Identification of CD177/NB1^{+} specified anterior definitive endoderm improves human pancreatic and β-like cell differentiation *in vitro*

Pallavi U. Mahaddalkar[1,2#], Katharina Scheibner[1#], Sandra Pfluger[1§], Ansarullah[1§], Michael Sterr[1,2], Julia Beckenbauer[1], Martin Irmler[4], Johannes Beckers[4,5,7], Sebastian Knöbel[3] & Heiko Lickert[1,2,6,7*]

1. **Institute of Diabetes and Regeneration Research, Helmholtz Diabetes Center, Helmholtz Zentrum München, 85764 Neuherberg, Germany**
2. **Institute of Stem Cell Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany**
3. **Miltenyi Biotec, 51429 Bergisch Gladbach, Germany**
4. **Institute of Experimental Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany**
5. **Chair of Experimental Genetics, School of Life Sciences Weihenstephan, Technische Universität München, 85354 Freising, Germany**
6. **Chair of β-Cell Biology, Technical University of Munich, School of Medicine, Klinikum Rechts der Isar, Ismaningerstraße 22, 81675 München, Germany**
7. **German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany**

 # contribute equally to this work

§ contribute equally to this work

***Corresponding author: Tel. no.: +49 89 3187 3760**

E-mail: heiko.lickert@helmholtz-muenchen.de

Abstract

Morphogen gradients pattern the endoderm and specify liver and pancreatic progenitors *in vivo*. However, if specified organ progenitors can be identified and isolated during human pluripotent stem cell (hPSC) differentiation is unknown. Here, we report the identification of two novel surface markers, CD177/NB1 glycoprotein and inducible T cell co-stimulatory ligand CD275/ICOSL, that isolate specified organ progenitors from seemingly homogenous endoderm differentiations *in vitro*. These markers allow assessing anterior definitive endoderm (ADE) patterning and specification in human revealing different morphogen requirements and induction efficiencies for the generation of specified pancreatic and liver progenitors using known and novel differentiation paradigms. Furthermore, molecular profiling and characterisation of $CD177^+$ and $CD275^+$ ADE subpopulations identified differential expression of signalling components and inverse activation of canonical and non-canonical WNT signalling. This signalling milieu specifies $CD275^+$ ADE progenitors towards the liver fate. In contrast, $CD177^+$ ADE progenitors express and synthesize the secreted WNT, NODAL and BMP antagonist CERBERUS1 and are specified towards the pancreatic fate. Strikingly, isolated $CD177^+$ ADE progenitors differentiate more homogenously into pancreatic progenitors as well as into functionally, more mature and glucose-responsive β-like cells *in vitro*, when compared to bulk endoderm differentiations. Overall, the identification of novel surface markers allowed us to isolate, monitor and understand human organ progenitor formation for the improved differentiation of β-like cells from hPSC.

Keywords: endoderm patterning, progenitors, pancreas differentiation, stem cell-derived β-like cell differentiation, Wnt signalling, surface marker, CD177, CD275 and CXCR4

Introduction

Endoderm patterning in the mouse embryo directly translates into formation of endoderm-derived organs along the anterior-posterior (A-P) axis, including the thymus, thyroid, lungs, liver, pancreas and gastro-intestinal tract[1-3]. High Nodal and Wnt signalling during embryonic patterning and germ layer formation activates the endoderm transcription factors FoxA2, Sox17, Eomes and Gata4/6 which leads to the execution of an endoderm program[4]. A series of morphogenetic events promotes the transformation of naïve endoderm into the primitive gut tube. During this period, the primitive gut tube is patterned along the A-P and dorso-ventral (D-V) axis and is divided into fore-, mid- and hindgut domains[1]. Fate mapping studies have revealed the fate of endoderm-derived organs is specified already shortly after gastrulation[5,6]. It is thought that liver and pancreas progenitors are specified and originate from a multipotent population of anterior definitive endoderm (ADE) depending on the inductive cues of Fgf, Bmp and Wnt received from neighbouring tissues[7]. Moderate doses of Fgf signalling induces albumin (Alb) expression and liver bud formation, whereas low levels of Fgf promotes Pdx1 expression and the formation of the ventral pancreatic bud[8]. Similarly, a gradient of Bmp and retinoic acid (RA) signalling further specifies liver versus pancreatic fate. High Bmp signalling promotes hepatic differentiation and represses pancreatic fate, while lack of RA fails to induce the dorsal pancreatic bud and generates impaired liver growth[7,9]. Likewise, the presence of non-canonical Wnt signalling in the foregut endoderm and pancreatic progenitors but not in liver progenitors suggested cell type-specific activation of the non-canonical Wnt signalling for progenitor specification and determination[10]. However, if human endoderm derived organ progenitors can be identified in culture is not known. Furthermore, the upstream signals that direct these progenitors into the $PDX1^{+}$ pancreatic lineage remain elusive, especially in human development which is not accessible at this stage[11].

Translation of knowledge gained from *in vivo* studies of endoderm and pancreas development have guided the *in vitro* differentiation of hPSCs towards glucose-responsive and insulin-secreting β-like cells[12,13]. Over the years, the differentiation protocols were steadily improved and recently up to 50% of insulin-producing β-like cells can be generated routinely[14-16]. However, many β-like cells still remain immature and their response to static glucose-stimulated insulin secretion (GSIS) is not comparable to β cells in human islets[14]. One common problem of bulk stem cell differentiation is the remarkable heterogeneity of both undifferentiated and differentiated hPSCs[17,18]. Enrichment of correct populations can enhance the differentiation towards a desired lineage and reduce unwanted differentiation towards other lineages[19,20]. Previous surface marker screens at pancreatic endoderm stage showed that enrichment of pancreatic endoderm and endocrine cells can enhance pancreatic progenitor differentiation towards hormone-positive and glucose-responsive cells[21-23]. Thus, we hypothesize that cellular heterogeneity has a significant impact on endoderm patterning and pancreatic differentiation.

To resolve endoderm heterogeneity we aimed to ascertain novel surface markers for the identification and characterisation of endoderm subpopulations with the aim of detecting endoderm-derived organ progenitors in culture. The recognition of CD177 and CD275 as specific surface markers for ADE subpopulations allowed us to confirm that bulk endoderm differentiation cultures from human induced pluripotent stem cells (hiPSCs) and human embryonic stem cells (hESC) are very heterogeneous in nature. Detailed analysis of the $CD177^+$ and $CD275^+$ ADE progenitors during early differentiation further allowed us to understand the signals and factors that induce ADE subpopulations and guide liver and pancreatic differentiations in human. The enrichment of $CD177^+$ ADE permitted for more homogenous pancreatic differentiation which resulted in improved maturation and function of insulin-producing β-like cells.

Results

Identification of novel surface markers to distinguish specified ADE subpopulations

Self-organisation and spatial patterning has been recently studied using hESCs[24]. Current hESC differentiation protocols induce a seemingly homogenous population of definitive endoderm (DE) when measured by specific marker genes, i.e. CXCR4, CXCR4/CD117 or FOXA2/SOX17 at day 4 of differentiation (D4) (Fig. 1a-c). To investigate whether endoderm differentiation is homogenous we used fluorescent activated cell sorting (FACS) based on CXCR4 (CD184) and c-KIT (CD117) marker expression and analysed the induced endoderm by quantitative PCR (qPCR) (Fig. 1b and 1d-g). Surprisingly, CD117 marks only a subset of $CXCR4^{+}$ endoderm (63.3%) and the $CXCR4^{+}/CD117^{+}$ double positive population can be divided into $CXCR4^{high}/CD117^{high}$ (21.3%), $CXCR4^{mid}/CD117^{mid}$ (15.4%), $CXCR4^{low}/CD117^{low}$ (26.6%) and $CXCR4^{+}/CD117^{-}$ (15.3%) subpopulations (Fig. 1b). We found inverse expression levels of *FOXA2* and *SOX17* in the sorted $CXCR4^{high}/CD117^{high}$ to $CXCR4^{low}/CD117^{low}$ subpopulations (Fig. 1d, e) and increased expression of the ADE markers *CERBERUS1* (*CER1*) and hematopoietically expressed homeobox (*HHEX*) in $CXCR4^{+}/CD117^{-}$ and $CXCR4^{high}/CD117^{high}$ subpopulations, respectively (Fig. 1f, g). FoxA2 and Sox17 are expressed along opposite A-P gradients in the endoderm during mouse gastrulation[25]. These results suggest that the endoderm is molecular heterogeneous and receives some kind of pattern information in culture, likely through neighbouring cell-type interactions[24], similar to the endoderm *in vivo*[26].

We speculated that the heterogeneity in endoderm can be a result of specified organ progenitors generated in the culture. Endoderm fate mapping studies in mouse embryos have presented already specified liver, pancreas and intestinal progenitors shortly after gastrulation[5,27]. To resolve endoderm heterogeneity and isolate different endoderm organ progenitors *in vitro*, we performed a screen to find novel surface markers that distinguish specified endoderm

subpopulations. We screened a library of 330 monoclonal antibodies (mAbs) directed against surface epitopes to identify mAbs that detect subpopulations of DE ($FOXA2^+/CXCR4^+$) in H9 hESC at D4 (Fig. 1h, Supplementary Fig. 1a,b). These antibodies included known cluster of differentiation markers, integrins and other cell surface markers. In general, CXCR4 marks the entire DE in mouse and humans and is commonly used as a standard marker for quantifying DE during *in vitro* differentiations[13,28]. The use of FOXA2 not only allowed us to identify novel antibodies that identify DE ($FOXA2^{high}/CXCR4^+$), but also exclude mesendoderm progenitors ($FOXA2^{low}/CXCR4^-$) present in differentiation cultures (Supplementary Fig. 1b). From the initial screen, we discovered 30 mAbs which marked distinct subpopulations of the DE (Supplementary Table 1, Supplementary Fig. 1a). These markers were selected based on mean intensity of expression in the $FOXA2^{high}/CXCR4^{high}$ population, while being preferentially lower expressed in the $FOXA2^{low}$ and $FOXA2^-/CXCR4^-$ populations. As we were specifically interested in ADE-derived organ progenitors, the secondary screen was focused on mAbs that identify ADE subpopulations and partially co-express the ADE marker CER1 (Supplementary Fig. 1e). This identified glycoprotein CD177/NB1, a molecule required for activation of neutrophils[29], as a candidate marker due to its highly dynamic expression range within the DE population ($FOXA2^{high}/CXCR4^{high}$) and marking ~49% of $CER1^+$ ADE subpopulation in $CXCR4^+$ DE (Supplementary Fig. 1c, f). Additionally, CD275/ICOSL, a factor required for T cell generation[30], showed a somewhat broader distribution marking a major subpopulation of $CXCR4^+$ cells and ~24% of the $CER1^+$ ADE subpopulation in $CXCR4^+$ DE (Supplementary Fig. 1d, f). mRNA quantification of the enriched $CD177^+$ and $CD275^+$ subpopulations by magnetic activated cell sorting (MACS) revealed high expression of the endoderm marker genes *FOXA2* and *SOX17*, when compared to the $CXCR4^+$ sorted DE population (Supplementary Fig. 1g). Interestingly, the $CD177^+$-enriched subpopulation showed higher levels of the ADE marker *CER1*, while the $CD275^+$-enriched subpopulation showed higher levels of the ADE and liver marker *HHEX*, suggesting that these are distinct ADE

subpopulations (Supplementary Fig. 1h)[26]. Furthermore, in DE cells that are >90% CXCR4$^+$, CD177 marks ~37% and CD275 marks ~20% of cells in H9 hESC differentiation at D4 (Fig. 1h,i). Analysis of several hiPSC and hESC lines affirm that the induction efficiency of pan-endoderm marker CXCR4 reaches up to ~90% in all cell lines making it seemingly homogenous endoderm (Supplementary Fig. 2a,b). However, further measuring of CD177 or CD275 ADE markers in CXCR4$^+$ DE revealed marked differences, suggesting that different ratios of ADE subpopulations are generated in different hiPSC and hESC lines (Supplementary Fig. 2c). When we compared three previously established endoderm differentiation schemes[14,15,31], we observed differences in the efficiencies of the two ADE subpopulations at almost identical levels of CXCR4$^+$ DE expression in the H9 hESC and HMGUi001-A hiPSC (Supplementary Fig. 3). Taken together, these results suggest that we discovered two novel surface markers that identify distinct ADE subpopulations in the bulk endoderm population, which are induced to different extend depending on cell line and differentiation protocol.

CD177 and CD275 mark molecular distinct ADE subpopulations

To further characterize CD177$^+$-, CD275$^+$-, and CXCR4$^+$- ADE subpopulations we used MACS followed by global gene expression profiling and qPCR (Fig. 2a). Gene expression profiles followed by principal component analysis (PCA) revealed that the enriched CD177$^+$ and CD275$^+$ ADE subpopulations were not only different from each other, but also different to the CXCR4$^+$ DE, demonstrating heterogeneity in endoderm differentiation at D4 (Fig. 2b). Gene ontology (GO) and differential gene expression analysis in three-way comparisons of CD177$^+$, CD275$^+$ and CXCR4$^+$ populations identified genes related to pathways regulating endoderm patterning and proliferation as well as pancreatic cell-fate specification (Fig. 2c-e). These included positive regulation of TGF-β[32], downstream signalling of activated FGFR1 pathway[7], positive regulation of MAPK cascades[33], retinoic acid (RA)-mediated signalling[9] and regulation of canonical and non-canonical WNT/planar cell polarity (PCP) signalling[10].

Non-canonical and canonical Wnt signalling were described to be differentially regulated in mouse pancreas vs liver progenitors, respectively[10]. Due to differential expression of CER1 (Supplementary Fig. 1), a Wnt, Bmp and Nodal signalling antagonist[34] and other WNT signalling components enriched in ADE subpopulations (Fig. 2d-e), we tested whether the CD177$^+$ and CD275$^+$ ADE subpopulations receive differential WNT signalling. Specifically, components of the non-canonical WNT/PCP pathway, such as *DVL2*, *CELSR1*, and *JNK* as well as WNT/PCP ligands, i.e. *WNT5A* and *WNT4* were upregulated in the CD177$^+$ ADE subpopulation (Fig. 2f). In contrast, the expression of canonical ligand *WNT3A* and its target gene *AXIN2* were upregulated in the CD275$^+$ ADE subpopulation (Fig. 2g). Western blot analysis showed increased DVL2 and phospho-JNK levels proposing signalling activation of the WNT/PCP pathway in the CD177$^+$ ADE subpopulation (Fig. 2h, i). Nuclear translocation of β-catenin is a hallmark of canonical WNT activation[35], thus we tested for the distribution of β-catenin in the nucleus and cytoplasm by immunofluorescence in the isolated subpopulations. Remarkably, β-catenin was highly enriched in adherens junctions of CD177$^+$ ADE progenitors, suggesting degradation in the cytoplasm. Conversely, β-catenin accumulation was observed in the nucleus and cytoplasm of CD275$^+$ and CXCR4$^+$ populations (Fig. 2j). These results suggest that CD275$^+$ ADE progenitors receive canonical WNT/β-catenin signalling and might be specified towards the liver fate, while CD177$^+$ ADE progenitors receive non-canonical WNT/PCP signalling and might be specified towards the pancreatic fate[5,10].

CD177 identifies specified pancreatic progenitors in hPSC differentiation

Next, we studied in detail the surface marker expression of CXCR4, CD177 and CD275 during hiPSC-derived pancreatic and endocrine differentiation *in vitro*. The expression of CD177 and CD275 peaked during ADE formation and patterning and gradually decreased during gut tube and liver and pancreatic progenitor formation, where the expression of CD177 and CD275 was almost negligible (Supplementary Fig. 4a). In contrast, we observed high expression of CXCR4

during pancreatic and endocrine differentiation until the formation of hiPSC-derived β-like cells. To test whether the differential activation of cell fate specification pathways in $CD177^+$ and $CD275^+$ ADE subpopulations (Fig. 2) leads to preferential lineage allocation, we explored the *in vitro* differentiation potential towards the liver and pancreatic fate. To test for liver potency we differentiated the MACS sorted cells towards hepatocytes[36] and tested the expression of early liver progenitor genes *AFP*, *TTR* and *HHEX* (Supplementary Fig. 4b, c). Remarkably, the expression of mRNA transcripts for *HHEX*, *AFP* and *TTR* were upregulated in $CXCR4^+$- and $CD275^+$-derived early hepatic progenitors, when compared to $CD177^+$-derived progenitors (Supplementary Fig. 4c). On the other hand, when we tested for pancreatic differentiation potency (Fig. 3a), we noticed more uniform expression of GATA6, an important regulator for pancreatic development[37-39], and PDX1 in $CD177^+$-derived pancreatic progenitors at stage 3 (CD177-PP1) when compared to $CXCR4^+$- and $CD275^+$-derived PP1 (Supplementary Fig. 4d). We also observed that $CD177^+$-derived pancreatic progenitors at stage 4 (CD177-PP2) expressed very high mRNA levels of *PDX1* and *NKX6.1* (Fig. 3b). Immunofluorescence (Fig. 3c) and FACS analysis (Fig. 3d-e) substantiated these findings, suggesting that sorting of the specified $CD177^+$ ADE subpopulations can enhance the differentiation potential towards the pancreatic lineage. To exclude the undirected generation of other endoderm-derived organ progenitors we tested lung ($SOX2^+$) and gut ($CDX2^+$) markers, which were not expressed in any of the sorted subpopulations at S4 (Supplementary Fig. 4e). Moreover, bulk endoderm differentiation of several hiPSC and hESC lines consistently showed a positive correlation of $CD177^+$ ADE at S1 with pancreatic differentiation efficiency at S3 (Supplementary Fig. 5). For further analyses we used the H1 hESC and HMGUI001-A hiPSC cell lines due to their superior pancreatic differentiation efficiencies (Supplementary Fig. 5). Taken together, these results suggest that CD177 not only predicts pancreatic differentiation efficiency at DE stage, but also shows that CD177-sorted ADE subpopulation differentiates more efficiently into pancreatic progenitors.

