TEC differentiation, maturation, and maintenance arise from their 3D microenvironment, as suggested by the rapid downregulation of *FOXP1* in isolated primary TECs placed into 2D culture^{34,35}, our iTECs stably express *FOXP1* and MHC class II molecules even after 133 days in 2D conditions, indicating that these crucial signals were successfully recapitulated and maintained in our culture. This long-term stability overcomes the limitation of previous induction protocols maintaining TEC-like cells for only up to 30 days with low expression of these markers in 2D when compared to primary TECs^{37–49}. Moreover, in an important advance for PSC-based modeling of the thymus, our iTECs are able to differentiate into *AIRE*⁺ mTEC-like cells and recapitulate *AIRE*-dependent neuronal mimetic mTEC development in thymocyte co-culture, indicating that they are able to respond to signals from thymocytes for further differentiation up to the terminal stages similarly to primary TECs. Together, these capabilities demonstrate the fidelity of our system and provide an opportunity to shed light on developmental processes that have otherwise been exceedingly difficult to observe.

Therefore, we used our system to explore important gaps in our understanding of human thymus organogenesis, especially how the remarkably diverse TEC lineages arise. We found that a precise dose of RA not only specifies PE, as previously reported¹³⁰, but also drives the differentiation of the thymic lineage in the absence of posteriorizing signals, as it led to swift upregulation of *FOXP1*, whose expression alone was enough to downregulate *GCM2* over time. This finding is consistent with murine and avian studies suggesting that both an under- and overabundance of RA at the PE stage may disrupt thymus organogenesis^{60,131}. Furthermore, since parathyroid-like cells were no longer detected after day 49, this indicates that the thymic lineage may constitute the default lineage in the 3rd PP in the absence of other signals, such as NOTCH inhibition. We were surprised to see a potential mTEC-like lineage divergence as soon as day 18 when *FOXP1* and *GCM2* only started to be expressed, marking the separation of the thymic and parathyroid primordia. In the mouse, separate domains of *Foxn1* and *Gcm2* expression are seen by E11.5¹³², when a transient *Foxn1* *Sox2*⁺ *Foxe1*⁺ population is also present¹². We were able to observe cells of this stage over several weeks, likely corresponding to the period up to 6–7 weeks of gestation when both *FOXP1* and *FOXE1* are expressed in the human thymic primordia^{4,133}. Our analyses suggest *FOXP1* *SOX2*⁺ *FOXE1*⁺ cells may be the earliest progenitors of an *EHF*⁺ *ELF3*⁺ *GRHL3*⁺ secretory mimetic mTEC-like population, with some similarities to the *Sox2*⁺ proximal lung lineage, which also develops under lower BMP signaling and eventually gives rise to secretory and ciliated cells^{134,135}. Not all early TFs are shared between these lineages, however, as *FOXE1* is not expressed in the developing lung^{133,136,137}, and *NKX2-1*, required for lung specification¹³⁸, is not present in our system.

After *SOX2*⁺ *FOXE1*⁺ cells branched off, *FOXP1*⁺ progenitors by day 28 started assuming characteristics of cTECs, from which the two mTEC-like populations expressing either *CLDN4* or *CCL21* subsequently appeared to diverge. In mice, embryonic progenitors show preferential differentiation to cTECs, while postnatal progenitors assume a bias toward mTECs²⁷. Consistent with this, single cell profiling of human thymi has shown that cTECs are more abundant in early fetal samples, while mTECs are the major TEC population in postnatal samples³⁰. Taken together with these previous findings, our results indicate that early cTEC-like cells arise before mTEC lineage commitment and may still have multi-lineage potential until the downregulation of *FOXP1* and cTEC markers. In our system, *FOXP1* downregulation coincided with increased expression of mTEC markers in the mTEC I and mTEC II clusters, suggesting the presence of a transcriptional program regulating *FOXP1* during mTEC lineage commitment. Indeed, human primary mTECs have much lower *FOXP1* expression compared to cTECs³⁵, and *GRHL1*, a potential repressor of *FOXP1*¹⁸, was inferred to be highly active in the mTEC II and mTEC III clusters in our GRN analysis. The mTEC I lineage was the last to diverge from the cTEC lineage in our system, even expressing cTEC-specific *PSMB11* before switching to high expression of mTEC-specific *CCL21* by day 133. This stage likely corresponds to week 11 or 12 in the human fetus, when *CCL21*⁺ cells start to arise, but *AIRE* is not yet expressed⁴. After co-culture with thymocytes, we were able to detect *AIRE*⁺ mTEC-like cells, possibly corresponding to the earliest *AIRE*-expressing TECs during week 13 of fetal development⁴.

Overall, these data suggest a default pathway to cTECs, from which mTECs sequentially diverge over time during embryonic development, consistent with the asymmetrical serial progression model proposed in mice^{3,115,139}. Previous reports have indicated that this serial progression may be exclusive to the embryonic period, given that in the postnatal murine thymus, the mTEC compartment is mostly maintained by unipotent mTEC-restricted stem cells, which arise from *Psmb11*-expressing progenitors only during early development^{25,26}. While lineage tracing studies have demonstrated the bipotent potential of embryonic cTEC-like cells^{22–24}, adult bipotent progenitors have been detected in various subsets including *Ly51*⁺ *MHCII*^{high} *Plet1*⁺ and *Ly51*^{low} *MHCII*^{low} *Sca-1*^{high} $\alpha 6$ -integrin^{high} in mice^{140,141}. Recently, it has also been shown that embryonic but not postnatal *Ccl21*⁺ cells are capable of converting into both cTECs and *Aire*⁺ mTECs¹⁴², and embryonic *Krt19*⁺ and *Sox9*⁺ subsets may act as mTEC-biased progenitors^{14,143}. Together, these murine studies indicate multiple possible pathways toward the cTEC and mTEC fates, with differing progenitor identities and levels of plasticity between the fetal and postnatal periods, although whether the same developmental dynamics are present in the human thymus, where certain markers such as *PLET1* are not expressed¹⁴⁴, remains unclear. As iPSC-based systems likely reflect the embryonic rather than the postnatal state, our iTECs offer a unique look into early human thymus

development that postnatal thymic epithelial stem cells or TEC lines cannot provide. Indeed, our single cell profiling allowed us to see developmental shifts in the TEP cluster, with RNA velocity further hinting at plasticity between closely related TEC lineages. While the heterogeneity of the terminally differentiated mimetic population is not yet fully represented in our system, with most mimetic mTEC-like cells by day 133 having a secretory or keratinized phenotype, and all likely post-*AIRE* cells in thymocyte co-culture assuming a neuronal or neuroendocrine phenotype, further improvements in modeling the lymphostromal crosstalk of the thymus in vitro may yet yield other subtypes of mimetic mTECs and allow further exploration of their origins in humans.

Apart from offering important biological insights, our iTECs also show promise for clinical application. For severely immunodeficient patients of DiGeorge syndrome¹⁹, Pignata Guarino syndrome²⁰, and other disorders of athymia, thymus transplantation is an established life-saving treatment, but it also carries risks of autoimmune complications and mortality through infections before sufficient T cell reconstitution occurs^{145,146}. As our system allows the differentiation of mature iTECs capable of generating naïve T cells with diverse TCR repertoires, we can envisage in vitro production of either T cells or thymic epithelium from patient-derived iPSCs or HLA-matched iPSC stocks as an alternative to allogeneic transplantation for immune reconstitution. This approach ensures patient MHC restriction, a major challenge for T cell therapies, and may further be applied to treatments of other genetic immunodeficiencies and cancers^{147–149}. The emergence of *AIRE*⁺ and *FOXP3*⁺ cells after thymocyte co-culture also raises the possibility that iTECs could support negative selection and Treg generation, which will be an important avenue to explore for applications to autoimmune diseases. Additionally, through its capability to recapitulate embryonic TEC development, our system will enable in vitro modeling of congenital athymias to better understand how these pathologies perturb early developmental processes and how they could be alleviated. Thymic dysfunction resulting from congenital diseases, chronic infections, chemotherapy, and age-related thymic involution may be mitigated with drugs promoting TEC proliferation and regeneration¹⁵⁰. Due to the long-term stability of TEC phenotypes and *FOXP3* expression over several months even in 2D culture, our iTECs will serve as an excellent candidate for screening drugs that could affect thymus development, pathology, and involution.

Taken together, our system, through its capability of generating diverse human cTEC- and mTEC-like cells, which up to now have been highly difficult to obtain from primary or other sources, presents an important platform to unveil the complexities of human thymus biology and is expected to contribute to broad applications in both basic and clinical science.

Methods

Collection of human thymus tissue

Thymi were collected from 23 donors, 12 female and 11 male, ranging between 1 month and 4 years of age (median = 4 months), undergoing pediatric cardiac surgery at the Kyoto University Hospital. Collected thymi were placed in cold, sterile PBS for up to 1 h until further processing for stromal cell and thymocyte isolation. This study complies with all relevant ethical guidelines and was approved by the Ethics Committee of Kyoto University Graduate School, Faculty of Medicine, and Kyoto University Hospital. Written informed consent was obtained from the parents of the donors. No compensation was provided. Sex, defined as the sex assigned at birth, was not considered in the study design, and disaggregated sex or gender information was not collected. Both male and female samples were used in each experiment including two or more donors, but due to their low number, the influence of sex on the results cannot be determined.

iPSC and other cell line culture

The iPSC lines 201B7⁶³ (human, female, RRID:CVCL_A324, kindly gifted by Dr. Masato Nakagawa, Kyoto University), 409B2⁷⁰ (human, female, RRID:CVCL_K092, kindly gifted by Dr. Keisuke Okita, Kyoto University), and 1383D6^{71,72} (human, male, RRID:CVCL_UP39, kindly gifted by Dr. Masato Nakagawa, Kyoto University) were previously established with written informed consent and used for this study with the approval by the Ethics Committee of the Center for iPS Cell Research and Application, Kyoto University. The iPSCs were cultured in feeder-free conditions in StemFit AK02N medium (Ajinomoto) at 37 °C and 5% CO₂. They were passaged once per week and seeded at a density of 1100 cells/cm² onto plates coated with iMatrix-511 silk (Nippi). At the time of passage, the medium was supplemented with 10 μM Y-27632 (Nacalai Tesque), which was removed on the following day. The previously established stromal cell line MS5-hDLL1⁹² (mouse, male, RRID:CVCL_VR88, purchased from Sigma-Aldrich, Cat#SCC167) was passaged twice a week at a ratio of 1:6 and maintained in D-MEM (Wako) supplemented with 10% fetal bovine serum (FBS, Sigma-Aldrich) and 1% Penicillin-Streptomycin Mixed Solution (Nacalai Tesque) at 37 °C and 5% CO₂. 201B7, 409B2, and 1383D6 were authenticated by STR analysis at the time of deposition to the RIKEN BioResource Research Center, and their *FOXP3*^{mCherry} reporters were authenticated by karyotyping by Chromocenter (Yonago, Japan) within five passages of their establishment. MS5-hDLL1 was verified by the vendor to be of mouse origin by a Contamination Clear panel by Charles River Animal Diagnostic Services.

Isolation of human primary thymic stroma

Similarly to a previously published protocol¹⁵¹, thymi were manually cut into 2–3 mm pieces and enzymatically digested for 60–90 min in Liberase solution containing 100 μg/ml Liberase TM (Roche), 200 μg/ml DNase I (Worthington), and 1% FBS in RPMI 1640 (Wako) at 37 °C on the Eppendorf ThermoMixer C with 800 rpm mixing frequency. Upon complete dissolution, the cells were centrifuged at 350 × g for 5 min, and the supernatant was removed. The cell pellet was resuspended in 83% v/v Percoll (Cytiva) in PBS, after which a lighter layer of 47% v/v Percoll in PBS and a further layer of only PBS were overlaid. The resuspended cells were centrifuged at 1490 × g for 30 min, 4 °C, and the stromal cells enriched between the PBS and lighter Percoll layer were collected. Cells were resuspended in TEC buffer containing 1% FBS and 5 mM EDTA (Dojindo) in PBS, counted, and incubated on ice with 50 μl MojoSort Human CD45 Nanobeads (BioLegend) per 1 ml TEC buffer for 1 × 10⁸ cells for 15 min. After one wash in TEC buffer, cells were resuspended in TEC buffer and depleted of CD45⁺ cells on a MagniSort Magnet (eBioscience) for 5 min. The stromal CD45⁺ fraction was collected and frozen for subsequent purification by flow cytometry.

TEC induction from iPSCs

iPSCs were seeded at a density of 1300–3200 cells/cm² onto plates coated with iMatrix-511 silk in AK02N medium supplemented with 10 μM Y-27632. On the next day, the medium was changed to AK02N only. Four days after seeding, the induction was started using the same medium and factors as previously reported for AFE induction⁵⁰ up until day 7.

Briefly, the base medium was changed to CDM2, which is composed of a 1:1 mixture of IMDM (Thermo Fisher) and F-12 (Thermo Fisher), with added 1 mg/ml polyvinyl alcohol (Sigma-Aldrich), 0.7 μg/ml insulin (Wako), 15 μg/ml transferrin (Nacalai Tesque), 450 μM monothioglycerol (Sigma-Aldrich), and 1:100 chemically defined lipid (Thermo Fisher). On day 0, CDM2 was supplemented with 100 ng/ml Activin A (Nacalai Tesque), 2 μM CHIR99021 (Nacalai Tesque), and 50 nM PI-103 (Tocris) to induce primitive streak (PS) cells. Activin A is known to drive pluripotent stem cells toward the PS and its derivatives when PI3K signaling is suppressed¹⁵² and WNT signaling is active⁵⁰.

PI-103, a potent PI3K inhibitor, and CHIR99021, a GSK3 inhibitor promoting WNT activation, are capable of effectively specifying the PS⁵⁰. On day 1, the medium was changed to CDM2 with 100 ng/ml Activin A and 250 nM LDN-193189 (Reprocell) to induce DE. Since BMP signaling at this stage leads to the formation of mesoderm¹⁵³, BMP inhibitor LDN-193189 is used to specifically induce DE⁵⁰. On day 3, the medium was changed to CDM2 with 250 nM LDN-193189 and 1 μ M A-83-01 (Tocris) for AFE induction. The AFE fate is promoted by the combined inhibition of BMP and TGF β signaling¹⁵⁴, with A-83-01 being a potent inhibitor of the TGF β pathway effective at specifying AFE⁵⁰.

Next, to induce PE from AFE, the medium was changed to CDM2 supplemented with 150 nM RA (Sigma-Aldrich) and 50 ng/ml FGF8 (R&D Systems) on day 7. RA is responsible for the patterning of the posterior pharyngeal pouches^{58–60} and FGF8 promotes the normal development and morphology of the pharyngeal arches^{61,62}. From day 18 onwards, medium changes were done using CDM2 only to allow for undirected differentiation of all 3rd PP derivatives. The medium was changed every 2–3 days. Throughout the induction, cells were kept in the same plates without passage. Details of all reagents are in Supplementary Data 1.

Establishment of *FOXN1*^{mCherry} reporter lines

FOXN1^{mCherry} reporters were established from the iPSC lines 201B7, 409B2, and 1383D6 using the CRISPR-Cas9 system. Guide RNA (gRNA) with spacer sequence 5'-GCTCCAGCGTTGCTGGTC-3' was designed to target the 3' UTR near the *FOXN1* stop codon. DNA templates for the gRNA were amplified with PCR using Phusion High-Fidelity DNA Polymerase (Thermo Fisher) and purified using the FastGene Gel/PCR Extraction Kit (Nippon Genetics). gRNA was transcribed in vitro using the MEGAshortscript Kit (Thermo Fisher) and purified using the RNeasy MinElute Cleanup Kit (QIAGEN). The donor plasmid DNA, containing the homology arms, the T2A-mCherry-pA sequence, and a puromycin resistance cassette under the EF1 α promoter flanked by two loxP sites, was assembled using the In-Fusion HD Cloning Kit (Takara Bio) with the pENTR-Donor-MCS2 vector¹⁵⁵ as the backbone. The Competent High DH5 α Kit (Toyobo) was used for transformation of *E. coli* with the vector, which was purified using the NucleoBond Xtra Midi EF Kit (Takara Bio).

For genome editing, using previously optimized conditions¹⁵⁶, 1 μ g donor plasmid DNA, 1.25 μ g gRNA, and 5 μ g TrueCut Cas9 Protein v2 (Thermo Fisher) were transfected into the iPSC lines with the P4 Primary Cell 4D-Nucleofector X Kit S (Lonza). After 3–5 days, 0.5 μ g/ml puromycin (Nacalai Tesque) was added to select for candidate clones incorporating the insert. Emerged colonies were randomly picked, expanded, and treated with Cre Recombinase Vesicles (Takara Bio) to remove the floxed puromycin cassette. The resulting cells were re-seeded sparsely, colonies were randomly picked again, and each resulting clone was analyzed by PCR with PrimeSTAR GXL DNA Polymerase (Takara Bio) after DNA extraction with the NucleoSpin Tissue XS Kit (Macherey-Nagel) to determine the presence of the insert. After gel electrophoresis, bands were detected using the FAS-Digi (Nippon Genetics) with an attached ILCE-5100 camera (SONY). Sequencing was performed by Eurofins Genomics (Tokyo, Japan) to confirm the absence of indels in the allele without the insert and the presence of T2A-mCherry in the allele with the insert. Primers used for PCR and sequencing are listed in Supplementary Data 2. Karyotyping was performed by Chromocenter (Yonago, Japan) to confirm the cell identity and the absence of major chromosomal alterations. All *FOXN1*^{mCherry} reporter iPSC lines generated in this study are available upon request.