Canonical WNT inhibits, whereas non-canonical WNT signalling promotes the pancreatic fate

CD275$^+$ ADE receives canonical WNT signalling and is biased towards liver differentiation, whereas CD177$^+$ ADE receives non-canonical WNT signalling and is biased towards the pancreatic fate, consistent with *in vivo* data from mouse liver and pancreas development[10]. To directly test whether modulation of WNT signalling has an impact on pancreatic *in vitro* differentiation from hiPSCs, we added IWP2, a Porcupine inhibitor that selectively inhibits palmitoylation and secretion of Wnt ligands,[40] to our differentiation cultures. We modulated WNT signalling during primitive gut tube (PGT) formation and patterning at S2, when the expression of CD177 and CD275 reach peak levels (Supplementary Fig. 4), and then further differentiated towards PP2 stage (Fig. 4a). Under standard differentiation conditions without IWP2, FACS analysis for PDX1$^+$/NKX6.1$^+$ showed ~75% PP2 cells derived from CD177$^+$-ADE, whereas ~23% were generated from CD275$^+$-ADE and ~45% from CXCR4$^+$-DE (Fig. 4b, Supplementary Table 2). After the addition of IWP2, the percentages of PDX1$^+$/NKX6.1$^+$ PP2 cells derived from CXCR4$^+$-DE (~70%) and CD275$^+$-ADE (~50%) increased drastically, suggesting that the inhibition of WNT ligand secretion promotes pancreatic differentiation (Fig. 4b, c). No differences were observed in the percentages of PDX1$^+$/NKX6.1$^+$ cells in CD177$^+$-PP2s, suggesting that CD177$^+$ADE progenitors, which express CER1, are already shielded from WNT signalling activation, consistent with our previous results (Fig. 2). In addition, CD177$^+$-PP2 showed significantly higher percentages of PDX1$^+$/NKX6.1$^+$ pancreatic progenitors when compared with previously published protocols (Supplementary Table 2).

As IWP2 impedes overall WNT secretion including canonical and non-canonical WNT ligands, we decided to directly analyse the impact of canonical and non-canonical WNT signalling on pancreatic differentiation using signalling specific ligands and small molecules (Fig. 4d). Therefore, we added the canonical WNT ligand WNT3A (20 ng/ml) and the GSK3β inhibitor

CHIR (3 µM) to activate canonical WNT signalling and WNT5A (400 ng/ml) to stimulate non-canonical WNT signalling pathways. Addition of IWP2 increased the percentages of PDX1$^+$/NKX6.1$^+$ PP2 cells from ~30% to ~60% (Fig. 4e-g). In contrast, addition of CHIR (~10%) and WNT3A (~20%) stalled the generation of PP2 cells, while exposure to WNT5A ligand (~40%) improved PP2 differentiation (Fig. 4d-g). These results together suggest that canonical WNT signalling inhibits, whereas non-canonical WNT signalling promotes pancreatic differentiation.

Another hallmark of canonical WNT signalling is the activation of cell-cycle regulators triggering cell proliferation[41] and of non-canonical WNT signalling to promote cell-cycle exit[42]. Cell proliferation assessment in CD177$^+$-, CD275$^+$- and CXCR4$^+$-PP2s by EdU pulse labelling revealed ~15% of CD177$^+$-PP2s were EdU$^+$ compared to ~50% of CD275$^+$- and ~60% of CXCR4$^+$-PP2s (Fig. 4h, i). Suppression of NGN3 during the early stages of pancreatic differentiation is vital to reduce the generation of polyhormonal cells expressing insulin and glucagon[43] and cell-cycle lengthening is essential to induce NGN3 expression[44]. Hence, we tested for the induction of *NGN3* mRNA in CD177$^+$-, CD275$^+$- and CXCR4$^+$-PP2s at S4 and in endocrine progenitors (EPs) at S5. *NGN3* mRNA was tightly regulated in CD177$^+$-PP2s and very efficiently induced in CD177$^+$-EPs at S5, when compared to CXCR4$^+$- and CD275$^+$-PP2s and -EPs (Fig. 4j).

CD177$^+$ specified ADE generates efficiently stem cell-derived β-like cells (SC-β-cells)

To determine whether hiPSC-derived CD177$^+$-ADE not only generates more efficiently pancreatic and endocrine progenitors, but also SC-β-cells we differentiated MACS-enriched DE and ADE towards INS$^+$/NKX6.1$^+$ cells at S7 (Fig. 5a). Since CXCR4$^+$-DE represents the bulk endoderm population, we decided to compare CD177$^+$-ADE to CXCR4$^+$-DE and additionally used unsorted (US) heterogeneous endoderm differentiations to benchmark the state-of-the art differentiation protocols. We added IWP2 to our cultures to optimize the overall

pancreatic differentiation efficiencies. In average, ~62% of CD177$^+$- and ~46% of CXCR4$^+$-β-cells synthesized both INS and NKX6.1 protein (Fig. 5b-d). We also observed similar percentages of INS$^+$/NKX6.1$^+$ cells (~45%) in the US-derived β-like cells (US-β-cells) after the treatment with IWP2 (Fig. 5b-d). Immunostaining and FACS of sectioned clusters revealed most of the cells to be INS$^+$ and only a small subset of cells (~10%) of CD177$^+$-β-cells and (~20%) of US-β-cells were polyhormonal (INS$^+$/GCG$^+$) cells (Fig. 5d, f, g, Supplementary Table 2). We observed only a small population of polyhormonal INS$^+$/SST$^+$ cells in the US-β-cells and CD177$^+$-β-cells (Fig. 5e-g). Finally, we repeated these experiments with the H1 hESC line and confirmed that CD177$^+$-ADE as compared to bulk differentiations shows more homogenous and efficient pancreatic and β-like cell differentiation (Supplementary Fig. 6a-c, f-g).

CD177$^+$ specified ADE generates more mature and functional SC-β-cells

Subsequently, we determined whether enrichment of ADE progenitors during early stages of differentiation not only improved SC-β-cell differentiation efficiency, but also maturation and function. As the establishment of polarity and compaction plays an important role in maturation of β-cells[42,45], we decided to test if 2D vs 3D culture of the sorted cells influences the induction and expression of PDX1 and NKX6.1 during pancreatic differentiation. Therefore, we sorted CD177$^+$ and CXCR4$^+$ DE/ADE and further differentiated these into pancreatic progenitors either in 2D adherent cultures or after re-aggregation in 3D clusters (Supplementary Fig. 7a). We observed a slight increase in the amount of PDX1$^+$/NKX6.1$^+$ double positive PP2s in 3D culture (Supplementary Fig. 7b). Interestingly, we noticed the formation of dense and compact aggregates with defined E-CAD$^+$ adherens junctions in CD177$^+$-β-cells (Supplementary Fig. 7c), in line with the enrichment of the cell-cell adhesion molecule β-catenin in CD177$^+$ ADE progenitors (Fig. 2l). Next, we analysed the expression of maturation markers in SC-β-cells and found elevated levels of *MAFA* and *GLUT1* mRNA transcripts in CD177$^+$-β-cells compared to

US-β-cells, but still significant lower expression when compared to human islets (Fig. 6b). Immunocytochemistry data showed the presence of MAFA in nuclei (Fig. 6c) and GLUT1 in both membrane and cytoplasm (Fig. 6e) in higher percentages in CD177^{+}-β-cells at S7. Quantification by FACS revealed that CD177^{+}-ADE generates more SC-β-cells co-expressing INS and MAFA (~42%) as well as co-expressing INS and GLUT1 (~38%) (Fig. 6d, f). In line with this, H1-derived CD177^{+}-β-cells showed higher percentages of β-cells expressing INS and GLUT1 (~31%) and INS and MAFA (~17%) compared to US-β-cells (Supplementary. Fig. 6d-h). The increased maturation state was further supported by the comparison to previous published protocols[14,15] (Supplementary Table 2), implicating that the isolation of pancreatic specified ADE promotes more homogenous differentiation towards β-like cells when compared to heterogeneous bulk DE (Fig. 6f).

Finally, we tested the functionality of US-β-cells and CD177^{+}-β-cells by assessing their response to glucose stimulation in sequential static and dynamic challenges. At S7, the US- and CD177^{+}-β-cells in 3D clusters strongly stained red for dithizone (DTZ), a zinc chelating reagent that stains the insulin granules (Fig. 7a). In line with the increased number and maturation status (Fig. 6), we recorded significant more insulin content in H1- and hiPSC-derived CD177^{+}-β-cells (Fig. 7b, Supplementary Fig. 8a). Upon exposure to increasing concentrations of glucose (5.6, 11.1 and 20 mM), US-and CD177^{+}-β-cells showed glucose-stimulated insulin secretion (GSIS), but US-β-cells failed to prolong the response with 20 mM glucose (Fig. 7c), implying exhaustion of insulin granules with repeated higher glucose stimuli. Low glucose stimulus of 2.8 mM glucose drastically reduced GSIS in CD177^{+}-β-cells almost to the basal levels (dashed line) pinpointing tight sensation and glucose regulation compared to US-β-cells (Fig. 7c). Next, we performed sequential static GSIS with both hiPSC and H1 hESC lines. We observed that CD177-β-cells were capable of sensing repeated low and high glucose impulses, thus regulating the insulin release depending upon glucose stimulation in both cell lines (Fig. 7d,

Supplementary Fig. 8b). Though hiPSC-derived US-β-cells could regulate the initial shift from 2.8 mM to 20 mM glucose, they failed to show increase GSIS in the 2nd glucose challenge (Fig. 7d). H1 hESC-derived US-β-cells completely failed to regulate insulin release upon different glucose stimulations (Supplementary Fig. 8b).

Ultimately, we supplemented our static assays with dynamic GSIS in two different cell lines; hiPSC and H1 hESC. We thus tested the response of US-β-cells and CD177^{+}-β-cells to secretagogues, such as 20 mM glucose, 10 nM Exendin-4 (Ex-4) and 25 mM KCl in a perifusion system. Upon exposure to 20 mM glucose, CD177^{+}-β-cells consistently displayed a rapid first-phase insulin release as well as a secretory response to the secretagogues Ex-4 and KCl similar to human islets in both cell lines (Fig. 7e, g, Supplementary Fig. 8d). However, after Ex-4 stimulation SC-β-cells were not capable of shutting off the insulin release rapidly in the low glucose condition, indicating either a still immature β-cell status or lacking inhibitory α- and/or δ-cells, when compared to human islets. In agreement with the insulin content of CD177^{+}-β-cells (Fig. 7b, Supplementary Fig. 8a), we consistently observed less insulin secretion compared to human islets in SC-β-cells (Fig. 7e, g, Supplementary Fig. 8d). In comparison, we did not observe a first-phase insulin release in US-β-cells on 20 mM glucose stimulation, but a reliable response to additional Ex-4 induction in the presence of high glucose and KCl hyperpolarization (Fig. 7f, Supplementary Fig.8c). This data collectively suggests that CD177^{+}-β-cells are significantly more mature and functional than SC-β-cells derived by bulk differentiation using hiPSC and hESC cell lines.

Discussion

In this study, by means of novel surface markers (CD177 and CD275), we report remarkable functional heterogeneity of the ADE *in vitro*, the signalling requirements for pancreatic progenitor specification and how sorting of specified organ progenitors at the endoderm stage can improve the differentiation, maturation and function of SC-derived β-like cells *in vitro*.

Although cellular heterogeneity is an obvious origin for uncontrolled pluripotent stem cell differentiations resulting in low numbers of terminally differentiated cell types, it has not been addressed systematically before. We resolved this endoderm heterogeneity by identifying 30 mAbs directed against surface epitopes. Here, we specifically focused on $CD177^+$ and $CD275^+$ ADE subpopulations. This revealed very different quality and quantity of endoderm generated *in vitro* using several hiPSC and hESC lines and different published protocols[14,15,31], suggesting that quality control at early steps of differentiation is warranted to direct ADE and further endoderm-derived organ differentiation. *In vivo,* high TGF-β/Nodal activity promotes the anterior fate, while lower Nodal signalling strength, Bmp signalling from the extra-embryonic region and high canonical Wnt/β-catenin signalling at the posterior side specifies posterior fates[1,46-48]. Interestingly, we observed higher expression of *CER1*, a Nodal, Bmp and Wnt signalling antagonist[49], in $CD177^+$ ADE progenitors, suggesting an active auto-regulatory feedback loop modulating ligand-receptor interactions in these cells and thus fine-tuning the morphogen requirements to maintain the anterior characteristic and fate specification of the $CD177^+$ ADE progenitors[1,50,51]. Hence, we propose that the endoderm is inherently heterogeneous and depending upon the signalling it receives, the endoderm is patterned by intrinsic and extrinsic cues which translate into the formation of distinct organ progenitors. In the future it will be interesting to explore the nature of the other endoderm subpopulations that can be marked by one of our 30 mAbs to identify organ progenitors for the lung, thymus and thyroid and the related signalling cues to induce fate specification and allocation.

373 Endoderm patterning in mice leads to fate specification of organ progenitors[5]. In human,
374 gastrulation and very early lineage commitment cannot be studied as it has happened already
375 before pregnancy can be confirmed[11]. Our study has helped to get first insights into how fine
376 tuning of morphogen gradients lead to patterning and fate specification during human endoderm
377 development *in vitro*. The secretion of the extracellular WNT, BMP and NODAL inhibitor
378 CER1, positive regulation of TGF-β signalling, regulation of canonical and non-canonical
379 Wnt/planar cell polarity prompted the specification of $CD177^+$ ADE progenitors towards the
380 pancreatic fate. On the other hand, upregulation of transcription factor *HHEX* and activation of
381 canonical WNT signalling specified $CD275^+$ ADE subpopulations towards the liver fate.
382 Interestingly, blocking the secretion of WNT ligands promoted, whereas activation of the
383 canonical WNT pathway inhibited the pancreatic fate, thus underpinning the role of canonical
384 WNT inhibition during pancreatic fate specification[10]. The efficient generation of $CD177^+$
385 ADE-derived pancreatic progenitors showed that depleting cell populations that are not
386 specified towards the pancreatic fate can help in increasing the homogeneity of organ
387 progenitors and improve directed differentiation. This is an important consideration when
388 choosing an 2D adherent or 3D cluster differentiation paradigm. For example, even if endoderm
389 induction efficiency is >90%, dragging along mesoderm (contaminating) cells in bulk
390 differentiations can vastly influence the specification and identity of endoderm-derived organ
391 progenitors. Priming of sorted $CD177^+$ ADE progenitors towards pancreatic fate by non-
392 canonical WNT/PCP signalling further translated into efficient pancreatic and endocrine
393 specification of these cells. In mice, loss of Celsr2 and Celsr3, two principle components of
394 WNT/PCP signalling reduced endocrine specification, β cell differentiation and glucose
395 homeostasis[52].

396 Islet architecture and compaction as well as the WNT/PCP pathway trigger β cell maturation
397 and function[42,45]. Recently published protocols for the generation of β-like cells *in vitro* used

an air-liquid interface (ALI) system to maintain apical-basal polarity to generate glucose-responsive β-like cells[14,53]. The pancreatic progenitors generated from CD177$^+$ ADE were more polarised and formed tighter and compact clusters likely due to the high expression and activity of the Wnt/PCP pathway and increased cell-cell adhesion. This together with the enrichment of specified pancreatic progenitors potentially explains the more homogenous pancreatic and endocrine differentiation. The evident increased expression of β cell maturation marker MAFA and GLUT1 in CD177$^+$ ADE-derived β-like cells suggested that cluster architecture and polarisation have an impact on maturation and functionality. Another striking result was the expression of glucose transporter GLUT1 by CD177$^+$ ADE generated β-like cells. In mouse β cells, the presence of GLUT2 during maturation of β cells is necessary for the proper uptake and sensing of glucose to initiate insulin secretion[45]. Upon static glucose stimulation, the insulin secretion of the *in vitro* generated β-like cells using previously published protocols is ~1.5-2 fold[14,15,38,54,55]. The β-like cells generated from enriched CD177$^+$ ADE, however, showed increased insulin content and improved sequential static and dynamic GSIS with a rapid first-phase response. Nevertheless, in the light of recent publications the link between classical maturation markers (MAFA and GLUT1) and functionality of SC-β-cells needs to be revisited[56,57]. In both studies the authors failed to detect MAFA and GLUT1 expression in the SC-β-cells, but both studies showed dynamic GSIS. One study showed rapid first-[56] and the other rapid first- and second-phase insulin secretion[57] after glucose stimulation. Contrary to these findings, ours and other studies show co-expression of INS and MAFA in SC-β-cells[14,23]. A combined approach of enriching CD177-ADE progenitors and improved differentiation protocols[56,57] will further enhance the maturation and functionality of SC-β-cells. This points towards the actuality that hESC and hiPSC genetic background, differentiation protocols, methods and antibodies, as well as functional read-outs are inherently different and need to be harmonized for standardized quality control of SC-β-cells.

Recent publications also point to the fact that enrichment of certain subpopulations of cells (e.g. $CD177^+$ ADE or $GP2^+$ PP)[22,23] at different stages during differentiations can enhance the efficiency of the terminally differentiated phenotypes and increases the purity and safety of the differentiation product. Isolation of $CD177^+$-specified ADE progenitors will help in controlling the differentiation at early stages even before PDX1 can be measured thereby maintaining homogeneity at later stages of differentiation. Testing CD177 induction at DE/ADE stage at S1 can be used as an estimate to determine the differentiation potential of genetically different hiPSC and hESC lines towards the pancreatic lineage. This can be important for upscaling the differentiation process and generate large number of pancreatic progenitors for β-cell replacement therapy. Encouraging results on self-expanding endoderm have shown the capacity to enhance homogenous differentiation to hormone producing cell types by synchronizing the differentiation speed and reducing cellular heterogeneity[28].

Taken together, our results do not only broaden the understanding of human pancreatic development but also provides insights into how the translation of certain developmental cues can contribute to generation of mature β-like cells *in vitro*. $CD177^+$ and $CD275^+$ serve as tools to standardise early differentiations in terms of generation of specified organ progenitors before moving on with expensive and time-consuming long-term differentiation paradigms. In light of these recent findings, our study will help to globally accelerate upscaling the production of clinically relevant cells for disease modelling, drug testing and cell-replacement therapy.

Materials and Methods

Cell sources

Human islets were obtained from the Rudbecklaboratoriet C11 (Uppsala, Sweden) and islet core facility (Edmonton, Canada) with informed consent. The H9, H1 and HUES8 hESC lines was received from WiCell Research Institute, Inc. (Madison, WI). The Mel1-NKX6.1-GFP were obtained from Australian Stem Cell Centre (Clayton, Victoria). An episomal reprogrammed HMGUi-001 iPSC line was generated in lab from control group of MODY-4 patients[58] (Gibco Human Episomal iPSC, Cat#A18945, Life Technologies, CA). All cell lines have been authenticated by Cell Line Genetics (Madison, WI) and confirmed to be mycoplasma-free by using the Lonza MycoAlert Mycoplasma Detection Kit (Lonza, Cat#LT07-418). The hESC lines were used under the permission of Robert Koch Institute.

In vitro differentiation of human pluripotent stem cells towards pancreatic β-like cells.

H9, H1, Mel1-NKX6.1-GFP, HUES8 and HMGUi-001 iPSCs were cultured on 1:30 diluted Geltrex (Invitogen, U.K, Cat#A1413302) in StemMACS iPS-Brew medium (Miltenyi Biotec, Germany, Cat#130-104-368). At ~70% confluency, cultures were rinsed with 1× DPBS without Mg^{2+} and Ca^{2+} (Invitrogen, Cat#14190) followed by incubation with 0.5 mM EDTA (Applichem, Cat#12604-021) for 2–3 min at 37°C. Single cells were rinsed with iPS-Brew, and spun at 1300 r.p.m. for 3 min. The resulting cell pellet was suspended in iPS-Brew medium supplemented with Y-27632 (10 μM; Sigma-Aldrich; MO, Cat#Y0503) and the single cell suspension was seeded at ~1.5–2×10^5 cells/cm^2 on Geltrex-coated surfaces. Cultures were fed every day with iPS-Brew medium and differentiation was initiated 24 h following seeding, resulting in ~90% starting confluency. This confluency is a key factor to getting proper patterning of endoderm.

The cells were differentiated towards definitive endoderm using MCDB131 medium supplemented with 0.5% BSA (Sigma, Cat#10775835001), 100 ng/ml Activin A (Peprotech, Cat#120-14-300) and 25 ng/ml WNT3A (Peprotech, Cat#315-20) or 3 µM CHIR-99021 (Miltenyi Biotec, Cat#130-103-926) for the first day. For the next 2 days, the cells were treated with MCDB131 supplemented with 0.5% BSA and 100ng/ml Activin A.

For differentiation towards β cells, Rezania et al. 2014 β-cell differentiation protocol was used[14]. In nutshell, the cells were differentiated towards primitive gut tube with MCDB131 supplemented with 0.5 % BSA, 50 ng/ml of FGF7 (Peprotech, Cat#100-19-100) 0.25 mM ascorbic acid (Sigma, Cat#120-14-300) and 1.25 µM IWP2 (Tocris, Cat#3533-10) for 2 days. For the WNT signalling activation experiments only, 20 ng/ml of WNT3A or 3 µM CHIR was added to the cultures along with the S2 medium without IWP2. For differentiation towards posterior foregut, the cells were further exposed to MCDB131 medium supplemented with 1xGlutamax (Gibco, Cat#A12860-01), 2% BSA (Cat#10775835001), 0.25 mM ascorbic acid (Sigma, Cat#A4544-25G), 50 ng/ml FGF7, 0.25 µM SANT-1 (Sigma, Cat#S4572-5MG), 1 µM retinoic acid (Sigma, Cat#R2625-50MG), 100 nM LDN193189 (Sigma, Cat#04-0074), 1:200 ITS-X (Gibco, Cat#51500-056) and 200 nM TPB (Merk Millipore, Cat#565740-1MG) for 2 days. The cells were then further differentiated towards pancreatic endoderm using MCDB131 supplemented with 1xGlutamax, 10 mM final glucose concentration, 2% BSA, 0.25 mM ascorbic acid, 2 ng/ml FGF7, 0.25 µM SANT-1, 0.1 µM retinoic acid, 200 nM LDN193189, 1:200 ITS-X and 100 nM TPB for 3 days. For induction of pancreatic endocrine precursors, the cells were next exposed to MCDB131 medium supplemented with 1xGlutamax, 20 mM final glucose concentration, 2% BSA, 0.25 µM SANT-1, 0.05 µM retinoic acid, 100 nM LDN193189, 1:200 ITS-X, 1 µM T3 (Sigma, Cat#T6397-100MG), 10 µM ALK5 Inhibitor II (Enzo life sciences, Cat#ALX-270-445-M005), 10 µM zinc sulphate (Sigma, Cat#SI Z0251-100G) and 10 µg/ml heparin (Sigma, Cat#H3149) for 3 days. Hormone positive cells were

generated by exposing the endocrine progenitors from last step with MCDB131 supplemented with 1xGlutamax, 20 mM final glucose concentration, 2% BSA, 100 nM LDN193189, 1:200 ITS-X, 1µM T3, 10 mM ALK5 Inhibitor II, 10 µM zinc sulphate and 100 nM gamma secretase inhibitor XX (Merck, Cat#565789) for the first 7 days. For maturation of β-like cells, the cells from previous stage were treated with 2% BSA, 1:200 ITS-X, 1 µM T3, 10 µM ALK5 inhibitor II, 10 µM zinc sulphate, 1 mM N-acetyl cysteine (Sigma, Cat#A9165), 10 µM Trolox (EMD, Cat#648471), 2 µM R428 (SelleckChem, Cat# S2841) and 10 mg/ml of heparin for 15 days.

Hepatic lineage commitment protocol

For generation of hepatic progenitors from DE, cells were sorted for CD177, CD275 and CXCR4 at day 4 and were then seeded on Geltrex coated ibidi chambers at the cell density of $2x10^5$ cells per well. Briefly, the cells were washed the next day and culture medium was changed to hepatic commitment medium KO-DMEM (Knockout serum replacement medium) supplemented with 1 mM L-glutamine, 1% nonessential amino acids, 0.1 mM 2-mercaptoethanol, and 1% dimethyl sulfoxide for 3 days. For the maturation, the cells were then cultured in Iscove's modified Dulbecco's medium (IMDM) supplemented with 20 ng/mL Oncostatin M (Peprotech, Cat#300-10T), 0.5 µM dexamethasone (Sigma, Cat#D1756-1G) and 1:200 ITS-X supplement for 8 more days.

Sorting ADE subpopulations

On day 4 of differentiation (DE/ADE), the differentiated samples were collected and stained for surface markers CD177, CXCR4 and CD275. For staining of the surface markers, 10 µl Ab (conjugated to APC) was added per $1x10^6$ cells in 100 µl volume of MCDB1+0.5% BSA. The cells were stained in dark for 15 min on ice. For magnetic labelling of the antibody the stained cells were washed 3x with PBS to remove the antibody and then suspended in 80 µl of MCDB131+0.5% BSA medium with 20 µL of Anti-APC Microbeads per $10x10^6$ of total cells.

The cells were incubated for 15 min at 4°C. The cells were washed with 1-2 ml of PBS and then suspended up to $20x10^6$ cells in 500 µL of MCDB131+0.5% BSA and proceeded with magnetic sorting. For magnetic sorting of the ADE subpopulations, the LS column was placed in magnetic sorter. The column was rinsed with 3 ml of PBS. Cell suspension was applied to the column and the flow through containing unlabelled cells was discarded. The column was then washed 3x with 3 ml of PBS. The antibody positive cells were collected by removing the column from the separator and flushing it with 5 ml of medium. The cells were then seeded in iPS-Brew medium supplemented with 10 µM Y-compound at the seeding density of $2\text{-}10x10^3$ cells in 1 well of ultra-low attachment round bottom 96 well plates to form an aggregate or $4x10^5$ cells in one well of ibidi chamber for further differentiation and staining.

Quantitative qPCR analysis and gene profiling

Gene expression was assessed in differentiated cells by Taqman Arrays (Applied Biosystems). Data were analysed using Expression Suite Software (Applied Biosystems) and normalized to undifferentiated hESCs using ΔΔCt method. Refer to Supplementary Table 3 for primer details.

Immunofluorescence staining and Western blotting

For immunocytochemistry, the cells or aggregates were dissociated using Accutase and fixed in 4% PFA. The cells were then permeabilised using 0.5% Triton-X in blocking buffer for 30 minutes at room temperature. The cells were blocked for 1 h using donkey block and stained for interested markers at 4°C overnight. Secondary antibody staining was performed for 2 h at room temperature. the following day. DAPI staining was used to define nucleus. List of primary antibodies used for immunostaining are mentioned in Supplementary Table 4.

Nuclear fractionation was done from sorted cells using Subcellular protein Fractionation Kit for cultured cells from Thermo Scientific (Cat#78840). For Western blotting, cells were dissociated in RIPA buffer directly after sorting. Cell lysates were resolved by SDS-PAGE,

transferred to PVDF membrane (Biorad) and incubated with the antibodies (Primary antibody: o/n, 4°C; Secondary antibody 1 h, room temperature). Protein bands were visualized using HRP conjugated antibodies and chemiluminescence reagent (Millipore). The bands were quantified with ImageJ. Refer to Supplementary Table 4 for the dilutions and list of antibodies.

Flow cytometry analysis

hESCs, hiPSCs and differentiated cells were dissociated and single cell suspension was prepared. The cells were fixed, permeablilised and stained for DE, pancreatic endoderm, pancreatic progenitors and hormone positive markers. FACS gating was determined using isotype antibodies. Cells were analysed using BD FACS Aria III. The list of all the FACS antibodies and isotype controls are mentioned in Supplementary Table 4.

Sequential static glucose stimulated insulin secretion

Sequential static glucose stimulated insulin secretion (seqGSIS) of the generated β-like cells was performed based on previous described protocols[14,59]. Briefly, 5 aggregates (6,000-10,000 cells in total) (n=4-5, biological replicates) were picked and rinsed three times with KRBH buffer (129 mM NaCl, 4.8 mM KCl, 2.5 mM $CaCl_2$, 1.2 mM $MgSO_2$, 1 mM Na_2HPO_4, 1.2 mM KH_2PO_4, 5 mM $NaHCO_3$, 10 mM HEPES and 0.1% BSA in deionized water and sterile filtered) and then equilibrated in KRBH buffer at 37°C for 30 min. Aggregates then were incubated in KRBH buffer spiked with 2.8 mM glucose for 30 min at room temperature. Supernatants were collected and the aggregates were transferred to KRBH buffer spiked with 20 mM glucose for 30 min. Supernatants were collected again. The aggregates were then washed to remove left over high glucose with KRBH and another round of low glucose and high glucose stimulus was performed. At the end of the experiment, cell aggregates were dissociated into single cells and the cell numbers were counted to normalize the GSIS. Mercodia Human Insulin ELISA kit

(Mercodia, Cat#10-1113-01) was used to measure the insulin content in supernatant sample following manufacturer's protocol.

Dynamic glucose stimulated insulin secretion

Briefly, 25 SC-derived β-cells (iPSC (n=3), H1 hESC (n=5) biological replicates) or human islets from healthy donors (n=4 biological replicates) were pre-incubated in KRBH buffer (115 mM NaCl, 5 mM KCl, 2.5 mM $CaCl_2$, 1 mM $MgCl_2$, 24 mM $NaHCO_3$,10 mM HEPES and 0.1% BSA, pH 7.4) containing 2.8 mM Glucose for 30 min and then loaded on a nylon filter in a plastic perfusion chamber containing acrylamide-based microbead slurry (Bio-Gel P-4, Bio-Rad Laboratories). The SC-derived β-cells or human islets were then sequentially perifused with low glucose (2.8 mM) for 12 min, followed by high-glucose (20 mM) for 24 min, Exendin-4 (10 nM) + high glucose (20 mM) for 24 min, low glucose (2.8 mM) for 12 min and a final step with 25 mM KCl for 12 min at a constant flow rate of 100 µl/180 sec using the BioRep perifusion system (Model No. PERI-4.2) maintained at 37°C in a temperature controlled chamber. Flow through fractions were collected on a 96-well plate maintained at 4°C and quantified for insulin content using Human Insulin ELISA (Mercodia, Cat#10-1113-01) as per manufacturer's instruction. After the completion of run, SC-derived β-cells or human islets were recovered from perifusion chambers and assayed for DNA contents and quantified using Quant-IT PicoGreen dsDNA kit, (Thermo Fischer, Cat# P7581).

591 **Table 1: Demographic information and characteristics of healthy human islet donors**

Donor ID	**Age (Years)**	**Gender**	**BMI**	**HLA**	**Islet Purity (%)**	**HbA1c (%)**
R301	18	M	19	A:2, - B:65, 49 BW: 6, 4 CW:7, 8 DRB1: DR:11, 13 DQB1:7, 6 DQA1:5, 1 DPA1: 1, -DPB1:02;01, 04;01 Other: DRW - 52, 52	75	5
R305	60	M	21	A:11, 31 B:8, 44 BW:6, 4 CW:5, 7 DRB1:17, 4 DR: DQB1:2, 7 DQA1: 5, 3 DPA1: 1, 2 DPB1: 1, 04;01 Other: DR52present DR53 present	80	5.6
R306	22	M	21.1	A:1, 3 B:8, 65 BW:6, 6 CW: 7, 8 DRB1:15, 15 DR:51, -- DQB1:6, 6 DQA1:1, 1 DPA1: 1, 1 DPB1: 1, 01;02		
R309	47	F	27.4	A:2 ,3 B:7, 60 BW:6, 6 CW:10, 7 DRB1: DR:5, 15 DQB1:8, 6 DQA1:1, 3 DPA1:1, 1		

				DPB1: 04;01, 04;01 Other: DRW - 53, 51		
R310	25	M	26.4	A:02, 31 B:7, 60 BW: CW:10, 7 DRB1:13, 15 DR: DQB1:6, 6 DQA1:1, 1 DPA1: 1, 2 DPB1:0401, 5 Other: DRB3 - 52, --		

Human islets were obtained from the University of Alberta Diabetes Institute Islet Core (Edmonton, Alberta, Canada). Islets from healthy donors R305, R306, R309, R310 were used for dynamic GSIS, whereas R301, R305, R306, R309, R310 were used for insulin content.

Insulin content

S7 clusters from CD177, unsorted differentiations and human islets were washed with PBS and dissociated using Accutase. Cells were counted and 1000 cells were collected for measuring insulin content. The cells were resuspended in Acid-EtOH solution (1.5% HCL and 70% EtOH) and kept on a shaker at 4°C overnight. The tubes were centrifuged at 2100 g for 15 min and supernatant was collected and neutralised with an equal volume of 1 M Tris (pH 7.5). Human insulin was measured using Mercodia Human Insulin ELISA kit (Mercodia, Cat#10-1113-01).

DTZ staining

S7 clusters were washed with PBS and then suspended into DTZ staining for 2 min at RT. The cells were then washed with PBS and pictures were taken with Leica Stereoscope.

Affymetrix microarray analysis

606 For gene profiling, total RNA was extracted using miRNeasy Mini kit (Qiagen, Cat#217004),
607 RNA integrity was checked using Agilent 2100 Bioanalyzer (Agilent RNA 6000 Pico Kit) and
608 only high quality RNA (RIN>7) was used for microarray analysis. Total RNA (5 ng) was
609 amplified using the Ovation Pico WTA System V2 in combination with the Encore Biotin
610 Module (Nugen). Amplified cDNA was hybridised on Affymetrix Human Gene ST 2.0 arrays
611 (Affymetrix, 902113). Staining and scanning (GeneChip Scanner 3000 7G) was done according
612 to the Affymetrix expression protocol including minor modifications as suggested in the Encore
613 Biotin protocol. Expression console (v.1.4.1.46, Affymetrix) was used for quality control. All
614 subsequent computational analysis was performed in R using Bioconductor packages.
615 Expression data were RMA normalised using the oligo package (version 1.42.0) and probe sets
616 were annotated using the package hugene20sttranscriptcluster.db (version 8.7.0). Differential
617 expression analyses were performed using the limma package (version 3.34.9) and *P*-values
618 were adjusted for multiple testing by Benjamini-Hochberg correction. A gene was considered
619 as differentially expressed if the raw *P*-value was below a threshold of 0.01 and the fold-change
620 was greater than or equal to 1.5. Functional enrichments were conducted using HOMER[60]. All
621 microarray data is available at GEO under the accession number GSE113791
622 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE113791).

623 **Image analysis**

624 Images were acquired with Leica SP5 confocal microscope and Zeiss LSM 880 Airy Scan
625 confocal microscope. Images taken by Leica confocal were analysed using Leica LAS AF Lite.
626 Images taken by Zeiss confocal microscope were analysed using Zeiss Zen Blue software.

627 **Statistical analysis and reproducibility**

628 All values are depicted as means ±s.e.m. All statistical tests performed are mentioned in figure
629 legends for each data set. In brief, statistical significance is defined as $P<0.05$. Sample sizes are

provided in the figure legends. Comparison of 3 or more data sets were performed using ordinary one way ANOVA with Bonferroni's multiple comparison test for Figs. 1d-g, 2f-g, 3b/e, 4b/e/i, 6b, 7b and Supplementary Figs. 1g-h, 3e and 4c. Two tailed unpaired t-tests (Student's *t*-test) with Welch's correction was used in Fig. 4j, 5d/e, 6f, 7d and Supplementary Fig. 6g/h, 8a. Two tailed unpaired t-tests (Student's *t*-test) assuming equal standard deviation in Supplementary Fig. 8b. All statistics were performed using GraphPad Prism software 8 (GraphPad Software Inc., La Jolla, CA). No statistical methods were used to determine sample size. All *P* values are displayed in the figures.