Flow cytometry

The frozen CD45⁺ thymic stromal cells were defrosted and washed once in TEC buffer. Cells were stained with antibodies diluted 1:50 in TEC buffer for 15 min on ice. After washing with TEC buffer, the cells were transferred to 5 ml tubes through a 35 μ m mesh for analysis with

the BD FACSARIA II. Propidium iodide (PI, Nacalai Tesque) was used at 1 μ g/ml to stain dead cells. To collect primary TECs for qPCR controls, the EPCAM⁺CD31⁺CD45⁺CD235ab⁺ fraction was sorted, centrifuged for 10 min at 150 \times g, 4 $^{\circ}$ C, and resuspended in TRIzol Reagent (Thermo Fisher). To collect primary thymic stroma for scRNA-seq, the EPCAM⁺CD45⁺CD235ab⁺ and EPCAM⁺CD31⁺CD45⁺CD235ab⁺ fractions were sorted separately, centrifuged, resuspended in 0.04% UltraPure BSA (Thermo Fisher) in PBS, and mixed at a ratio of 1:2 to obtain stromal cells enriched for TECs.

For flow cytometry of induced cells on days 3 and 7, the cells were first washed with PBS and dissociated using Accutase (Nacalai Tesque) for 7–10 min at 37 $^{\circ}$ C. Cells were detached with cell scrapers and washed twice with FACS buffer consisting of 2% FBS in PBS. For cells on day 3, antibodies diluted 1:50 in FACS buffer were used to stain for 15 min on ice. After two washes in FACS buffer, cells were passed through a 35 μ m mesh for analysis with the Cytek NL-3000. DAPI was used at 0.5 μ g/ml to stain dead cells. For cells on day 7, Ghost Dye Red 710 (Tonbo Biosciences) was used at 1:1000 in PBS to stain dead cells for 30 min on ice before proceeding to further staining. After one wash in FACS buffer, cells were fixed with Fixation/Permeabilization solution (Thermo Fisher) at room temperature for 30 min. Cells were washed twice using Permeabilization Buffer (Thermo Fisher) and stained with antibodies diluted 1:10 (FOXA2) or 1:100 (SOX2) in Permeabilization Buffer for 30–45 min at room temperature. After two washes with Permeabilization Buffer, cells were resuspended in FACS buffer, passed through a 35 μ m mesh, and analyzed with the Cytek NL-3000.

For flow cytometry of induced cells on day 28–133, the cells were washed with PBS and dissociated using Accumax (Nacalai Tesque) for 20–30 min at 37 $^{\circ}$ C. Cells were detached with cell scrapers and collected in 50 ml tubes containing FACS buffer consisting of 0.1% bovine serum albumin (BSA, Nacalai Tesque) in PBS. Large cell clumps that could not be dissociated even after pipetting were allowed to sink to the bottom of the tubes for 30–60 s and the supernatant was transferred to new tubes. Cells were centrifuged for 5 min at 350 \times g, 4 $^{\circ}$ C, washed once with FACS buffer, and passed through a 35 μ m mesh. Antibodies diluted 1:10 in FACS buffer were used for staining on ice for 20 min. After washing once and resuspending the cells in FACS buffer, analysis was performed using the BD FACSARIA II. DAPI was used at 0.5 μ g/ml to stain dead cells. To collect induced cells for qPCR, the mCherry⁺, mCherry^{low}, and mCherry^{high} fractions were sorted separately, centrifuged for 10 min at 150 \times g, 4 $^{\circ}$ C, and resuspended in TRIzol Reagent. For scRNA-seq, the same fractions were sorted, centrifuged, resuspended in 0.04% UltraPure BSA in PBS, and mixed at a ratio of 1:1:2 to enrich for mCherry⁺ cells.

For flow cytometry of induced cells on days 7, 11, and 18 for scRNA-seq, the cells were processed and analyzed in the same way as described above for cells on day 28–133 with varying lengths of Accutase or Accumax dissociation (7–20 min) depending on cell density, but due to the absence of mCherry⁺ cells, the entire live fraction was sorted and resuspended in 0.04% UltraPure BSA in PBS.

Flow cytometry results were analyzed in BD FACS Diva Software, SpectroFlo, FlowJo, and Microsoft Excel. Gating strategies are shown in Supplementary Fig. 12 and 13. All antibodies, together with their dilutions, catalog numbers, and further details, are listed in Supplementary Data 3.

Organoid culture and TCR repertoire analysis

Human primary thymocytes were isolated by gently vortexing 2–3 mm thymus pieces, after which cells from the supernatant were incubated in RBC Lysis Buffer (BioLegend) for 10 min on ice, washed in 1% FBS in RPMI 1640, and frozen for future use. Stocked thymocytes were defrosted, washed in 1% FBS in PBS, and stained with antibodies diluted 1:10 in the same buffer for 20 min on ice. After washing once more in 1% FBS in PBS, cells were passed through a 35 μ m mesh and analyzed using the BD FACSARIA II or BD FACS SymphoNy S6. PI was used at 1 μ g/ml to

stain dead cells. To isolate DP and DN thymocytes for organoid co-culture, the CD4⁺CD8 α ⁺CD45⁺CD69⁺ and CD5⁺CD7⁺CD4⁺CD8 α ⁺CD11c⁺CD14⁺CD19⁺ populations were respectively sorted and resuspended in RPMI 1640 with 4% B-27 Supplement (Thermo Fisher). To isolate SP thymocytes as donor controls for TCR repertoire analysis ($n=1$ pediatric donor, male), the CD45⁺CD3⁺TCR $\alpha\beta$ ⁺CD4⁺CD8 α ⁺ and CD45⁺CD3⁺TCR $\alpha\beta$ ⁺CD4⁺CD8 α ⁺CD8 β ⁺ populations were sorted at 10,000 cells each and resuspended in ISOGEN-LS (Nippon Gene).

For organoid co-culture, 10,000 DP or 5,000 DN thymocytes and 100,000 MSS-hDLL1 cells or 100,000 sorted mCherry⁺ or mCherry⁻ iTECs from day 37 \pm 2 of induction were mixed, resuspended in 2.5 μ l Matrigel (Corning) per organoid, and seeded onto 0.8 μ m membranes (Cytiva) floating in 12 well plates filled with RPMI 1640, supplemented with 10 μ M Y-27632, 5 ng/ml FLT3L (PeproTech), 5 ng/ml IL7 (PeproTech), 30 μ M L-ascorbic acid 2-phosphate (Sigma-Aldrich), and 4% B-27 Supplement. Y-27632 was removed on the following day, after which the medium was changed every 2–3 days. To collect thymocytes for analysis after co-culture, organoids were homogenized using pestles. Cells from the supernatant were washed in 1% FBS in PBS and stained with antibodies diluted 1:10 in the same buffer for 20 min on ice. After one more wash, cells were passed through a 35 μ m mesh and analyzed with the Cytex Aurora. For TCR repertoire analysis of the SP thymocytes obtained after mCherry⁺ iTEC co-culture with DP thymocytes, the CD45⁺CD3⁺TCR $\alpha\beta$ ⁺CD4⁺CD8 α ⁺ and CD45⁺CD3⁺TCR $\alpha\beta$ ⁺CD4⁺CD8 α ⁺CD8 β ⁺ populations were sorted at 10,000 cells each using the BD FACS Aria II and resuspended in ISOGEN-LS.

TCR repertoire analysis was performed by Repertoire Genesis (Osaka, Japan) as previously described^{157,158}. Briefly, RNA was extracted using the RNeasy Lipid Tissue Mini Kit (Qiagen) and reverse transcribed with Superscript III reverse transcriptase (Thermo Fisher). TCR β genes were amplified with adapter-ligation PCR^{157,158}, index sequences were added using the Nextera XT Index Kit v2 Set A or Set D (Illumina), and sequencing was performed with the MiSeq (Illumina). Sequences of TCR β genes *TRBV* and *TRBJ* were assigned based on the international ImMunoGeneTics information system (IMGT) database¹⁵⁹. All data processing was performed using Repertoire Genesis analysis software^{157,158}. *TRBV* and *TRBJ* gene usage percent was calculated as the frequency of reads in total sequence reads. For the estimation of TCR repertoire diversity, the Shannon index⁹⁷ was calculated with the Repertoire Genesis analysis software.