Acknowledgements

We would like to thank Alireza Rezania for providing hMAFA and hGLUT1 antibodies. The authors are grateful to Ingo Burtscher, Aimée Bastidas-Ponce, Ciro Salinno and Marta Medina Tarquis for their invaluable inputs and technical assistance. We would like to thank Donna Marie Thomson and Mara Catani for reading and correction of the manuscript. This work was supported by the HumEN consortium funded by European Union's Seventh Framework Programme for Research, Technological Development and Demonstration under grant agreement no.602587 (http://www.hum-en.eu/). This work was also funded in part by the German Center for Diabetes Research (DZD e.V.) and the Helmholtz Alliance 'Aging and Metabolic Programming, AMPro'. P.U.M is funded by postdoctoral grant from HumEN consortium, EU. M.S, S.P and K.S are supported by Helena graduate school and Technische Universität München (TUM). S.K is an employee of Miltenyi Biotec GmbH, Bergisch Gladbach , Germany.

Author Contributions

P.U.M and H.L wrote the manuscript. P.U.M and K.S contributed equally to this work. P.U.M and H.L designed experiments, analysed and interpreted the results. K.S, S.P and A.A. performed experiments and analysed data. M.S provided bioinformatics assistance and analysed the results. J.B provided technical assistance. S.K designed, performed and analysed the antibody screen. M.I performed the microarray chips and analysed them. P.U.M and H.L conceived the project.

Competing financial interests

The authors declare competing financial interests. Helmholtz Zentrum München and Miltenyi Biotec own the patent (#PLA16A13, *Luxembourg patent application No.: 100320:* “"METHODS FOR PURIFYING ENDODERM AND PANCREATIC ENDODERM CELLS DERIVED FROM HUMAN EMBRYONIC STEM CELLS" for the use of CD177 and CD275 antibodies in pancreatic and liver differentiation protocols.

References and Bibliography

1 Zorn, A. M. & Wells, J. M. Vertebrate Endoderm Development and Organ Formation. *Annual review of cell and developmental biology* **25**, 221-251, doi:10.1146/annurev.cellbio.042308.113344 (2009).

2 Spence, J. R. & Wells, J. M. Translational embryology: Using embryonic principles to generate pancreatic endocrine cells from embryonic stem cells. *Developmental Dynamics* **236**, 3218-3227, doi:10.1002/dvdy.21366 (2007).

3 Beddington, R. S. P. & Robertson, E. J. Axis Development and Early Asymmetry in Mammals. *Cell* **96**, 195-209, doi:10.1016/S0092-8674(00)80560-7.

4 Stainier, D. Y. R. A glimpse into the molecular entrails of endoderm formation. *Genes & Development* **16**, 893-907, doi:10.1101/gad.974902 (2002).

5 Tremblay, K. D. & Zaret, K. S. Distinct populations of endoderm cells converge to generate the embryonic liver bud and ventral foregut tissues. *Dev Biol* **280**, 87-99, doi:10.1016/j.ydbio.2005.01.003 (2005).

6 Tam, P. P. L. *et al.* Sequential allocation and global pattern of movement of the definitive endoderm in the mouse embryo during gastrulation. *Development* **134**, 251-260, doi:10.1242/dev.02724 (2007).

7 Zaret, K. S. Genetic programming of liver and pancreas progenitors: lessons for stem-cell differentiation. *Nat Rev Genet* **9**, 329-340 (2008).

8 Serls, A. E., Doherty, S., Parvatiyar, P., Wells, J. M. & Deutsch, G. H. Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. *Development* **132**, 35-47, doi:10.1242/dev.01570 (2005).

9 Wang, Z., Dollé, P., Cardoso, W. V. & Niederreither, K. Retinoic acid regulates morphogenesis and patterning of posterior foregut derivatives. *Dev Biol* **297**, 433-445, doi:https://doi.org/10.1016/j.ydbio.2006.05.019 (2006).

10 Rodríguez-Seguel, E. *et al.* Mutually exclusive signaling signatures define the hepatic and pancreatic progenitor cell lineage divergence. *Genes & Development* **27**, 1932-1946, doi:10.1101/gad.220244.113 (2013).

11 Jennings, R. E. *et al.* Development of the Human Pancreas From Foregut to Endocrine Commitment. *Diabetes* **62**, 3514-3522, doi:10.2337/db12-1479 (2013).

12 Madsen, O. D. & Serup, P. Towards cell therapy for diabetes. *Nature Biotechnology* **24**, 1481, doi:10.1038/nbt1206-1481 (2006).

13 D'Amour, K. A., Agulnick, A. D., Eliazer, S., Kelly, O. G. & Kroon, E. Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat Biotechnol* **23**, doi:10.1038/nbt1163 (2005).

14 Rezania, A. *et al.* Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat Biotech* **32**, 1121-1133, doi:10.1038/nbt.3033 http://www.nature.com/nbt/journal/v32/n11/abs/nbt.3033.html#supplementary-information (2014).

15 Pagliuca, Felicia W. *et al.* Generation of Functional Human Pancreatic β Cells In Vitro. *Cell* **159**, 428-439, doi:https://doi.org/10.1016/j.cell.2014.09.040 (2014).

16 Russ, H. A. *et al.* Controlled induction of human pancreatic progenitors produces functional beta-like cells in vitro. *The EMBO Journal* **34**, 1759-1772, doi:10.15252/embj.201591058 (2015).

17 Canham, M. A., Sharov, A. A., Ko, M. S. H. & Brickman, J. M. Functional Heterogeneity of Embryonic Stem Cells Revealed through Translational Amplification of an Early Endodermal Transcript. *PLOS Biology* **8**, e1000379, doi:10.1371/journal.pbio.1000379 (2010).

18 Morrison, G. M. *et al.* Anterior Definitive Endoderm from ESCs Reveals a Role for FGF Signaling. *Cell Stem Cell* **3**, 402-415, doi:https://doi.org/10.1016/j.stem.2008.07.021 (2008).

19 Gadue, P. *et al.* Generation of Monoclonal Antibodies Specific for Cell Surface Molecules Expressed on Early Mouse Endoderm. *STEM CELLS* **27**, 2103-2113, doi:doi:10.1002/stem.147 (2009).
20 Holtzinger, A. *et al.* New markers for tracking endoderm induction and hepatocyte differentiation from human pluripotent stem cells. *Development (Cambridge, England)* **142**, 4253-4265, doi:10.1242/dev.121020 (2015).
21 Kelly, O. G. *et al.* Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells. *Nat Biotech* **29**, 750-756, doi:http://www.nature.com/nbt/journal/v29/n8/abs/nbt.1931.html#supplementary-information (2011).
22 Ameri, J. *et al.* Efficient Generation of Glucose-Responsive Beta Cells from Isolated GP2⁺ Human Pancreatic Progenitors. *Cell Reports* **19**, 36-49, doi:10.1016/j.celrep.2017.03.032.
23 Cogger, K. F. *et al.* Glycoprotein 2 is a specific cell surface marker of human pancreatic progenitors. *Nature Communications* **8**, 331, doi:10.1038/s41467-017-00561-0 (2017).
24 Deglincerti, A. *et al.* Self-organization of the in vitro attached human embryo. *Nature* **533**, 251, doi:10.1038/nature17948 https://www.nature.com/articles/nature17948#supplementary-information (2016).
25 Burtscher, I. & Lickert, H. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* **136**, 1029-1038, doi:10.1242/dev.028415 (2009).
26 Hoshino, H., Shioi, G. & Aizawa, S. AVE protein expression and visceral endoderm cell behavior during anterior–posterior axis formation in mouse embryos: Asymmetry in OTX2 and DKK1 expression. *Dev Biol* **402**, 175-191, doi:https://doi.org/10.1016/j.ydbio.2015.03.023 (2015).
27 Franklin, V. *et al.* Regionalisation of the endoderm progenitors and morphogenesis of the gut portals of the mouse embryo. *Mechanisms of Development* **125**, 587-600, doi:https://doi.org/10.1016/j.mod.2008.04.001 (2008).
28 Cheng, X. *et al.* Self-Renewing Endodermal Progenitor Lines Generated from Human Pluripotent Stem Cells. *Cell stem cell* **10**, 371-384, doi:10.1016/j.stem.2012.02.024 (2012).
29 Stroncek, D. F. Neutrophil-specific antigen HNA-2a, NB1 glycoprotein, and CD177. *Current opinion in hematology* **14**, 688-693, doi:10.1097/MOH.0b013e3282efed9e (2007).
30 Hu, H. *et al.* Noncanonical NF-κB regulates inducible costimulator (ICOS) ligand expression and T follicular helper cell development. *Proceedings of the National Academy of Sciences* **108**, 12827-12832, doi:10.1073/pnas.1105774108 (2011).
31 Rezania, A. *et al.* Enrichment of human embryonic stem cell-derived NKX6.1-expressing pancreatic progenitor cells accelerates the maturation of insulin-secreting cells in vivo. *STEM CELLS* **31**, 2432-2442, doi:doi:10.1002/stem.1489 (2013).
32 Xu, X., Browning, V. & Odorico, J. Activin, BMP and FGF pathways cooperate to promote endoderm and pancreatic lineage cell differentiation from human embryonic stem cells. *Mechanisms of development* **128**, 412-427, doi:10.1016/j.mod.2011.08.001 (2011).
33 Yap, C., Goh, H. N., Familari, M., Rathjen, P. D. & Rathjen, J. The formation of proximal and distal definitive endoderm populations in culture requires p38 MAPK activity. *Journal of Cell Science* **127**, 2204-2216, doi:10.1242/jcs.134502 (2014).
34 Piccolo, S. *et al.* The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature* **397**, 707-710, doi:10.1038/17820 (1999).
35 MacDonald, B. T., Tamai, K. & He, X. Wnt/β-catenin signaling: components, mechanisms, and diseases. *Developmental cell* **17**, 9-26, doi:10.1016/j.devcel.2009.06.016 (2009).
36 Chen, Y.-F. *et al.* Rapid Generation of Mature Hepatocyte-Like Cells from Human Induced Pluripotent Stem Cells by an Efficient Three-Step Protocol. *Hepatology (Baltimore, Md.)* **55**, 1193-1203, doi:10.1002/hep.24790 (2012).
37 Decker, K., Goldman, D. C., Grasch, C. L. & Sussel, L. Gata6 is an important regulator of mouse pancreas development. *Dev Biol* **298**, 415-429, doi:10.1016/j.ydbio.2006.06.046 (2006).

38 Shi, Z.-D. *et al.* Genome Editing in hPSCs Reveals GATA6 Haploinsufficiency and a Genetic Interaction with GATA4 in Human Pancreatic Development. *Cell Stem Cell* **20**, 675-688.e676, doi:https://doi.org/10.1016/j.stem.2017.01.001 (2017).
39 Tiyaboonchai, A. *et al.* GATA6 Plays an Important Role in the Induction of Human Definitive Endoderm, Development of the Pancreas, and Functionality of Pancreatic β Cells. *Stem Cell Reports* **8**, 589-604, doi:10.1016/j.stemcr.2016.12.026 (2017).
40 Chen, B. *et al.* Small molecule-mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nature chemical biology* **5**, 100-107, doi:10.1038/nchembio.137 (2009).
41 Pei, Y. *et al.* WNT signaling increases proliferation and impairs differentiation of stem cells in the developing cerebellum. *Development* **139**, 1724-1733, doi:10.1242/dev.050104 (2012).
42 Bader, E. *et al.* Identification of proliferative and mature β-cells in the islets of Langerhans. *Nature* **535**, 430-434, doi:10.1038/nature18624 http://www.nature.com/nature/journal/v535/n7612/abs/nature18624.html#supplementary-information (2016).
43 Johansson, K. A. *et al.* Temporal Control of Neurogenin3 Activity in Pancreas Progenitors Reveals Competence Windows for the Generation of Different Endocrine Cell Types. *Developmental Cell* **12**, 457-465, doi:https://doi.org/10.1016/j.devcel.2007.02.010 (2007).
44 Krentz, N. A. J. *et al.* Phosphorylation of NEUROG3 Links Endocrine Differentiation to the Cell Cycle in Pancreatic Progenitors. *Dev Cell* **41**, 129-142 e126, doi:10.1016/j.devcel.2017.02.006 (2017).
45 Roscioni, S. S., Migliorini, A., Gegg, M. & Lickert, H. Impact of islet architecture on [beta]-cell heterogeneity, plasticity and function. *Nat Rev Endocrinol* **12**, 695-709, doi:10.1038/nrendo.2016.147 (2016).
46 Vincent, S. D., Dunn, N. R., Hayashi, S., Norris, D. P. & Robertson, E. J. Cell fate decisions within the mouse organizer are governed by graded Nodal signals. *Genes & Development* **17**, 1646-1662, doi:10.1101/gad.1100503 (2003).
47 Perea-Gomez, A. *et al.* Nodal Antagonists in the Anterior Visceral Endoderm Prevent the Formation of Multiple Primitive Streaks. *Developmental Cell* **3**, 745-756, doi:https://doi.org/10.1016/S1534-5807(02)00321-0 (2002).
48 Robertson, E. J., Norris, D. P., Brennan, J. & Bikoff, E. K. Control of early anterior-posterior patterning in the mouse embryo by TGF-β signalling. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **358**, 1351-1358, doi:10.1098/rstb.2003.1332 (2003).
49 Brennan, J. *et al.* Nodal signalling in the epiblast patterns the early mouse embryo. *Nature* **411**, 965doi:10.1038/35082103https://www.nature.com/articles/35082103#supplementary-information (2001).
50 Arnold, S. J. & Robertson, E. J. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nature Reviews Molecular Cell Biology* **10**, 91, doi:10.1038/nrm2618 (2009).
51 Aykul, S., Ni, W., Mutatu, W. & Martinez-Hackert, E. Human Cerberus Prevents Nodal-Receptor Binding, Inhibits Nodal Signaling, and Suppresses Nodal-Mediated Phenotypes. *PLOS ONE* **10**, e0114954, doi:10.1371/journal.pone.0114954 (2015).
52 Cortijo, C., Gouzi, M., Tissir, F. & Grapin-Botton, A. Planar Cell Polarity Controls Pancreatic Beta Cell Differentiation and Glucose Homeostasis. *Cell Reports* **2**, 1593-1606, doi:10.1016/j.celrep.2012.10.016.
53 Pezzulo, A. A. *et al.* The air-liquid interface and use of primary cell cultures are important to recapitulate the transcriptional profile of in vivo airway epithelia. *American Journal of Physiology - Lung Cellular and Molecular Physiology* **300**, L25-L31, doi:10.1152/ajplung.00256.2010 (2011).

54 Kim, Y. *et al.* Islet-like organoids derived from human pluripotent stem cells efficiently function in the glucose responsiveness in vitro and in vivo. **6**, 35145, doi:10.1038/srep35145 https://www.nature.com/articles/srep35145#supplementary-information (2016).
55 Millman, J. R. *et al.* Generation of stem cell-derived β-cells from patients with type 1 diabetes. **7**, 11463, doi:10.1038/ncomms11463 https://www.nature.com/articles/ncomms11463#supplementary-information (2016).
56 Nair, G. G. *et al.* Recapitulating endocrine cell clustering in culture promotes maturation of human stem-cell-derived β cells. *Nature Cell Biology* **21**, 263-274, doi:10.1038/s41556-018-0271-4 (2019).
57 Velazco-Cruz, L. *et al.* Acquisition of Dynamic Function in Human Stem Cell-Derived β Cells. *Stem Cell Reports* **12**, 351-365, doi:https://doi.org/10.1016/j.stemcr.2018.12.012 (2019).
58 Wang, X. *et al.* Generation of a human induced pluripotent stem cell (iPSC) line from a patient with family history of diabetes carrying a C18R mutation in the PDX1 gene. *Stem Cell Research* **17**, 292-295, doi:https://doi.org/10.1016/j.scr.2016.08.005 (2016).
59 Shi, Z.-D. *et al.* Genome Editing in hPSCs Reveals GATA6 Haploinsufficiency and a Genetic Interaction with GATA4 in Human Pancreatic Development. *Cell Stem Cell* **20**, 675-688.e676, doi:10.1016/j.stem.2017.01.001.
60 Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell* **38**, 576-589, doi:https://doi.org/10.1016/j.molcel.2010.05.004 (2010).

Main Figure Legends:

Figure 1. **Identification of novel CD177$^+$ and CD275$^+$ ADE subpopulations** (a) Schematic representation of hESCs differentiation towards DE showing the growth factors and small molecules added. (b-c) Representative FACS plot of CXCR4$^+$/CD117$^+$ cells (b) showing a heterogeneous population and apparent homogenous FOXA2$^+$/SOX17$^+$ (c) DE. (d-e) Gene expression profiles of CXCR4$^+$/CD117$^-$, CXCR4^{high}/CD117^{high}, CXCR4^{mid}/CD117^{mid} and CXCR4^{low}/CD117^{low} cells for *FOXA2* and *SOX17*. (f-g) Gene expression profiles of CXCR4$^+$/CD117$^-$, CXCR4^{high}/CD117^{high}, CXCR4^{mid}/CD117^{mid} and CXCR4^{low}/CD117^{low} cells for *CER1* and *HHEX* (Fig. 1b-g n=3, biological replicates). Mean ±s.e.m, P<0.05 and P<0.01 is determined using ordinary one way ANOVA with Bonferroni's test for multiple analysis. (h) Summary of the antibody screen identifying and isolating CD177$^+$ and CD275$^+$ as novel markers to identify ADE subpopulations. CXCR4 and FOXA2 are used as controls to identify the whole DE. (i) hPSCs and hPSC-derived DE stained for CXCR4, CD177 and CD275 as shown by live cell FACS.

Figure 2. Molecular profiling of CD177$^+$, CD275$^+$ and CXCR4$^+$ DE subpopulations reveal distinct signatures (a) Summary of differentiation protocol towards DE/ADE followed by MACS sorting to enrich for CD177 and CD275 subpopulation. CXCR4 represents the non-enriched DE/ADE population. (b) PCA analysis showing mRNA derived transcriptome profiles are characteristic of different DE/ADE subpopulations. (c-e) Bar graphs of selected and significantly enriched gene ontology (GO) terms in CD275$^+$ versus CXCR4$^+$ (c), CD177$^+$ versus CD275$^+$ (d) and CD177$^+$ versus CXCR4$^+$ (e) DE populations. (f-g) Validation of the microarray analysis by real-time qPCR for non-canonical WNT/PCP components and ligands (f) and canonical WNT components and ligands (g) (n=3, biological replicates). Error bar represents ±s.e.m; P<0.05 and P<0.01 is determined using ordinary one way ANOVA with Bonferroni's test for multiple analysis. Data were normalized to *18S*. (h, i) Western blot analysis (h) and quantification (i) of WNT/PCP components such as p-JNK and DVL2 in ADE subpopulations. GAPDH is used as a loading control. (j-k) Nuclear fractionation (j) and quantification (k) of β-CATENIN in cytoplasm (C) and nucleus (N) of enriched ADE subpopulations. (l) Immunofluorescence analysis validated the exclusive localization of β-catenin in the membrane in CD177$^+$ ADE cells and in the cytoplasm and nucleus in CD275$^+$ ADE and CXCR4$^+$ DE cells. FOXA2 is used as a nuclear marker. Scale bars 50 μm and 20 μm in inset.

Figure 3. CD177$^+$ ADE efficiently differentiates into PDX1$^+$/NKX6.1$^+$ pancreatic progenitors (a) Summary of pancreatic differentiation protocol after the enrichment of

CD177$^+$, CD275$^+$ and CXCR4$^+$ at DE stage. (b) mRNA quantification of *PDX1* and *NKX6.1* in S4 cells generated from enriched subpopulations. Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using ordinary one-way ANOVA with Bonferroni's test for multiple analysis. Data were normalized to CXCR4-PP2. *GAPDH* was used as a control. (c) Immunofluorescence staining of CD177$^+$-, CD275$^+$- and CXCR4$^+$-PP2s for PDX1 and NKX6.1 showing most of the CD177-PP2 positive for PDX1/NKX6.1. Scale bars 50 µm. (d) PDX1/NKX6.1 intracellular FACS analysis and quantification (e) of CD177$^+$-, CD275$^+$- and CXCR4$^+$-PP2 showing percentage of PDX1$^+$/NKX6.1$^+$ cells. Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using ordinary one-way ANOVA with Bonferroni's test for multiple analysis. Exact *P* values are mentioned in the figure.

Figure 4. Inhibition of canonical WNT secretion promotes pancreatic differentiation (a) Overview of differentiation protocol towards pancreatic progenitors (PP2) where inhibition of WNT secretion is performed at S2 stage in CD177$^+$-, CD275$^+$- and CXCR4$^+$-DE using IWP2. (b-c) FACS quantification (b) and immunocytochemistry (c) for the percentage of cells expressing PDX1$^+$/NKX6.1$^+$ generated from CD177$^+$-, CD275$^+$- and CXCR4$^+$-ADE. Scale bars, 20 µm. (d-g) Protocol for activation of canonical and non-canonical WNT pathway using 20 ng/ml WNT3A or 3 µM CHIR and 400 ng/ml of WNT5A at S2 stage. (e-g) Analysis (e) of representative FACS plots (f) and immunofluorescent images (g) for PDX1/NKX6.1 at S4 stage showing activation of canonical WNT pathway hampering PP2 induction. Data in (b,e,i) is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using ordinary one-way ANOVA with Bonferroni's test for multiple analysis. Exact *P* values are mentioned in the figure. Scale bars, 50µm. (h) EdU staining (h) and quantification (i) for EdU$^+$ cells in different ADE subsets revealing the exit of cells from the cell cycle. Scale bars, 80 µm. (j) mRNA analysis of *NGN3* at S4 and S5 stage in different ADE subpopulations. (n=3, biological replicates). Error bar represents ±s.e.m; $P<0.05$, $P<0.01$, two-sided unpaired *t*-test. Data were normalized to *18S.*

Figure 5. CD177$^+$- ADE efficiently differentiates into NKX6.1$^+$/INS$^+$ β-like cells (a) Schematic representation of adapted Rezania et al. 2014 protocol for differentiation towards pancreatic β cells. (b, f) Representative immunofluorescent staining for CD177-β-cells, CXCR4-β-cells and unsorted (US-β-cells) for: (b) NKX6.1 (red) and C-peptide (green) and (f) GCG (red), INS (green) and SST (magenta). Nuclear DAPI staining is shown in blue. Scale bars, 20 µm (b) and 50 µm (f). (c, g) Representative FACS plots for CD177-β-cells, CXCR4-β-cells and US-β-cells for expression of INS and NKX6.1 (c) and INS, GCG and SST (g). (d-e) FACS quantification of the cells at S7 stage for the percentage of cells expressing INS,

NKX6.1 and GCG (d) and INS and SST (e). 4 biological independent experiments were used. Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using ordinary one-way ANOVA with Bonferroni's test for multiple analysis (d) and two-sided unpaired *t*-test (e). Exact *P* values are mentioned in the figure.

Figure 6. **CD177^{+}- enriched ADE generates more mature β-like cells *in vitro*** (a) Outline of the differentiation protocol. (b) Gene expression profile of β-like cells for maturation markers *MAFA* and *GLUT1*. Data are expressed as the fold change relative to US-β-cells. Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using ordinary one-way ANOVA with Bonferroni's test for multiple analysis. Exact *P* values are mentioned in the figure. Data were normalized to *18S*. (c, e) Immunohistochemistry for the expression of MAFA (red) and INS (green) and (e) INS (green) and GLUT1 (red) in US-β-cells and CD177-β-cells. DAPI is shown in blue. Insets show higher magnification images. For MAFA, staining was performed on monolayer due to antibody issues in PFA fixed sections. GLUT1 staining was performed on sectioned aggregates. Scale bars, in low magnification images, 50 µm (c) and high magnification images 5 µm (c) and 20 µm (e) and inset 80 µm (e). (d, f) FACS analysis (d) and quantification (f) of cells expressing MAFA and GLUT1 in differentiation at S7. Each point is a biological independent replicate. Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using non-parametric Student's *t* test. Exact *P* values are mentioned in the figure.

Figure 7. **Functional characterization of the US-β-cells and CD177-β-cells** (a) Images of US-β-cells and CD177-β-cells stained with DTZ (b) Comparison of insulin content in 1000 cells between US-β-cells (n=9), CD177-β-cells (n=9) and human islets (n=5). (c) Sequential static GSIS assay comparing the insulin secretion of US-β-cells and CD177-β-cells when subjected to 2.8, 5.6, 11.1, 20 and 2.8 mM glucose stimulations in the interval of 30 minutes (n=4, biological replicates) (d) Insulin secretion in static GSIS on 5 aggregates collected from US-β-cells (n=4; biologically independent experiments) and CD177-β-cells (n=5; biologically independent experiments) in response to multiple challenges with 30 min of 2.8 and 20 mM glucose. (e-g) Insulin secretion in response to dynamic glucose, Ex-4 and KCl challenges in a perifusion system on US-β-cells (f) and CD177-β-cells (n=3, biological replicates) (g) cells in comparison to human islets (n=4, biological replicates) (e). Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using ordinary one-way ANOVA with Bonferroni's test for multiple analysis (b) and non-parametric Student's *t* test (d).

Supplementary Figure legends:

Supplementary Fig. 1. Screening strategy for the identification of endoderm subpopulations (a) Screening work flow for the initial screen. (b-d) Representative FACS plots for CD177 and CD275 (b) labelling of differentiated day 4 DE cells with known endoderm markers (FOXA2 and CXCR4) revealed definitive endoderm ($FOXA2^+/CXCR4^+$) and mes-endoderm ($FOXA2^{low}/CXCR4^-$) subpopulations (c-d) CD177 and CD275 expression profiles reveal different endoderm subpopulations. (e) Immunofluorescent staining for CER1 (green) with FOXA2 (red) in DE cultures. Scale bar, 50 µm. (f) FACS analysis for $CD275^+/CER1^+$ and $CD177^+/CER1^+$ ADE cell populations in day 4 DE. (g-h) qPCR quantification for the mRNA expression of *FOXA2* and *SOX17* (g), *CER1* and *HHEX* (h) in enriched $CD177^+$ and $CD275^+$ ADE subpopulations. Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using ordinary one-way ANOVA with Bonferroni's test for multiple analysis.

Supplementary Fig. 2. Percentage of $CD177^+$ and $CD275^+$ ADE subpopulations induced in different hESC and hiPSC lines (a-b) Endoderm differentiation scheme from hESCs towards DE/ADE. (b) FACS plots represent the percentage of $CXCR4^+/CD177^+$ and $CXCR4^+/CD275^+$ subpopulations in hH1, hMEL1-NKX6.1, HMGUi001-A hiPSC, HUES8 and H9 at DE/ADE stage. (c) Quantification of flow cytometry data from (b). Each point represents a biologically independent experiment.

Supplementary Fig. 3. Induction efficiency of $CD177^+$ and $CD275^+$ ADE shows variation using different protocols (a-c) Adaptation of previously published endoderm differentiation protocols from hESCs. (d) FACS quantification for the percentage of total population expressing CXCR4 in DE cells derived from HMGUi001-A hiPSC using 3 different endoderm induction protocols (n=3 biological replicates). (e) FACS quantification for the percentage of cells expressing $CXCR4^+/CD177^+$ and $CXCR4^+/CD275^+$ in DE generated using previously published protocols.

Supplementary Fig. 4. Differentiation of enriched $CD177^+$, $CD275^+$ and $CXCR4^+$ ADE subpopulations towards liver and pancreas fate (a) Expression of $CD177^+$-, $CD275^+$-, and $CXCR4^+$ during differentiation of hESCs towards pancreatic β-like cells. (b) Liver differentiation protocol. (c) qPCR quantification of the expression of early liver progenitor markers *HHEX, TTR* and *AFP* in enriched ADE subpopulations. Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using ordinary one-way ANOVA with Bonferroni's test for multiple analysis. (d) Immunofluorescent staining of pancreatic progenitor cells derived from

enriched ADE subpopulations for the co-expression of posterior foregut marker GATA6 and PDX1. Scale bars, 50µm**.** e) Immunofluorescent staining of pancreatic progenitor cells derived from enriched ADE subpopulations for the co-expression of lung marker SOX2, intestinal marker CDX2 and PDX1. Scale bars, 50µm**.**

Supplementary Fig. 5. CD177$^+$ ADE positively correlates with PP1 induction (a) Pancreatic induction protocol. (b) FACS analysis of H1, HMGUi001-A hiPSC, HUES8 and MEL1-NKX6.1 for PDX1 at S3 stage. (c) Quantification of CD177$^+$ cells generated at S1 and PDX1$^+$ cells generated at S3 stage showing correlation between CD177 and PDX1 induction. Each point on the graph depicts a biologically independent data set.

Supplementary Fig. 6. H1 hESC pancreatic and endocrine differentiations of CD177$^+$ and US-DE (a) Overview of differentiation protocol used to generate CD177/US-β-cells. (b-e) Immunostainings for INS and NKX6.1 (b), GCG, INS and SST (c) and C-peptide and GLUT1 (d), INS, MAFA and NKX6.1 (e) in CD177- and US-β-cells. Scale bars, 50 µm. (f) Representative flow cytometry contour plots of S4 and S7 cells generated from CD177- and US-ADE/DE cells on H1 line and stained for indicated markers. (g,h) Percentage of cells expressing indicated markers. Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using non-parametric Student's *t* test. Exact *P* values are mentioned in the figure.

Supplementary Fig. 7. Comparison of 2D and 3D culture system on pancreatic differentiation (a) Overview of differentiation protocol used. (b) Comparison of PDX1$^+$/NKX6.1$^+$ generated from CD177$^+$- and CXCR4$^+$-ADE in 2D and 3D settings. (c) Morphology of CD177 and CXCR4-β-cells; DAPI (blue) and E-CAD (green). Scale bars, 20 µm. Graph represents the size of the aggregates in µm. Mean ±s.e.m, $P<0.05$ determined using non-parametric Student's *t* test. Exact *P* value is mentioned in the graph.

Supplementary Fig. 8. H1 hESC-derived CD177-ADE generates more functional β-like cells *in vitro*. (a) Insulin content of US-β-cells and CD177-β-cells (n=10, biological replicates). (b) Comparison of insulin secretion of US-β-cells and CD177-β-cells in sequential static GSIS (n= 5, biological replicates). (c,d) Insulin secretion in response to dynamic glucose, Ex-4 and KCl challenges in a perifusion system on US-β-cells (c) and CD177-β-cells (d) (n=5, biological replicates). Data is represented as ±s.e.m; $P<0.05$ and $P<0.01$ is determined using non-parametric Student's *t* test. Exact *P* values are mentioned in the figure.

Supplementary Table 1: List of antibodies selected for secondary screen

No	Antibody	Isotype
1	Mouse IgG1-isotype control	
2	Mouse IgG2a-isotype control	
3	Mouse IgG2b-isotype control	
4	Mouse IgM-isotype control	
5	Rat IgG2a-isotype control	
6	Rat IgG2b- isotype control	
7	CD44	Mouse IgG1
8	CD90	Mouse IgG1
9	CD111	Mouse IgG1
10	CD133/1 (AC133)	Mouse IgG1
11	CD146	Mouse IgG1
12	CD262	Mouse IgG1
13	CD275 (B7-H2)	Mouse IgG1
14	Anti-PTK7 (CCK-4)	Mouse IgG2a
15	CD133/2 (293C3)	Mouse IgG2b
16	Anti-PSA-NCAM	Mouse IgM
17	CD15	Mouse IgM
18	CD49f	Rat IgG2a
19	ANTI CX3CR1	Rat IgG2b
20	Anti CCR10	Recombinant human IgG1
21	Anti HLA-DQ	Recombinant human IgG1
22	Anti-SSEA1	Recombinant human IgG1
23	Anti- SSEA4	Recombinant human IgG1
24	CD46	Recombinant human IgG1

25	CD47	Recombinant human IgG1
26	CD49b	Recombinant human IgG1
27	CD51	Recombinant human IgG1
28	CD82	Recombinant human IgG1
29	CD131	Recombinant human IgG1
30	CD138 (44F9)	Recombinant human IgG1
31	CD171 (LCAM)	Recombinant human IgG1
32	CD177	Recombinant human IgG1
33	CD244 (2B4)	Recombinant human IgG1
34	CD270 (HVEM)	Recombinant human IgG1
35	CD184	Mouse IgG2a
36	DCIR	Recombinant human IgG1
37	CD234	Recombinant human IgG1
38	Anti- LGR5	Rat IgG2b
39	CD166	Recombinant human IgG1
40	CD195 CCR5	Recombinant human IgG1
41	Anti SSEA-5	Mouse IgG1κ
42	Anti-NOTCH1	Recombinant human IgG1
43	CD41a (ITGA2b)	Recombinant human IgG1
44	CD49c (ITGA3)	Recombinant human IgG1
45	CD140b	Recombinant human IgG1
46	CD181 (CXCR1)	Mouse IgG2bκ

1022

1023

1024

1025

Supplementary Table 2: Stage-wise comparison of pancreatic progenitors and β-like cells (S3-S7) generated from enriched CD177$^+$ ADE and CXCR4$^+$ ADE vs already published protocols*.

Stage (% of cells from total population)	CD177$^+$ derived pancreatic precursors	CXCR4$^+$ derived pancreatic precursors	Bulk differentiationaccording to Rezania et.al., 2014 (as published)	Bulk differentiation according to Pagliuca et.al., 2014 (as published)
S3: PDX1$^+$ cells	>70%	>50%	>89%	>85%
S4: PDX1$^+$/NKX6.1$^+$	>60%	>50%	>62%	>55%
S6: INS$^+$/NKX6.1$^+$ GCG$^+$/INS$^+$	>60% >13%	>50% >10%	>44% >20%	**C-peptide$^+$/NKX6.1$^+$:** >38% **GCG$^+$/C-peptide$^+$:** >8%
S7: INS$^+$/MAFA$^+$ INS$^+$/GLUT1$^+$	>30% >30%	>20% >20%	NA NA	NA NA

* NA: not available

A direct comparison between GLUT1$^+$/INS$^+$ β-like cells derived from CXCR4$^+$-, CD177$^+$-ADE compared to already published protocols at S7 stage was not possible as both the protocols did not check for the expression of GLUT1.