For qPCR and scRNA-seq, organoids of mCherry⁺ iTECs and DP thymocytes co-cultured for 14 days were dissociated in Accumax for 15 min at 37 °C, washed in 1% FBS in PBS, and analyzed using the BD FACS Symphony S6. DAPI was used at 0.5 μ g/ml to stain dead cells. For qPCR, the mCherry⁻ and mCherry⁺ populations were sorted, centrifuged, and resuspended in TRIzol Reagent. For scRNA-seq, the mCherry⁻ and mCherry⁺ populations were sorted, centrifuged, resuspended in 0.04% UltraPure BSA in PBS, and mixed at a ratio of 1:2. This scRNA-seq dataset includes thymocytes from $n=4$ pediatric donors, 3 female and 1 male.

RNA extraction and quantitative PCR

RNA was extracted from cells resuspended in TRIzol Reagent using the Direct-zol RNA MicroPrep Kit (Zymo Research). 300 ng of extracted total RNA was reverse transcribed using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara Bio). The cDNA was diluted 10 times and used in duplicate or triplicate reactions for qPCR with the THUNDERBIRD Next SYBR qPCR Mix (Toyobo). The results were analyzed with the QuantStudio 12K Flex Real-Time PCR System (Thermo Fisher), Microsoft Excel, and GraphPad Prism. All results were normalized by *ACTB* expression. For all mRNA expression values of induced cells shown as relative to primary TECs, the values were normalized by the mean mRNA expression values of FACS-purified primary TECs (EPCAM⁺CD31⁺CD45⁺CD235ab⁺) from $n=4$ pediatric donors, 2 male and 2 female.

Immunofluorescence staining and microscopy

Induced cells in 24 well plates were washed once in PBS and fixed in 4% paraformaldehyde solution (PFA, Nacalai Tesque) for 15 min at room temperature. After two PBS washes, cells were permeabilized for 15 min at room temperature using 0.4% Triton X-100 (Sigma-Aldrich) in PBS. Cells were washed once in PBST consisting of PBS with 0.1% Tween 20 (Nacalai Tesque) and blocked for 1 h at room temperature with blocking buffer consisting of 1% BSA in PBS. Staining was performed overnight at 4 °C with primary antibodies diluted in blocking buffer. For day 18, EPCAM, TBX1, and HOXA3 antibodies were all used in 1:300 dilutions. For days 28, 80, and 133, the dilutions were 1:150 for CLDN4 and KRT5 (Proteintech), 1:600 for KRT8, PSMB11, EPCAM, HLA-DR, KRT5 (BioLegend), VIM, and Ki-67, and 1:1500 for mCherry. On the next day, cells were washed three times in PBST and stained with secondary antibodies diluted 1:500 and DAPI at 0.2 μ g/ml in blocking buffer for 1 h at room temperature. After one PBST and two PBS washes, cells were stored in PBS at 4 °C until observation using the Keyence BZ-X810.

All phase-contrast images of non-stained and live cells were also taken using the same microscope. For the structured illumination microscopy of 3D-like nodules, images with a pitch of 1.2 μ m over a range of 50–150 μ m were taken using the sectioning and Z-stack functions of the Keyence BZ-X810. 3D displays of Z-stacks, representative of $n=4$ (live cells) or $n=3$ (fixed, stained cells) independent experiments with 3 or more imaged nodules each, were created using the Keyence BZ-X810 analysis software, recorded using OBS Studio, and arranged into videos using Microsoft Clipchamp. For the 3D nodules with co-staining of VIM and mCherry, the original red color of the mCherry staining was changed to magenta for viewer accessibility.

Cell and library preparation for scRNA-seq

Sorted induced cells and primary thymic stroma were applied at 10,000 cells per sample to the Chromium Next GEM Chip K (10x Genomics). Reverse transcription, cDNA amplification, and library construction were performed using the Chromium Next GEM Single Cell 5' GEM Kit v2, Chromium Next GEM Single Cell 5' Gel Bead Kit v2, Library Construction Kit, and Dual Index Kit TT Set A according to the manufacturer's instructions (10x Genomics). Sequencing was performed by the CiRA Foundation (Kyoto, Japan) using the NovaSeq 6000 (Illumina) at 50,000 read pairs per cell.

scRNA-seq data processing and integration

FASTQ files were made and aligned with Cell Ranger using default settings and the refdata-gex-GRCh38-2020-A human reference provided by 10x Genomics. Doublets identified by Scrublet¹⁶⁰ were excluded from further analysis, which was performed with the R package Seurat¹⁶¹. Cells with total RNA counts and unique feature counts in the top or bottom 5%, as well as cells containing >10% mitochondrial counts, were filtered out. This resulted in datasets ranging from 5782–9207 cells per sample and 747–8357 unique features per cell. SCTransform was performed to normalize and scale the data, while regressing out the mitochondrial mapping percentage and cell cycle score.

All samples of the main induction from day 7–133 were merged and PCA was run on the merged object using the union of the top 3000 variable features of each sample as input. The top 50 PCs were retained for UMAP dimensional reduction and clustering with the Louvain method (resolution = 0.5). Clusters were manually annotated based on the expression of canonical markers of each cell type and the module scores of cTEC and mTEC markers defined by Park et al.³⁰. When the same cell type was assigned to more than one cluster, clusters were arbitrarily numbered using Roman numerals I to VI. For primary thymic stroma, obtained from $n=2$ female and $n=1$ male pediatric donors, samples were integrated with anchors identified using RPCA. PCA was run on the integrated object using the top 2000 variable features, with

the top 100 PCs kept for UMAP dimensional reduction and clustering with the Louvain method (resolution = 0.1). Several small lymphocyte clusters containing T cells, B cells, and others were merged into one Lympho cluster for visual clarity. For the iTECs co-cultured with DP thymocytes in organoids, PCA was run using the top 2000 variable features, with the top 100 PCs being retained for UMAP dimensional reduction and clustering with the Louvain method (resolution = 0.4). Clusters were manually annotated based on the expression of known markers and the co-clustering with the day 133 induced cell and primary thymic stroma datasets above, as well as the dataset of thymic epithelium reported in Park et al.³⁰.

To compare induced monoculture or co-culture and primary TECs, SCTransformed samples of induced cells from day 133, primary thymic stroma, and organoid co-cultured cells were each integrated with RPCA using the top 2000 variable features, after which their original cluster annotations were visualized on UMAP. To compare cells from day 38 and 81 of two different inductions, the two SCTransformed samples of the respective induction day were merged, PCA was run using the top 50 PCs with the union of the top 3000 variable features as input, and the result was visualized using UMAP. For subclustering of the TEP cluster of the merged day 7–133 induced dataset, as well as the mTEC II cluster of the iTEC with DP thymocyte co-culture dataset, PCA and UMAP were re-run, clustering was performed with resolution = 0.3 and 0.2, respectively, and the new cluster annotations were added back onto the original UMAP. To compare our co-cultured iTECs with the dataset of thymic epithelium by Park et al.³⁰, the published raw count matrix and cell annotations were downloaded from Zenodo (<https://doi.org/10.5281/zenodo.5500511>). The Park et al. dataset was subset for all cells annotated as thymic epithelium and SCTransformed identically to our dataset. Both datasets were integrated with CCA using the top 2000 variable features and visualized on UMAP with each of their original cluster annotations.

Differential expression and GRN analysis

Cluster markers were found using the Seurat function FindAllMarkers with the Wilcoxon Rank Sum test. All genes with Bonferroni-adjusted p -value < 0.05 and average \log_2 fold change > 0.25 compared to other clusters were deemed to be markers of that cluster. To assess the similarity of gene expression patterns in induced TECs and the Park et al. dataset³⁰, the reported top 20 markers of the cTEC cluster and the combined top 20 markers of the mTEC clusters were imported into a module and a score of each module in induced cells was calculated with Seurat's AddModuleScore function. Module scores were visualized on UMAP.

For GRN analysis, all cells of the TEC lineages were subset and the SCENIC pipeline¹⁶² was run with default parameters. GRNs are computational models of TFs regulating each other and their target genes, representing their interactions as networks. GRNs thereby allow the inference of transcriptional states and cell identities. GRNBoost2¹⁶³ was used to calculate putative GRNs based on gene co-expression, after which only direct-binding interactions were retained based on known TF-binding motifs. The activity of the regulons, which include the TFs and their potential target genes, was then scored using AUCell. The AUC (Area Under the Curve) scores for each cell were applied to the Seurat objects, and cluster-specific regulons were determined using the FindAllMarkers function. Regulons with Bonferroni-adjusted p -value < 0.05 and average \log_2 fold change > 0.01 were deemed to be markers of that cluster.