Supplementary Table 3: List of primers for qPCR

Primer	Sequence/TaqMan id
GAPDH	Hs02758991_g1
18S	Hs99999901_s1
FOXA2	Hs00232764_m1
SOX17	Hs00751752_s1
CER1	For: CCCATCAAAAGCCATGAAGT Rev: TTTCCCAAAGCAAAGGTTGT
HHEX	For: ACGGTGAACGACTACACGC Rev: CTTCTCCAGCTCGATGGTCT
CELSR1	Hs00947712_m1
WNT4	Hs01573504_m1
WNT5A	Hs01086864_m1
DVL2	Hs00182901_m1
WNT3A	Hs00263977_m1
AXIN2	Hs00610344_m1
NGN3	Hs01875204_s1
PDX1	Hs00236830_m1
NKX6.1	Hs01055914_m1
NEUROD1	Hs01922995_s1
MAFA	Hs01651425_s1/Hs04186804_s1
GLUT1	Hs00892681_m1
UCN3	Hs00846499_s1
TTR	For: ACTTGGCATCTCCCCATTC Rev:TAGGAGTAGGGGCTCAGCAG
AFP	Hs00173490_m1
NKX2.2	Hs00159616_m1
INS	Hs02741908_m1
GCG	Hs01031536_m1

Supplementary Table 4: List of Antibodies for Immunohistochemistry, FACS and Western blotting

Conjugated antibodies:

Antibody	Company	Catalogue No	Dilution FACS
Human CXCR4-PE Human CXCR4-APC	Miltenyi Biotec Miltenyi Biotec	130-098-354 120-010-802	1:40
Human CD117-APC Human CD117-PE	Miltenyi Biotec Miltenyi Biotec	130-091-733 130-091-734	1:40
FOXA2	R and D	IC2400G	1:10
SOX17	R and D	IC1924A	1:10
Human CD177-APC	Miltenyi Biotec	120-017-498	1:20
Human CD275-APC	Miltenyi Biotec	120-012-112	1:20
PE Mouse anti-PDX1	BD Pharmingen™	562161	1:40
Alexa Fluor® 647 Mouse anti-Nkx6.1	BD Pharmingen™	563338	1:40
Alexa Fluor® 647 Mouse IgG1 κ Isotype Control	BD Pharmingen™	563023	1:40

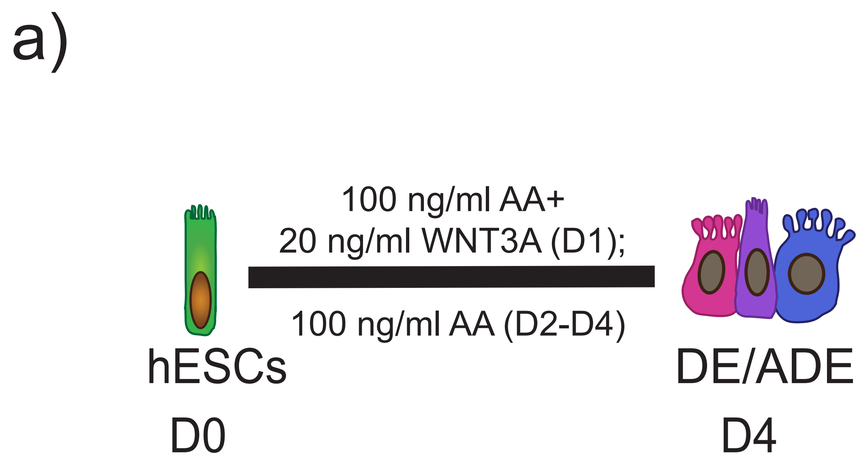
Unconjugated antibodies:

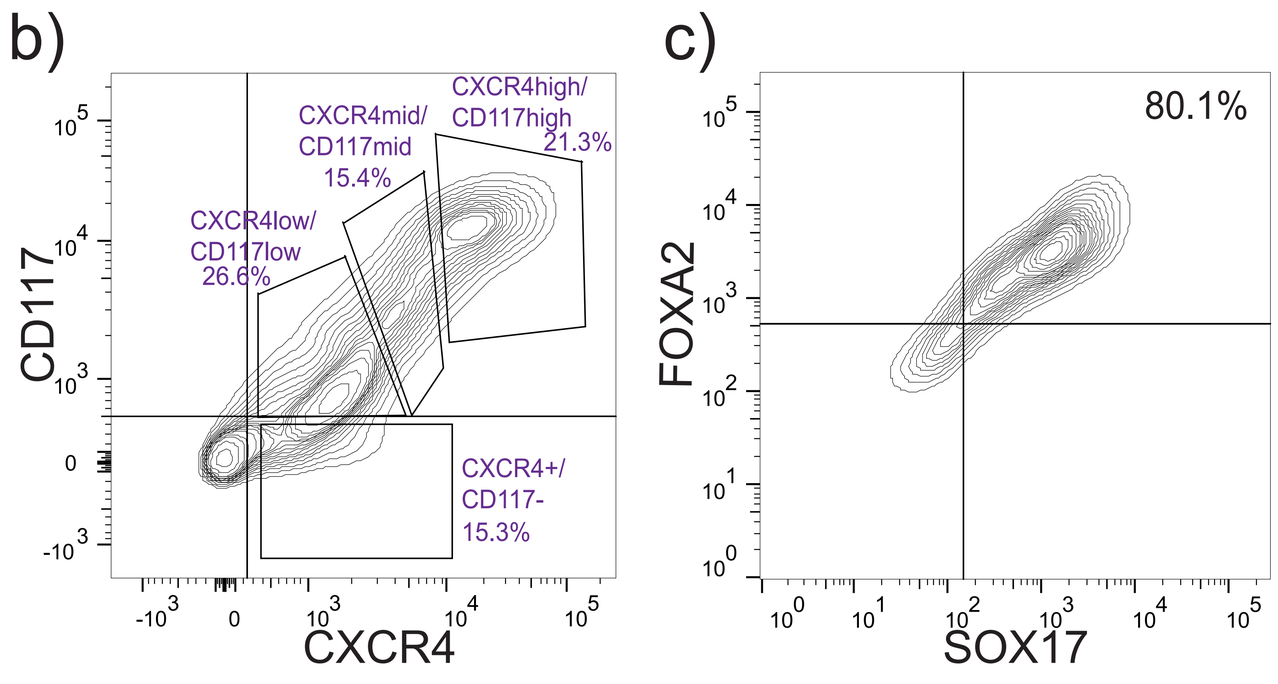
Antibody	Company	Catalogue No	Dilution FACS	Dilution IF
Rabbit FOXA2	Cell signalling	8186	1:1000	1:1000
Goat SOX17	Acris/Novus	GT15094	1:1000	1:1000
Goat CER1	R&D Systems	AF1075	1:1000	1:1000
Mouse β-catenin	BD	610154	1:1000	1:1000
Guinea pig INSULIN	Thermo Schientific	PA1-26938	1:100	1:100
Guinea pig C-peptide	Abcam	ab30477	1:300	1:300
Rabbit MAFA	Betalogics	LP9872	1:100	1:100
Rabbit GLUT1	Thermo Fisher	PA1-37782	1:100	1:100
Goat GATA6	R&D Systems	AF1700	1:1000	1:1000
Mouse SOX2	Abgent / Bio Cat	AM2048	1:1000	1:1000
Rabbit CDX2	Santa Cruz	sc-134468	1:1000	1:1000
Goat PTF1A	NCBI	AB2153	1:1000	1:1000
Mouse GCG	Sigma	G2654-.2ML	1:300	1:300
Goat PDX1	R&D Systems	AF2419	1:100	1:500
Rabbit NKX6.1 Goat NKX6.1	Novus biologicals R&D systems	NBP1-49672 AF5857	1:2000 1:300	1:5000 1:300

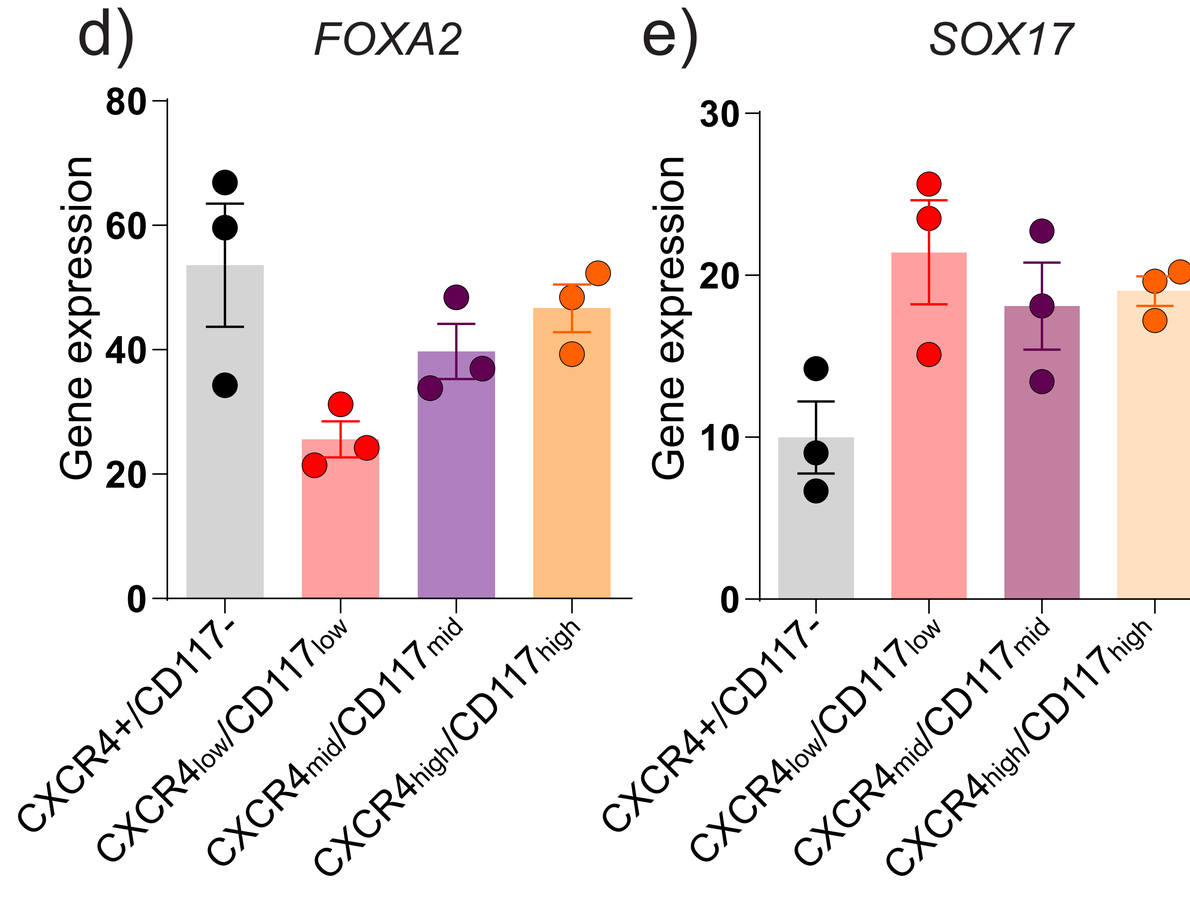
Unconjugated antibodies:

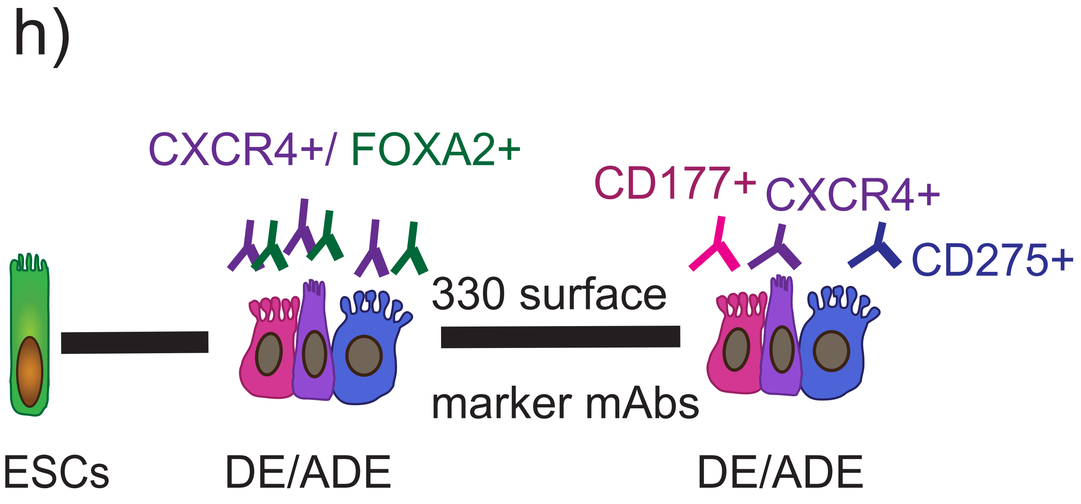
Antibody	Company	Catalogue No	Dilution
Rabbit p-JNK	Cell signalling	4668	1:1000
Rabbit DVL2	Cell signalling	3216	1:1000
Mouse β-catenin	BD	610154	1:2000
Mouse GAPDH	Merck Biosciences	CB1001	1:6000