Trajectory and RNA velocity analysis

Trajectory inference and pseudotime analysis were performed in R with Slingshot¹⁶⁴. For the induced cells from day 7–133, the root was set as the AFE I cluster and end points were set as the cTEC, mTEC I, mTEC II, mTEC III, mTEC IV, mTEC V, and KC clusters, due to their respective presence on days 7 and 133. To determine genes differentially

expressed along pseudotime between different lineages, the fitGAM and patternTest functions in tradeSeq¹⁶⁵ were used. Genes with $\text{waldStat} > 100$ and FDR-adjusted p -value < 0.05 were deemed to be differentially expressed. Differentially expressed genes were visualized using the plotSmoothers function in tradeSeq and the Heatmap function in ComplexHeatmap¹⁶⁶. For RNA velocity analysis, velocity¹⁶⁷ was used to recount spliced and unspliced reads by considering the mapping strands based on the BAM files generated by Cell Ranger. These count matrices were then applied to the preprocessed Seurat objects. The induced TEC lineage clusters from day 7–133 with the top 2000 variable genes were subset and used as input to the scvelo pipeline¹²⁷, which was run with default parameters using the dynamical model. The resulting velocity vector field was displayed on UMAP with scvelo.

For RNA velocity analysis of the iTEC with DP thymocyte co-culture dataset, the *EPCAM*⁺ iTEC cluster with the top 1000 variable genes was subset and analyzed with the scvelo pipeline as described above. The latent time inferred by scvelo was used to set the root as the mTEC II early subcluster and the end points as the mTEC-mim II and mTEC-mim III clusters for Slingshot pseudotime analysis of the neuronal mimetic mTEC-like cells. The fitGAM and associationTest functions in tradeSeq were used to determine the genes whose expression is associated with pseudotime, with mean log fold change > 0.5, $\text{waldStat} > 10$, and FDR-adjusted p -value < 0.05 being used as criteria. Genes were visualized using the plotSmoothers function in tradeSeq.

Ligand-receptor interaction analysis

The induced cTEC, mTEC I, mTEC II, mTEC III, mTEC IV, mTEC V, KC, and Mes clusters from the day 133 sample were subset and used as input for CellPhoneDB¹⁶⁸ with the statistical analysis method. Predicted ligand-receptor interaction pairs were ranked by their interaction score. Dot plots of interaction means and p -values were created with a manually curated selection from the top 50 interactions of ligands from the induced Mes cluster and receptors from the induced TEC lineage clusters.

Statistical analysis and reproducibility

All statistical analyses of scRNA-seq data were conducted within R, Python, and their respective packages as described in the Methods section. No custom algorithms or software were used. For statistical analyses of mRNA and protein expression data, unpaired two-sided t -tests were performed in Microsoft Excel and GraphPad Prism. Results with p -value < 0.05 were considered significant. Definitions of the center and dispersion, as well as other details, are given in the figure legends. “ n ” is defined as the number of independent experiments (biological replicates)—i.e., the number of separate inductions in the case of iPSCs and the number of donors in the case of primary cells—and its value for each experiment is provided in the figure legends. No data randomization, blinding, sample size estimation, test of statistical assumptions, or data exclusion was done.

Reporting summary

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

Data availability

scRNA-seq data of induced and primary TECs were deposited into the GEO database under the accession number GSE275981 and are publicly available at the following URL: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE275981>. RNA-seq data of TCR repertoire analysis were deposited into the GEO database under the accession number GSE275815 and are publicly available at the following URL: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE275815>. Source data are provided with this paper.

References

- Kadouri, N., Nevo, S., Goldfarb, Y. & Abramson, J. Thymic epithelial cell heterogeneity: TEC by TEC. *Nat. Rev. Immunol.* **20**, 239–253 (2020).
- Abramson, J. & Anderson, G. Thymic epithelial cells. *Annu. Rev. Immunol.* **35**, 85–118 (2017).
- Takahama, Y., Ohigashi, I., Baik, S. & Anderson, G. Generation of diversity in thymic epithelial cells. *Nat. Rev. Immunol.* **17**, 295–305 (2017).
- Farley, A. M. et al. Dynamics of thymus organogenesis and colonization in early human development. *Development* **140**, 2015–2026 (2013).
- Nehls, M. et al. Two genetically separable steps in the differentiation of thymic epithelium. *Science* **272**, 886–889 (1996).
- Nehls, M., Pfeifer, D., Schorpp, M., Hedrich, H. & Boehm, T. New member of the winged-helix protein family disrupted in mouse and rat nude mutations. *Nature* **372**, 103–107 (1994).
- Han, J. & Zúñiga-Pflücker, J. C. A 2020 view of thymus stromal cells in T cell development. *J. Immunol.* **206**, 249–256 (2021).
- Michel, C. et al. Revisiting the road map of medullary thymic epithelial cell differentiation. *J. Immunol.* **199**, 3488–3503 (2017).
- Lkhagvasuren, E., Sakata, M., Ohigashi, I. & Takahama, Y. Lymphotoxin β receptor regulates the development of CCL21-expressing subset of postnatal medullary thymic epithelial cells. *J. Immunol.* **190**, 5110–5117 (2013).
- Givony, T. et al. Thymic mimetic cells function beyond self-tolerance. *Nature* **622**, 164–172 (2023).
- Michelson, D. A., Hase, K., Kaisho, T., Benoist, C. & Mathis, D. Thymic epithelial cells co-opt lineage-defining transcription factors to eliminate autoreactive T cells. *Cell* **185**, 2542–2558 (2022).
- Gao, H. et al. The lineage differentiation and dynamic heterogeneity of thymic epithelial cells during thymus organogenesis. *Front. Immunol.* **13**, 805451 (2022).
- Klein, F. et al. Combined multidimensional single-cell protein and RNA profiling dissects the cellular and functional heterogeneity of thymic epithelial cells. *Nat. Commun.* **14**, 4071 (2023).
- Farley, A. M. et al. Thymic epithelial cell fate and potency in early organogenesis assessed by single cell transcriptional and functional analysis. *Front. Immunol.* **14**, 1202163 (2023).
- Michelson, D. A. & Mathis, D. Thymic mimetic cells: tolerogenic masqueraders. *Trends Immunol.* **43**, 782–791 (2022).
- Murata, S. et al. Regulation of CD8⁺ T cell development by thymus-specific proteasomes. *Science* **316**, 1349–1353 (2007).
- Ushio, A. et al. Functionally diverse thymic medullary epithelial cells interplay to direct central tolerance. *Cell Rep.* **43**, 114072 (2024).
- Kadouri, N. et al. Transcriptional regulation of the thymus master regulator Foxn1. *Sci. Immunol.* **7**, eabn8144 (2022).
- Kobrynski, L. J. & Sullivan, K. E. Velocardiofacial syndrome, DiGeorge syndrome: the chromosome 22q11.2 deletion syndromes. *Lancet* **370**, 1443–1452 (2007).
- Pignata, C. et al. Congenital alopecia and nail dystrophy associated with severe functional T-cell immunodeficiency in two sibs. *Am. J. Med. Genet.* **65**, 167–170 (1996).
- Peterson, P., Pitkänen, J., Sillanpää, N. & Krohn, K. Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED): a model disease to study molecular aspects of endocrine autoimmunity. *Clin. Exp. Immunol.* **135**, 348–357 (2004).
- Ohigashi, I. et al. Aire-expressing thymic medullary epithelial cells originate from β 5t-expressing progenitor cells. *Proc. Natl. Acad. Sci. USA* **110**, 9885–9890 (2013).
- Mayer, C. E. et al. Dynamic spatio-temporal contribution of single β 5t⁺ cortical epithelial precursors to the thymus medulla. *Eur. J. Immunol.* **46**, 846–856 (2016).
- Baik, S., Jenkinson, E. J., Lane, P. J. L., Anderson, G. & Jenkinson, W. E. Generation of both cortical and Aire(+) medullary thymic epithelial compartments from CD205(+) progenitors. *Eur. J. Immunol.* **43**, 589–594 (2013).
- Ohigashi, I. et al. Adult thymic medullary epithelium is maintained and regenerated by lineage-restricted cells rather than bipotent progenitors. *Cell Rep.* **13**, 1432–1443 (2015).
- Sekai, M., Hamazaki, Y. & Minato, N. Medullary thymic epithelial stem cells maintain a functional thymus to ensure lifelong central T cell tolerance. *Immunity* **41**, 753–761 (2014).
- Nusser, A. et al. Developmental dynamics of two bipotent thymic epithelial progenitor types. *Nature* **606**, 165–171 (2022).
- Odaka, C. et al. TGF- β type II receptor expression in thymic epithelial cells inhibits the development of Hassall's corpuscles in mice. *Int. Immunol.* **25**, 633–642 (2013).
- Bautista, J. L. et al. Single-cell transcriptional profiling of human thymic stroma uncovers novel cellular heterogeneity in the thymic medulla. *Nat. Commun.* **12**, 1096 (2021).
- Park, J.-E. et al. A cell atlas of human thymic development defines T cell repertoire formation. *Science* **367**, eaay3224 (2020).
- Ragazzini, R. et al. Defining the identity and the niches of epithelial stem cells with highly pleiotropic multilineage potency in the human thymus. *Dev. Cell* **58**, 2428–2446 (2023).
- Zeng, Y. et al. Single-cell RNA sequencing resolves spatio-temporal development of pre-thymic lymphoid progenitors and thymus organogenesis in human embryos. *Immunity* **51**, 930–948 (2019).
- Yayon, N. et al. A spatial human thymus cell atlas mapped to a continuous tissue axis. *Nature* **635**, 708–718 (2024).
- Mohtashami, M. & Zúñiga-Pflücker, J. C. Three-dimensional architecture of the thymus is required to maintain delta-like expression necessary for inducing T cell development. *J. Immunol.* **176**, 730–734 (2006).
- Campinoti, S. et al. Reconstitution of a functional human thymus by postnatal stromal progenitor cells and natural whole-organ scaffolds. *Nat. Commun.* **11**, 6372 (2020).
- Zhu, Z. & Huangfu, D. Human pluripotent stem cells: an emerging model in developmental biology. *Development* **140**, 705–717 (2013).
- Zeleniak, A. et al. De novo construction of T cell compartment in humanized mice engrafted with iPSC-derived thymus organoids. *Nat. Methods* **19**, 1306–1319 (2022).
- Sun, X. et al. Directed differentiation of human embryonic stem cells into thymic epithelial progenitor-like cells reconstitutes the thymic microenvironment in vivo. *Cell Stem Cell* **13**, 230–236 (2013).
- Gras-Peña, R. et al. Human stem cell-derived thymic epithelial cells enhance human T-cell development in a xenogeneic thymus. *J. Allergy Clin. Immunol.* **149**, 1755–1771 (2022).
- Ramos, S. A. et al. Generation of functional human thymic cells from induced pluripotent stem cells. *J. Allergy Clin. Immunol.* **149**, 767–781 (2022).
- Ramos, S. A. et al. Generation of functional thymic organoids from human pluripotent stem cells. *Stem Cell Rep.* **18**, 829–840 (2023).
- Parent, A. V. et al. Generation of functional thymic epithelium from human embryonic stem cells that supports host T cell development. *Cell Stem Cell* **13**, 219–229 (2013).
- Soh, C.-L. et al. FOXP1 (GFP/w) reporter hESCs enable identification of integrin- β 4, HLA-DR, and EpCAM as markers of human PSC-derived FOXP1(+) thymic epithelial progenitors. *Stem Cell Rep.* **2**, 925–937 (2014).
- Su, M. et al. Efficient in vitro generation of functional thymic epithelial progenitors from human embryonic stem cells. *Sci. Rep.* **5**, 9882 (2015).