Figure 1. Identification of novel CD177⁺ and CD275⁺ ADE subpopulations









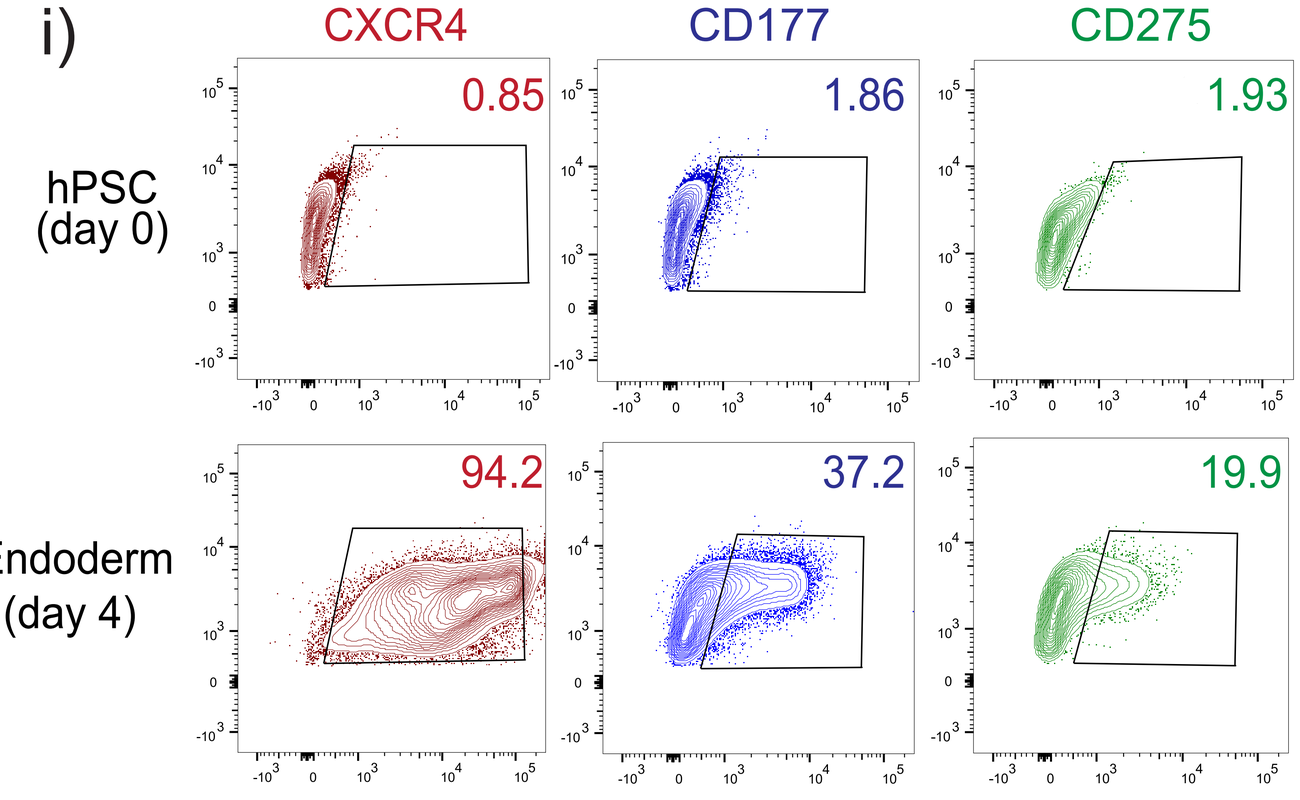
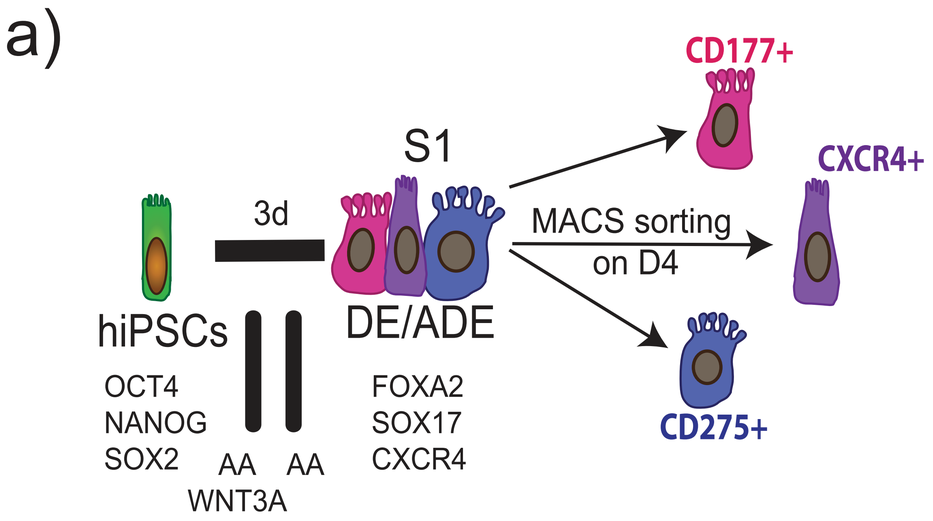
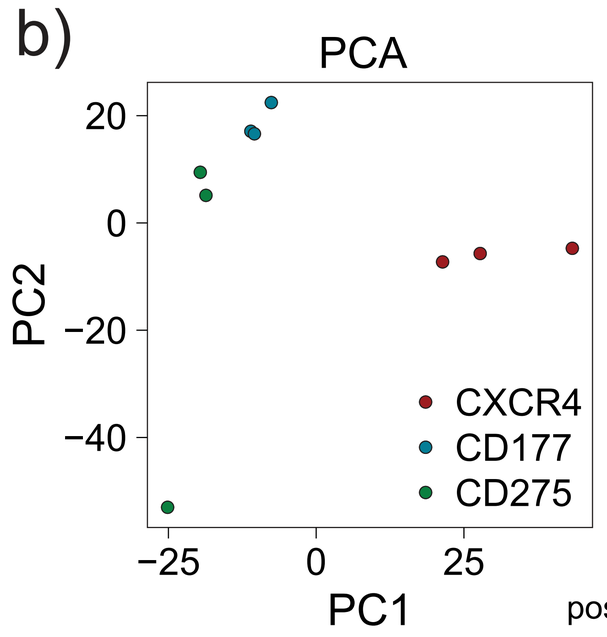
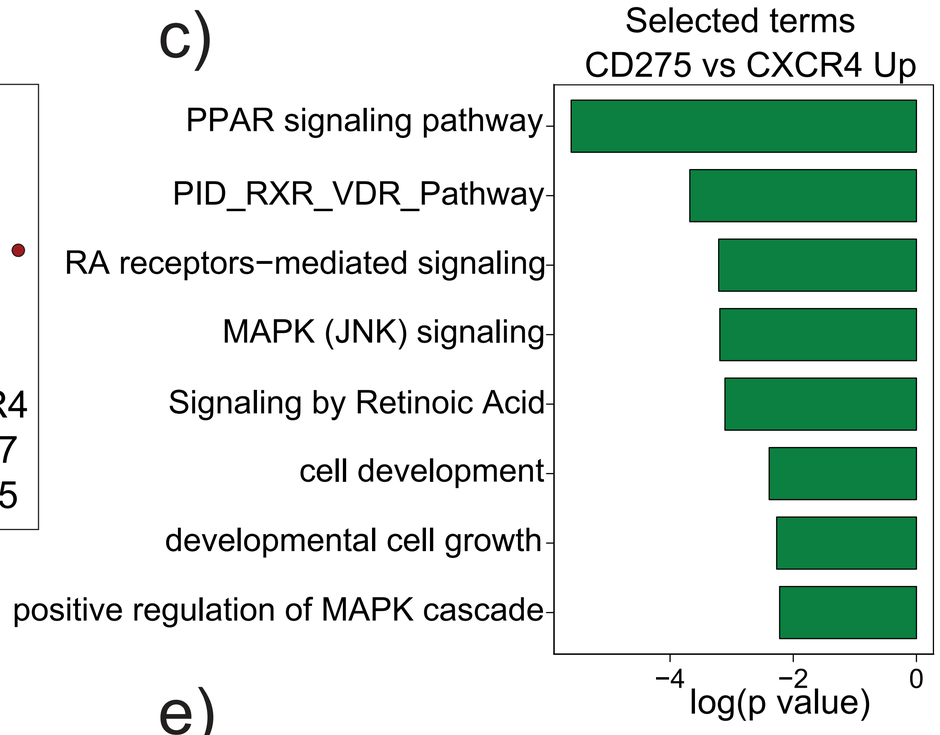
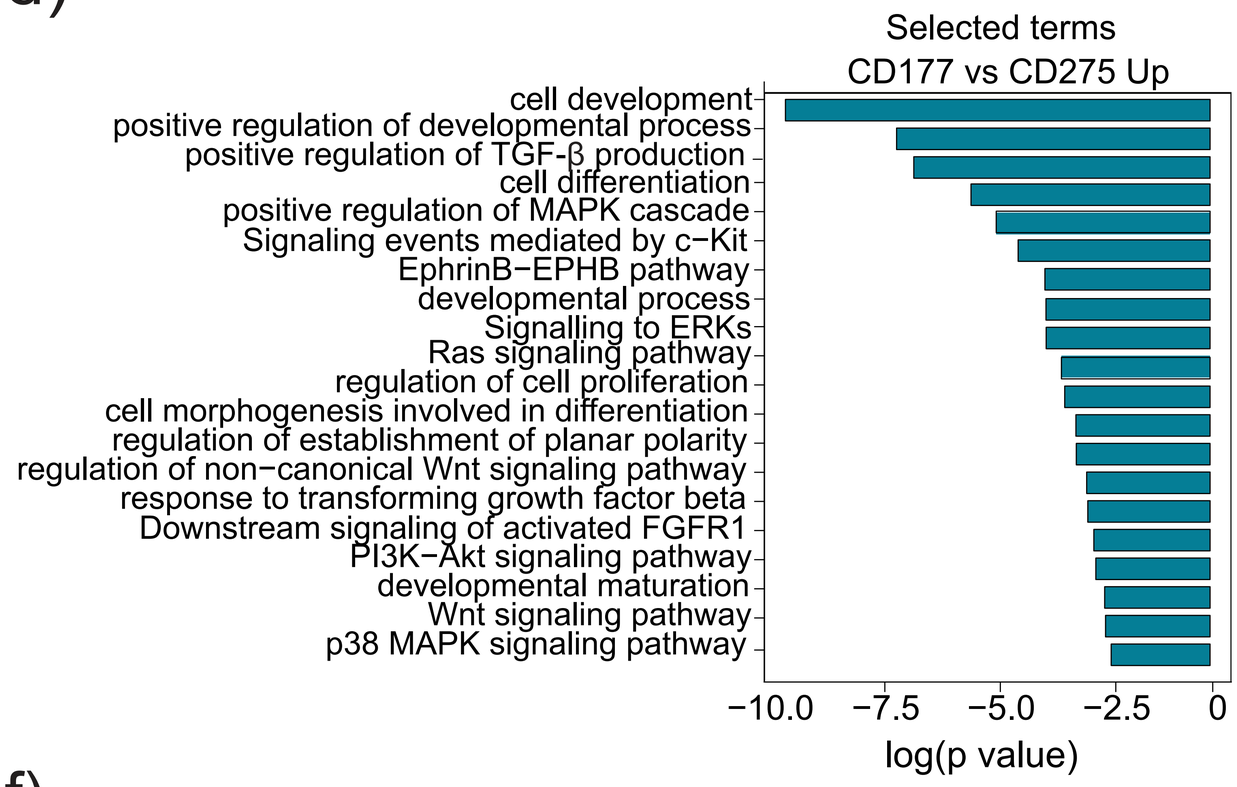


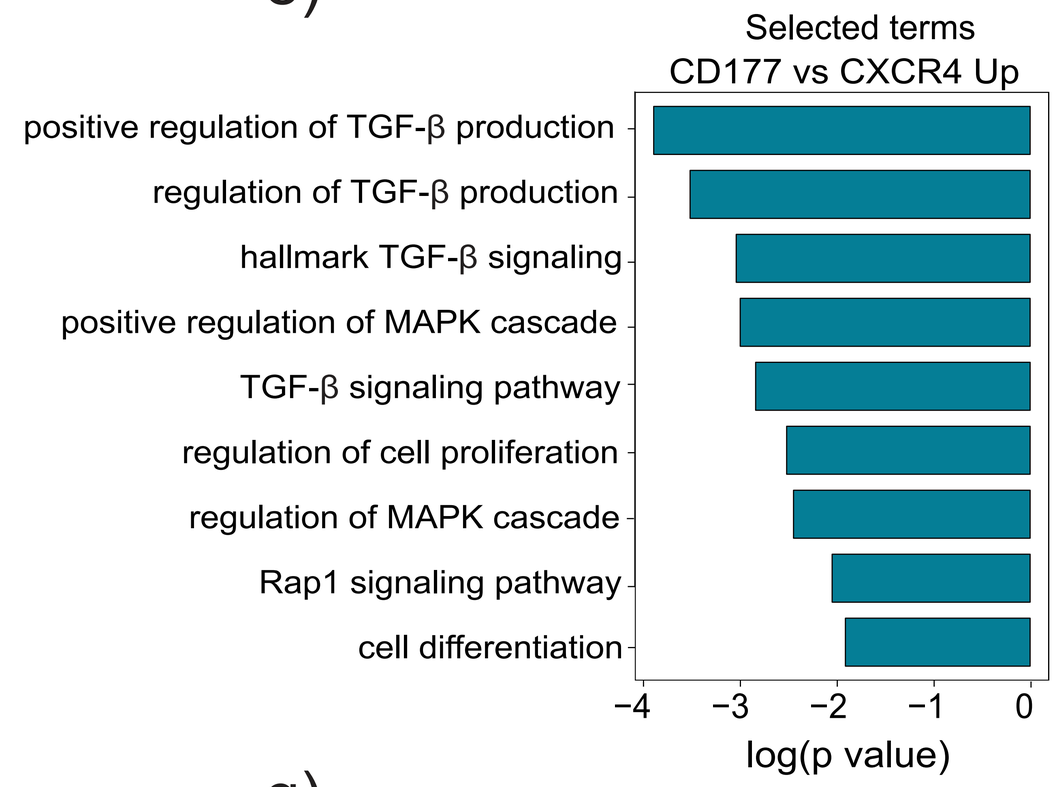
Figure 2. Molecular profiling of CD177⁺, CD275⁺ and CXCR4⁺ DE subpopulations reveal distinct signatures