45. Lai, L. & Jin, J. Generation of thymic epithelial cell progenitors by mouse embryonic stem cells. *Stem Cells* **27**, 3012–3020 (2009).
46. Okabe, M., Ito, S., Nishio, N., Tanaka, Y. & Isobe, K.-I. Thymic epithelial cells induced from pluripotent stem cells by a three-dimensional spheroid culture system regenerates functional T cells in nude mice. *Cell. Reprogram.* **17**, 368–375 (2015).
47. Inami, Y. et al. Differentiation of induced pluripotent stem cells to thymic epithelial cells by phenotype. *Immunol. Cell Biol.* **89**, 314–321 (2011).
48. Otsuka, R. et al. Efficient generation of thymic epithelium from induced pluripotent stem cells that prolongs allograft survival. *Sci. Rep.* **10**, 224 (2020).
49. Provin, N. & Giraud, M. Differentiation of pluripotent stem cells into thymic epithelial cells and generation of thymic organoids: applications for therapeutic strategies against APECED. *Front. Immunol.* **13**, 930963 (2022).
50. Loh, K. M. et al. Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell* **14**, 237–252 (2014).
51. Zorn, A. M. & Wells, J. M. Vertebrate endoderm development and organ formation. *Annu. Rev. Cell Dev. Biol.* **25**, 221–251 (2009).
52. Bleul, C. C. & Boehm, T. BMP signaling is required for normal thymus development. *J. Immunol.* **175**, 5213–5221 (2005).
53. Gordon, J., Patel, S. R., Mishina, Y. & Manley, N. R. Evidence for an early role for BMP4 signaling in thymus and parathyroid morphogenesis. *Dev. Biol.* **339**, 141–154 (2010).
54. Balciunaite, G. et al. Wnt glycoproteins regulate the expression of FoxN1, the gene defective in nude mice. *Nat. Immunol.* **3**, 1102–1108 (2002).
55. Gotoh, S. et al. Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells. *Stem Cell Rep.* **3**, 394–403 (2014).
56. Jacob, A. et al. Differentiation of human pluripotent stem cells into functional lung alveolar epithelial cells. *Cell Stem Cell* **21**, 472–488 (2017).
57. Serra, M. et al. Pluripotent stem cell differentiation reveals distinct developmental pathways regulating lung- versus thyroid-lineage specification. *Development* **144**, 3879–3893 (2017).
58. Wendling, O., Dennefeld, C., Chambon, P. & Mark, M. Retinoid signaling is essential for patterning the endoderm of the third and fourth pharyngeal arches. *Development* **127**, 1553–1562 (2000).
59. Diman, N. Y. S.-G. et al. A retinoic acid responsive Hoxa3 transgene expressed in embryonic pharyngeal endoderm, cardiac neural crest and a subdomain of the second heart field. *PLoS One* **6**, e27624 (2011).
60. Vermot, J., Niederreither, K., Garnier, J.-M., Chambon, P. & Dollé, P. Decreased embryonic retinoic acid synthesis results in a DiGeorge syndrome phenotype in newborn mice. *Proc. Natl. Acad. Sci. USA* **100**, 1763–1768 (2003).
61. Abu-Issa, R., Smyth, G., Smoak, I., Yamamura, K.-I. & Meyers, E. N. Fgf8 is required for pharyngeal arch and cardiovascular development in the mouse. *Development* **129**, 4613–4625 (2002).
62. Frank, D. U. et al. An Fgf8 mouse mutant phenocopies human 22q11 deletion syndrome. *Development* **129**, 4591–4603 (2002).
63. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
64. Yamazaki, Y. et al. PAX1 is essential for development and function of the human thymus. *Sci. Immunol.* **5**, eaax1036 (2020).
65. Zou, D. et al. Patterning of the third pharyngeal pouch into thymus/parathyroid by Six and Eya1. *Dev. Biol.* **293**, 499–512 (2006).
66. Peters, H., Neubüser, A., Kratochwil, K. & Balling, R. Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. *Genes Dev.* **12**, 2735–2747 (1998).
67. Xu, P.-X. et al. Eya1 is required for the morphogenesis of mammalian thymus, parathyroid and thyroid. *Development* **129**, 3033–3044 (2002).
68. Chojnowski, J. L. et al. Multiple roles for HOXA3 in regulating thymus and parathyroid differentiation and morphogenesis in mouse. *Development* **141**, 3697–3708 (2014).
69. Liu, Z., Yu, S. & Manley, N. R. Gcm2 is required for the differentiation and survival of parathyroid precursor cells in the parathyroid/thymus primordia. *Dev. Biol.* **305**, 333–346 (2007).
70. Okita, K. et al. A more efficient method to generate integration-free human iPS cells. *Nat. Methods* **8**, 409–412 (2011).
71. Nakagawa, M. et al. A novel efficient feeder-free culture system for the derivation of human induced pluripotent stem cells. *Sci. Rep.* **4**, 3594 (2014).
72. Sasaki, K. et al. Robust in vitro induction of human germ cell fate from pluripotent stem cells. *Cell Stem Cell* **17**, 178–194 (2015).
73. Swann, J. B., Happe, C. & Boehm, T. Elevated levels of Wnt signaling disrupt thymus morphogenesis and function. *Sci. Rep.* **7**, 785 (2017).
74. Goldfarb, Y. et al. HDAC3 is a master regulator of mTEC development. *Cell Rep.* **15**, 651–665 (2016).
75. Liu, D. et al. Canonical Notch signaling controls the early thymic epithelial progenitor cell state and emergence of the medullary epithelial lineage in fetal thymus development. *Development* **147**, dev178582 (2020).
76. Onder, L. et al. Alternative NF- κ B signaling regulates mTEC differentiation from podoplanin-expressing precursors in the cortico-medullary junction. *Eur. J. Immunol.* **45**, 2218–2231 (2015).
77. Shen, H. et al. Medullary thymic epithelial NF- κ B-inducing kinase (NIK)/IKK α pathway shapes autoimmunity and liver and lung homeostasis in mice. *Proc. Natl. Acad. Sci. USA* **116**, 19090–19097 (2019).
78. Sasaki, K. et al. Thymoproteasomes produce unique peptide motifs for positive selection of CD8(+) T cells. *Nat. Commun.* **6**, 7484 (2015).
79. Gommeaux, J. et al. Thymus-specific serine protease regulates positive selection of a subset of CD4+ thymocytes. *Eur. J. Immunol.* **39**, 956–964 (2009).
80. Ueno, T. et al. CCR7 signals are essential for cortex-medulla migration of developing thymocytes. *J. Exp. Med.* **200**, 493–505 (2004).
81. Kozai, M. et al. Essential role of CCL21 in establishment of central self-tolerance in T cells. *J. Exp. Med.* **214**, 1925–1935 (2017).
82. Lopes, N., Sergé, A., Ferrier, P. & Irla, M. Thymic crosstalk coordinates medulla organization and T-cell tolerance induction. *Front. Immunol.* **6**, 365 (2015).
83. Hikosaka, Y. et al. The cytokine RANKL produced by positively selected thymocytes fosters medullary thymic epithelial cells that express autoimmune regulator. *Immunity* **29**, 438–450 (2008).
84. O'Neill, K. E. et al. Foxn1 is dynamically regulated in thymic epithelial cells during embryogenesis and at the onset of thymic involution. *PLoS One* **11**, e0151666 (2016).
85. Odaka, C. et al. Keratin 8 is required for the maintenance of architectural structure in thymus epithelium. *PLoS One* **8**, e75101 (2013).
86. Hamazaki, Y. et al. Medullary thymic epithelial cells expressing Aire represent a unique lineage derived from cells expressing claudin. *Nat. Immunol.* **8**, 304–311 (2007).
87. Li, Y. et al. Unraveling the spatial organization and development of human thymocytes through integration of spatial transcriptomics and single-cell multi-omics profiling. *Nat. Commun.* **15**, 7784 (2024).
88. Grigorjeva, I. V. & Thakker, R. V. Transcription factors in parathyroid development: lessons from hypoparathyroid disorders. *Ann. N. Y. Acad. Sci.* **1237**, 24–38 (2011).

89. Shitara, S. et al. IL-7 produced by thymic epithelial cells plays a major role in the development of thymocytes and TCR $\gamma\delta$ + intraepithelial lymphocytes. *J. Immunol.* **190**, 6173–6179 (2013).
90. Hozumi, K. et al. Delta-like 4 is indispensable in thymic environment specific for T cell development. *J. Exp. Med.* **205**, 2507–2513 (2008).
91. Shakib, S. et al. Checkpoints in the development of thymic cortical epithelial cells. *J. Immunol.* **182**, 130–137 (2009).
92. Seet, C. S. et al. Generation of mature T cells from human hematopoietic stem and progenitor cells in artificial thymic organoids. *Nat. Methods* **14**, 521–530 (2017).
93. Montel-Hagen, A. et al. Organoid-induced differentiation of conventional T cells from human pluripotent stem cells. *Cell Stem Cell* **24**, 376–389 (2019).
94. Trotman-Grant, A. C. et al. DL4- μ beads induce T cell lineage differentiation from stem cells in a stromal cell-free system. *Nat. Commun.* **12**, 5023 (2021).
95. Cheroutre, H. Starting at the beginning: new perspectives on the biology of mucosal T cells. *Annu. Rev. Immunol.* **22**, 217–246 (2004).
96. Van Coppenolle, S. et al. Notch induces human T-cell receptor $\gamma\delta$ + thymocytes to differentiate along a parallel, highly proliferative and bipotent CD4 CD8 double-positive pathway. *Leukemia* **26**, 127–138 (2012).
97. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
98. Oh, S. et al. Distinct subpopulations of DN1 thymocytes exhibit preferential $\gamma\delta$ T lineage potential. *Front. Immunol.* **14**, 1106652 (2023).
99. Hale, L. P. & Markert, M. L. Corticosteroids regulate epithelial cell differentiation and Hassall body formation in the human thymus. *J. Immunol.* **172**, 617–624 (2004).
100. White, A. J. et al. Lymphotoxin signals from positively selected thymocytes regulate the terminal differentiation of medullary thymic epithelial cells. *J. Immunol.* **185**, 4769–4776 (2010).
101. Langbein, L. et al. Tight junction-related structures in the absence of a lumen: occludin, claudins and tight junction plaque proteins in densely packed cell formations of stratified epithelia and squamous cell carcinomas. *Eur. J. Cell Biol.* **82**, 385–400 (2003).
102. Wang, J. et al. Hassall's corpuscles with cellular-senescence features maintain IFN α production through neutrophils and pDC activation in the thymus. *Int. Immunol.* **31**, 127–139 (2019).
103. Ohigashi, I., Matsuda-Lennikov, M. & Takahama, Y. Peptides for T cell selection in the thymus. *Peptides* **146**, 170671 (2021).
104. Dooley, J., Erickson, M. & Farr, A. G. Alterations of the medullary epithelial compartment in the Aire-deficient thymus: implications for programs of thymic epithelial differentiation. *J. Immunol.* **181**, 5225–5232 (2008).
105. Alawam, A. S., Anderson, G. & Lucas, B. Generation and regeneration of thymic epithelial cells. *Front. Immunol.* **11**, 858 (2020).
106. Kadakia, T. et al. E-protein-regulated expression of CXCR4 adheres preselection thymocytes to the thymic cortex. *J. Exp. Med.* **216**, 1749–1761 (2019).
107. Zhang, X. et al. CD24 on thymic APCs regulates negative selection of myelin antigen-specific T lymphocytes. *Eur. J. Immunol.* **42**, 924–935 (2012).
108. Magaletta, M. E. et al. Integration of single-cell transcriptomes and chromatin landscapes reveals regulatory programs driving pharyngeal organ development. *Nat. Commun.* **13**, 457 (2022).
109. Lammers, S. et al. Ehf and Fezf2 regulate late medullary thymic epithelial cell and thymic tuft cell development. *Front. Immunol.* **14**, 1277365 (2024).
110. Sun, S. et al. CD90 marks a mesenchymal program in human thymic epithelial cells in vitro and in vivo. *Front. Immunol.* **13**, 846281 (2022).
111. Brunk, F., Augustin, I., Meister, M., Boutros, M. & Kyewski, B. Thymic epithelial cells are a nonredundant source of Wnt ligands for thymus development. *J. Immunol.* **195**, 5261–5271 (2015).
112. James, K. D., Jenkinson, W. E. & Anderson, G. Non-epithelial stromal cells in thymus development and function. *Front. Immunol.* **12**, 634367 (2021).
113. Yang, J. et al. Epithelial-mesenchymal transition in age-associated thymic involution: mechanisms and therapeutic implications. *Ageing Res. Rev.* **92**, 102115 (2023).
114. Liu, Z. et al. Thymus-associated parathyroid hormone has two cellular origins with distinct endocrine and immunological functions. *PLoS Genet.* **6**, e1001251 (2010).
115. Hamazaki, Y., Sekai, M. & Minato, N. Medullary thymic epithelial stem cells: role in thymic epithelial cell maintenance and thymic involution. *Immunol. Rev.* **271**, 38–55 (2016).
116. Hamazaki, Y. Adult thymic epithelial cell (TEC) progenitors and TEC stem cells: models and mechanisms for TEC development and maintenance. *Eur. J. Immunol.* **45**, 2985–2993 (2015).
117. Ferreira, P. et al. A novel method to identify Post-Aire stages of medullary thymic epithelial cell differentiation. *Eur. J. Immunol.* **51**, 311–318 (2021).
118. Demarchi, F., Verardo, R., Varnum, B., Brancolini, C. & Schneider, C. Gas6 anti-apoptotic signaling requires NF- κ B activation. *J. Biol. Chem.* **276**, 31738–31744 (2001).
119. Wang, X. et al. Crosstalk between Akt and NF- κ B pathway mediates inhibitory effect of gas6 on monocytes-endothelial cells interactions stimulated by P. gingivalis-LPS. *J. Cell. Mol. Med.* **24**, 7979–7990 (2020).
120. Golson, M. L. & Kaestner, K. H. Fox transcription factors: from development to disease. *Development* **143**, 4558–4570 (2016).
121. Brancaccio, A. et al. Requirement of the forkhead gene Foxe1, a target of sonic hedgehog signaling, in hair follicle morphogenesis. *Hum. Mol. Genet.* **13**, 2595–2606 (2004).
122. Eichberger, T. et al. FOXE1, a new transcriptional target of GLI2 is expressed in human epidermis and basal cell carcinoma. *J. Invest. Dermatol.* **122**, 1180–1187 (2004).
123. Hollnagel, A., Oehlmann, V., Heymer, J., R  ther, U. & Nordheim, A. Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. *J. Biol. Chem.* **274**, 19838–19845 (1999).
124. Yang, J. et al. Id proteins are critical downstream effectors of BMP signaling in human pulmonary arterial smooth muscle cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **305**, L312–L321 (2013).
125. Kousa, A. I. et al. Age-related epithelial defects limit thymic function and regeneration. *Nat. Immunol.* **25**, 1593–1606 (2024).
126. Chen, R. et al. CD147 deficiency in T cells prevents thymic involution by inhibiting the EMT process in TECs in the presence of TGF β . *Cell. Mol. Immunol.* **18**, 171–181 (2021).
127. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).
128. Heimli, M. et al. Multimodal human thymic profiling reveals trajectories and cellular milieu for T agonist selection. *Front. Immunol.* **13**, 1092028 (2023).
129. Tao, W. et al. Insm1 regulates mTEC development and immune tolerance. *Cell. Mol. Immunol.* **20**, 1472–1486 (2023).
130. Kearns, N. A. et al. Generation and molecular characterization of human pluripotent stem cell-derived pharyngeal foregut endoderm. *Dev. Cell* **58**, 1801–1818 (2023).
131. Roberts, C., Ivins, S. M., James, C. T. & Scambler, P. J. Retinoic acid down-regulates Tbx1 expression in vivo and in vitro. *Dev. Dyn.* **232**, 928–938 (2005).
132. Gordon, J., Bennett, A. R., Blackburn, C. C. & Manley, N. R. Gcm2 and Foxn1 mark early parathyroid- and thymus-specific domains in the developing third pharyngeal pouch. *Mech. Dev.* **103**, 141–143 (2001).

133. Trueba, S. S. et al. PAX8, TITF1, and FOXE1 gene expression patterns during human development: new insights into human thyroid development and thyroid dysgenesis-associated malformations. *J. Clin. Endocrinol. Metab.* **90**, 455–462 (2005).
134. Herriges, M. & Morrisey, E. E. Lung development: orchestrating the generation and regeneration of a complex organ. *Development* **141**, 502–513 (2014).
135. Weaver, M., Yingling, J. M., Dunn, N. R., Bellusci, S. & Hogan, B. L. M. Bmp signaling regulates proximal-distal differentiation of endoderm in mouse lung development. *Development* **126**, 4005–4015 (1999).
136. Kuwahara, A. et al. Delineating the early transcriptional specification of the mammalian trachea and esophagus. *eLife* **9**, e55526 (2020).
137. Dathan, N., Parlato, R., Rosica, A., De Felice, M. & Di Lauro, R. Distribution of the *tif2/foxe1* gene product is consistent with an important role in the development of foregut endoderm, palate, and hair. *Dev. Dyn.* **224**, 450–456 (2002).
138. Minoo, P., Su, G., Drum, H., Bringas, P. & Kimura, S. Defects in tracheoesophageal and lung morphogenesis in *Nkx2.1(-/-)* mouse embryos. *Dev. Biol.* **209**, 60–71 (1999).
139. Alves, N. L. et al. Serial progression of cortical and medullary thymic epithelial microenvironments. *Eur. J. Immunol.* **44**, 16–22 (2014).
140. Ulyanchenko, S. et al. Identification of a bipotent epithelial progenitor population in the adult thymus. *Cell Rep.* **14**, 2819–2832 (2016).
141. Wong, K. et al. Multilineage potential and self-renewal define an epithelial progenitor cell population in the adult thymus. *Cell Rep.* **8**, 1198–1209 (2014).
142. Ohigashi, I. et al. Developmental conversion of thymocyte-attracting cells into self-antigen-displaying cells in embryonic thymus medulla epithelium. *eLife* **12**, RP92552 (2024).
143. Lucas, B. et al. Embryonic keratin19+ progenitors generate multiple functionally distinct progeny to maintain epithelial diversity in the adult thymus medulla. *Nat. Commun.* **14**, 2066 (2023).
144. Depreter, M. G. L. et al. Identification of Plet-1 as a specific marker of early thymic epithelial progenitor cells. *Proc. Natl. Acad. Sci. USA* **105**, 961–966 (2008).
145. Davies, E. G. et al. Thymus transplantation for complete DiGeorge syndrome: European experience. *J. Allergy Clin. Immunol.* **140**, 1660–1670 (2017).
146. Kreins, A. Y. et al. European Society for Immunodeficiencies guidelines for the management of patients with congenital athymia. *J. Allergy Clin. Immunol.* **154**, 1391–1408 (2024).
147. Fox, T. A., Houghton, B. C. & Booth, C. Gene edited T cell therapies for inborn errors of immunity. *Front. Genome Ed.* **4**, 899294 (2022).
148. Sun, D. et al. CAR-T cell therapy: a breakthrough in traditional cancer treatment strategies (review). *Mol. Med. Rep.* **29**, 47 (2024).
149. Iriguchi, S. et al. A clinically applicable and scalable method to regenerate T-cells from iPSCs for off-the-shelf T-cell immunotherapy. *Nat. Commun.* **12**, 430 (2021).
150. Duah, M. et al. Thymus degeneration and regeneration. *Front. Immunol.* **12**, 706244 (2021).
151. Ohigashi, I., Matsuda-Lennikov, M. & Takahama, Y. Large-scale isolation of mouse thymic epithelial cells. *Methods Mol. Biol.* **2580**, 189–197 (2023).
152. McLean, A. B. et al. Activin A efficiently specifies definitive endoderm from human embryonic stem cells only when phosphatidylinositol 3-kinase signaling is suppressed. *Stem Cells* **25**, 29–38 (2007).
153. Bernardo, A. S. et al. BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. *Cell Stem Cell* **9**, 144–155 (2011).
154. Green, M. D. et al. Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. *Nat. Biotechnol.* **29**, 267–272 (2011).
155. Li, H. L. et al. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Rep.* **4**, 143–154 (2015).
156. Kagita, A. et al. Efficient ssODN-mediated targeting by avoiding cellular inhibitory RNAs through precomplexed CRISPR-Cas9/sgRNA ribonucleoprotein. *Stem Cell Rep.* **16**, 985–996 (2021).
157. Kitaura, K., Shini, T., Matsutani, T. & Suzuki, R. A new high-throughput sequencing method for determining diversity and similarity of T cell receptor (TCR) α and β repertoires and identifying potential new invariant TCR α chains. *BMC Immunol.* **17**, 38 (2016).
158. Kitaura, K. et al. Different somatic hypermutation levels among antibody subclasses disclosed by a new next-generation sequencing-based antibody repertoire analysis. *Front. Immunol.* **8**, 389 (2017).
159. Lefranc, M.-P. et al. IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res.* **37**, D1006–D1012 (2009).
160. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291 (2019).
161. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
162. Van de Sande, B. et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020).
163. Moerman, T. et al. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* **35**, 2159–2161 (2019).
164. Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
165. Van den Berge, K. et al. Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* **11**, 1201 (2020).
166. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
167. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
168. Vento-Tormo, R. et al. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018).

Acknowledgements

We are grateful to Erina Yamaguchi, Yurika Matsuda, Tomoko Ikari, and Miho Shimada for experimental assistance, to Satoko Sakurai for assistance with scRNA-seq data processing, to Tomoya Matsumoto for imaging assistance, to Dr. Izumi Ohigashi and Dr. Yousuke Takahama for technical help, and to Dr. Yoshiya Kawaguchi for helpful discussions. This study was supported by JSPS KAKENHI Grant Numbers JP23KK0136 to Y.H., T.Y., and Y.P., JP22J13683 to Y.P., JP22KJ1841 to Y.P., JP25K18473 to Y.P., and JP20K06469 to K.Kometani; AMED Grant Numbers JP23bm1323001 to Y.H., T.Y., and A.H., JP20bm0704043 to Y.H., and JP23bm1123022 to Y.H.; JST SPRING Grant Number JPMJSP2110 to Y.G. and K.Kanai; the JSPS WISE program “The Graduate Program for Medical Innovation (MIP)” to Y.G.; the Takeda Science Foundation to Y.H.; and the iPS Cell Research Fund to Y.H.

Author contributions

Conceptualization, Y.P. and Y.H.; Methodology, Y.P., K.Kometani, H.X., and A.H.; Software, Y.P. and T.Y.; Validation, Y.P., K.Kanai, M.O., and K.N.; Formal Analysis, Y.P., Y.G., and T.Y.; Investigation, Y.P., Y.G., and K.Kanai; Resources, T.Y., T.I., and A.H.; Data Curation, Y.P.; Writing – Original Draft, Y.P.; Writing – Review & Editing, Y.P. and Y.H.; Visualization, Y.P.;

Supervision, Y.H.; Project Administration, Y.P. and Y.H.; Funding Acquisition, Y.P., K.Kometani, and Y.H.

Competing interests

Y.P., Y.G., and Y.H. are named inventors of patent application no. PCT/JP2023/021888, related to the induction method of thymic epithelial cells, filed by Kyoto University. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-62523-1>.

Correspondence and requests for materials should be addressed to Yoko Hamazaki.

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

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

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025