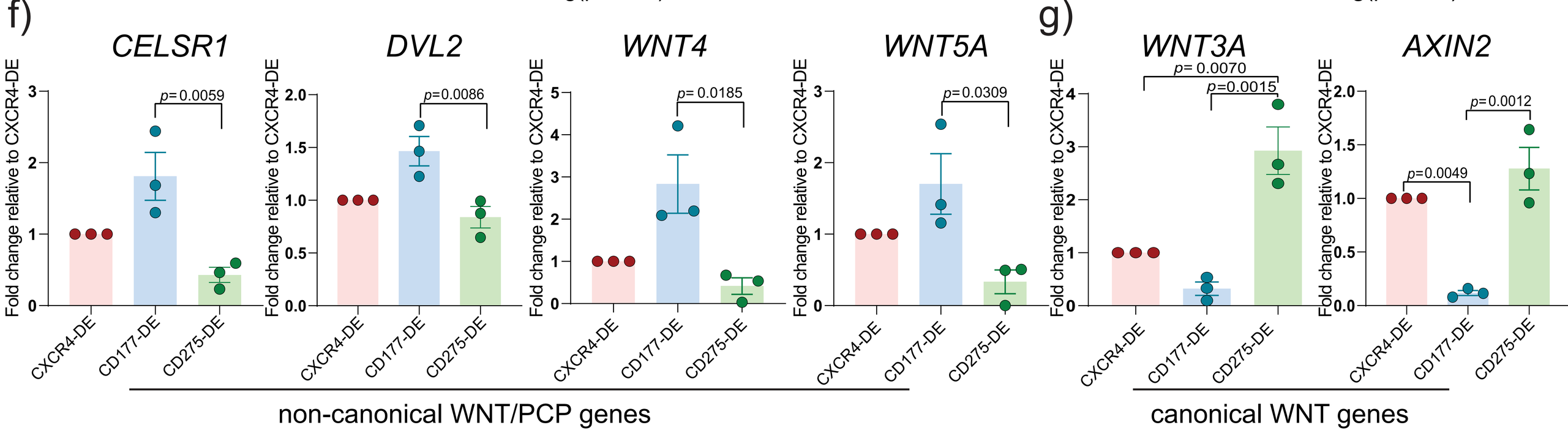


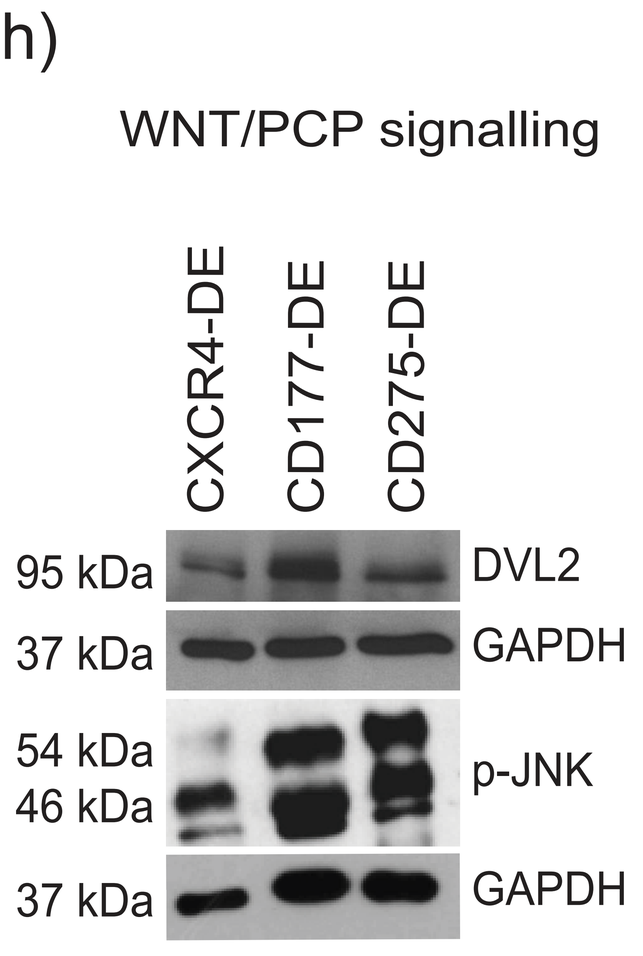


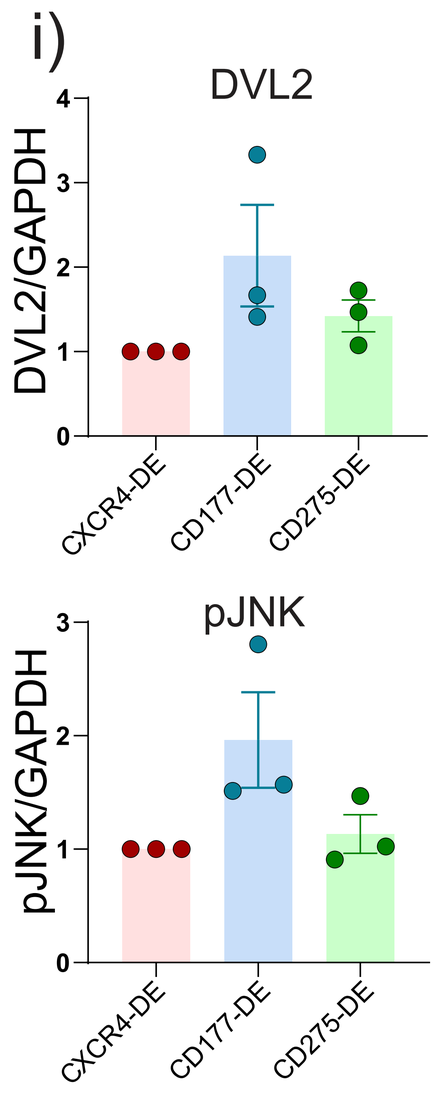












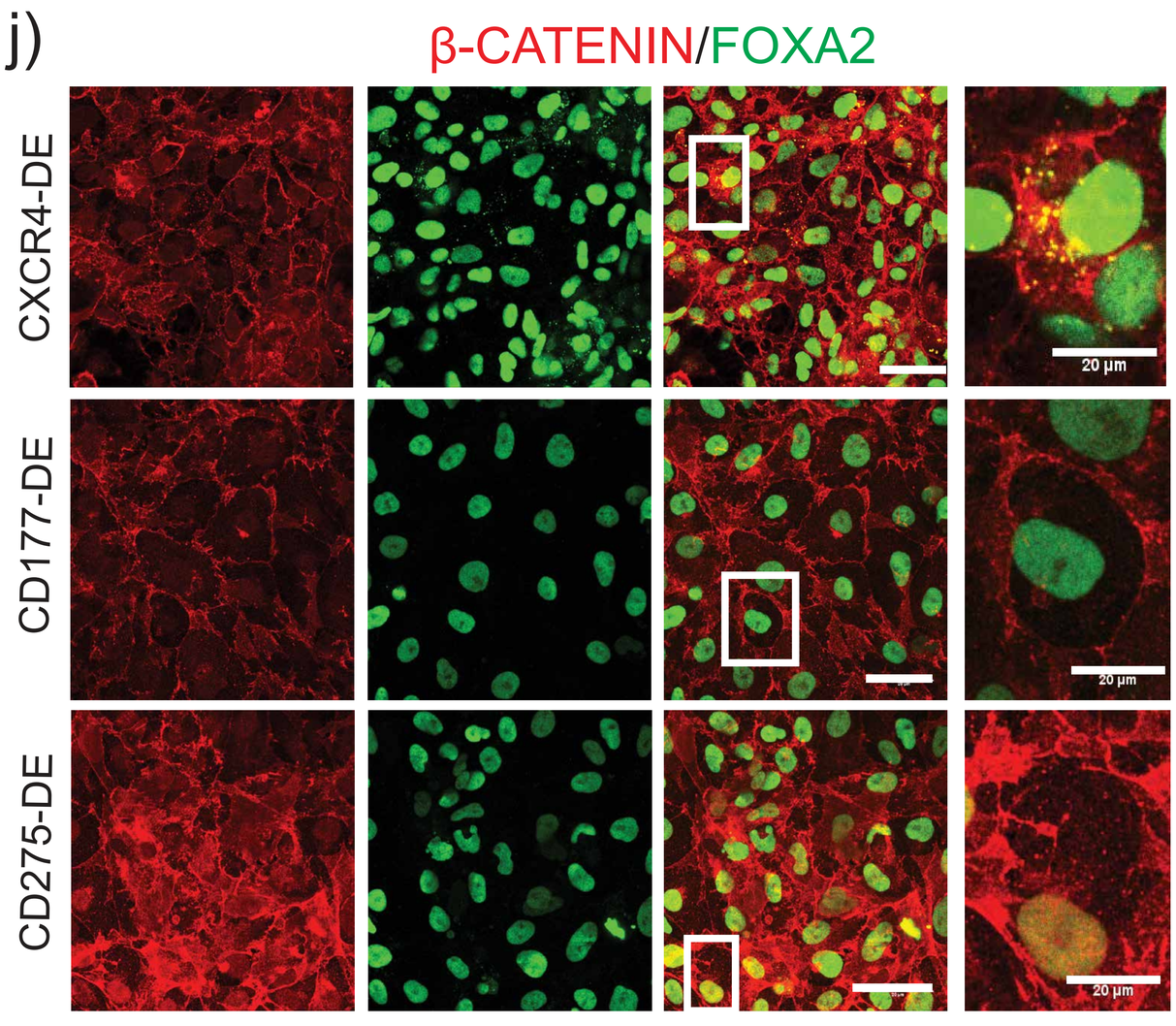
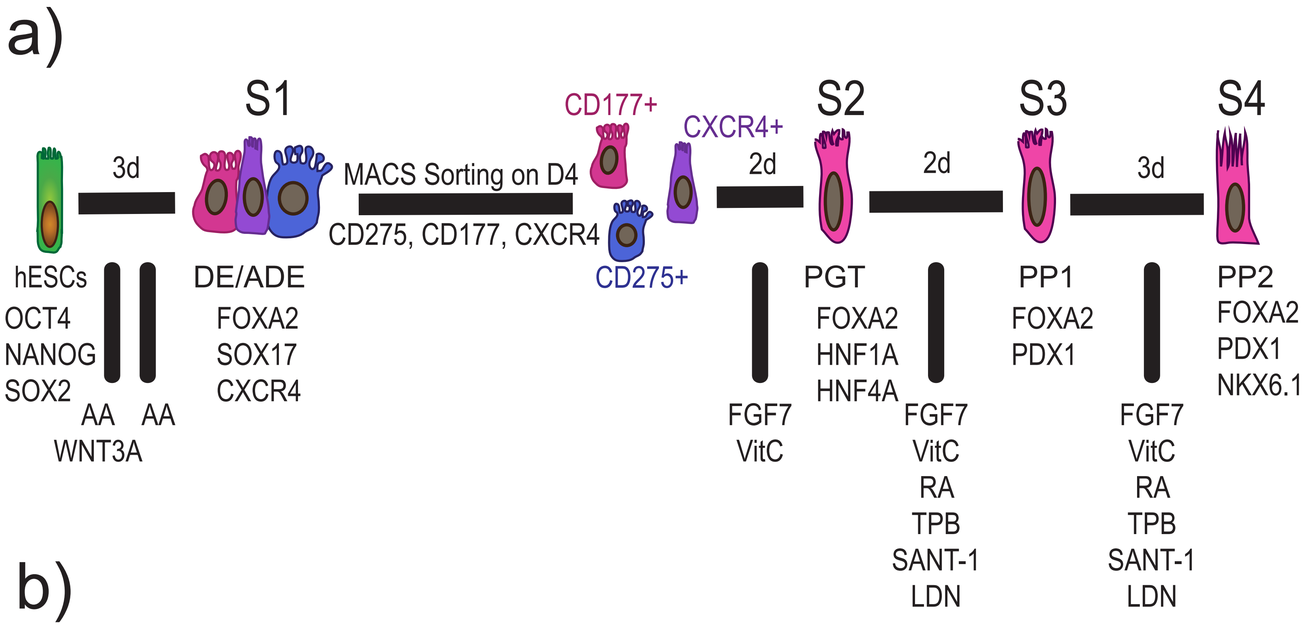
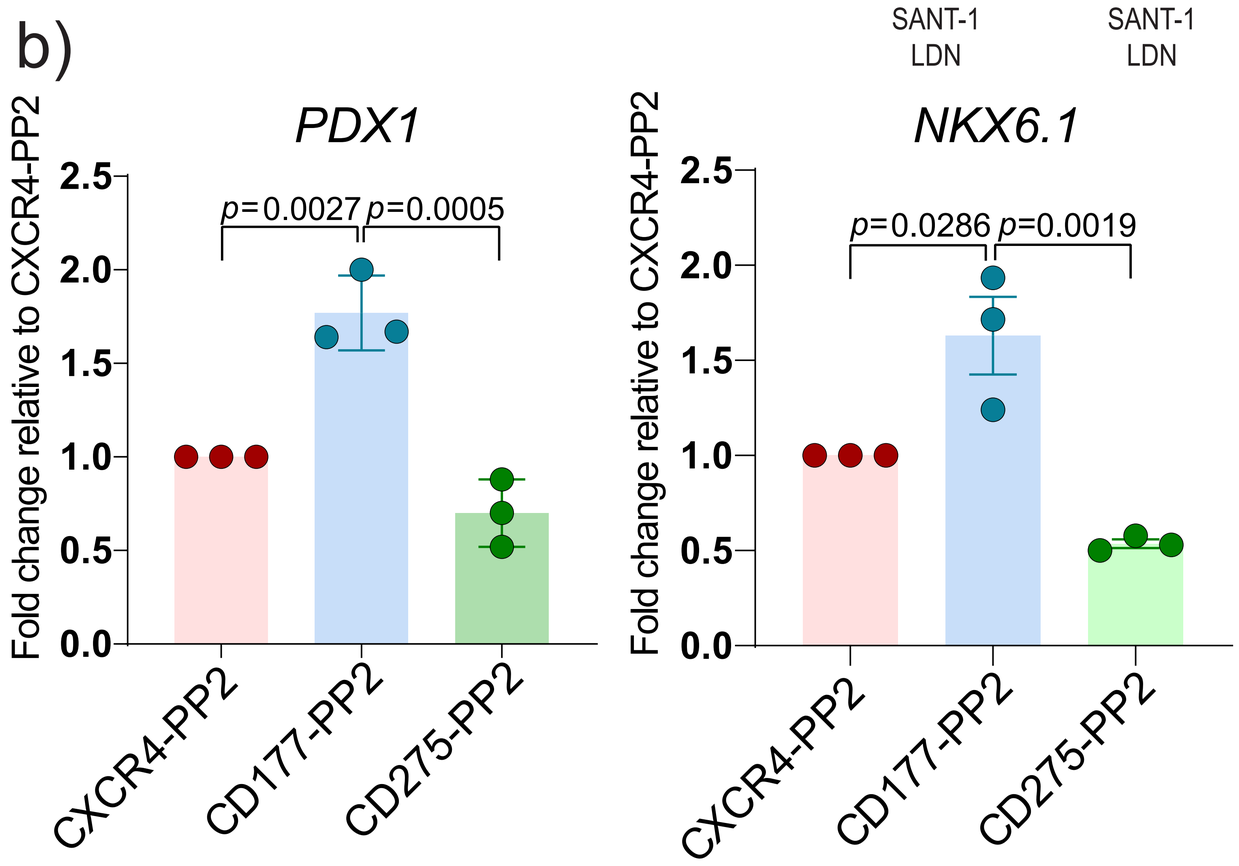
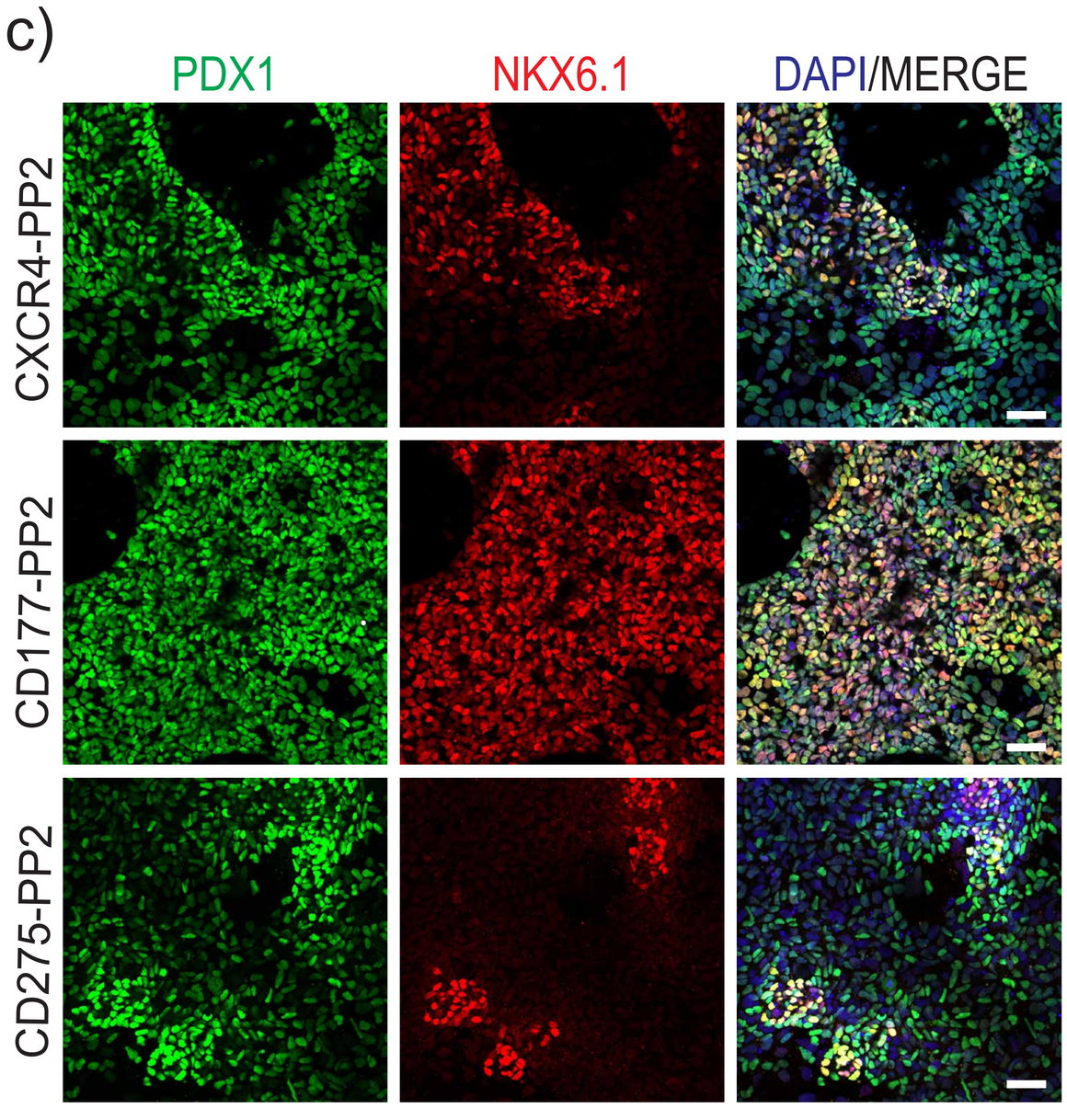
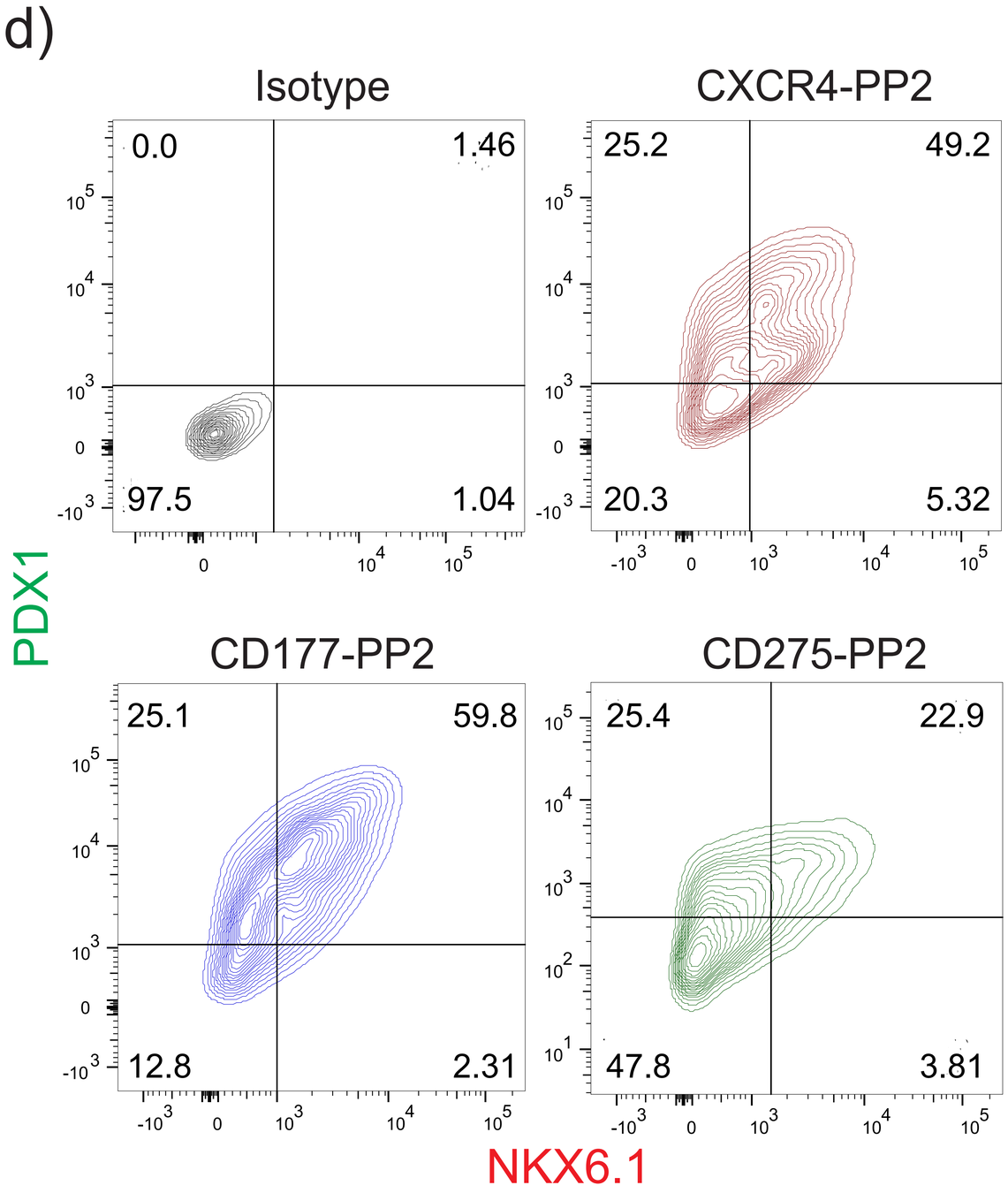


Figure 3. CD177⁺ ADE efficiently differentiates into pancreatic progenitors









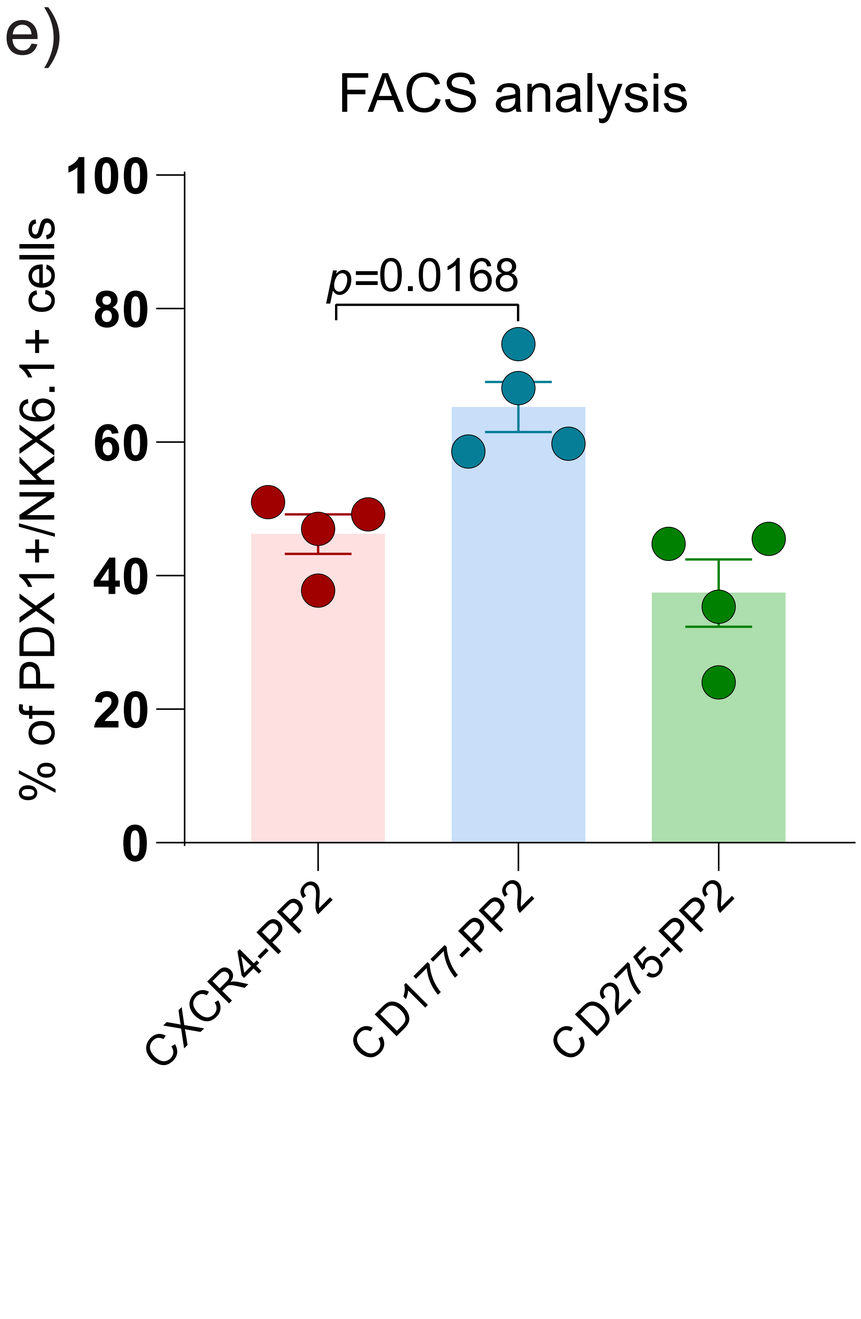
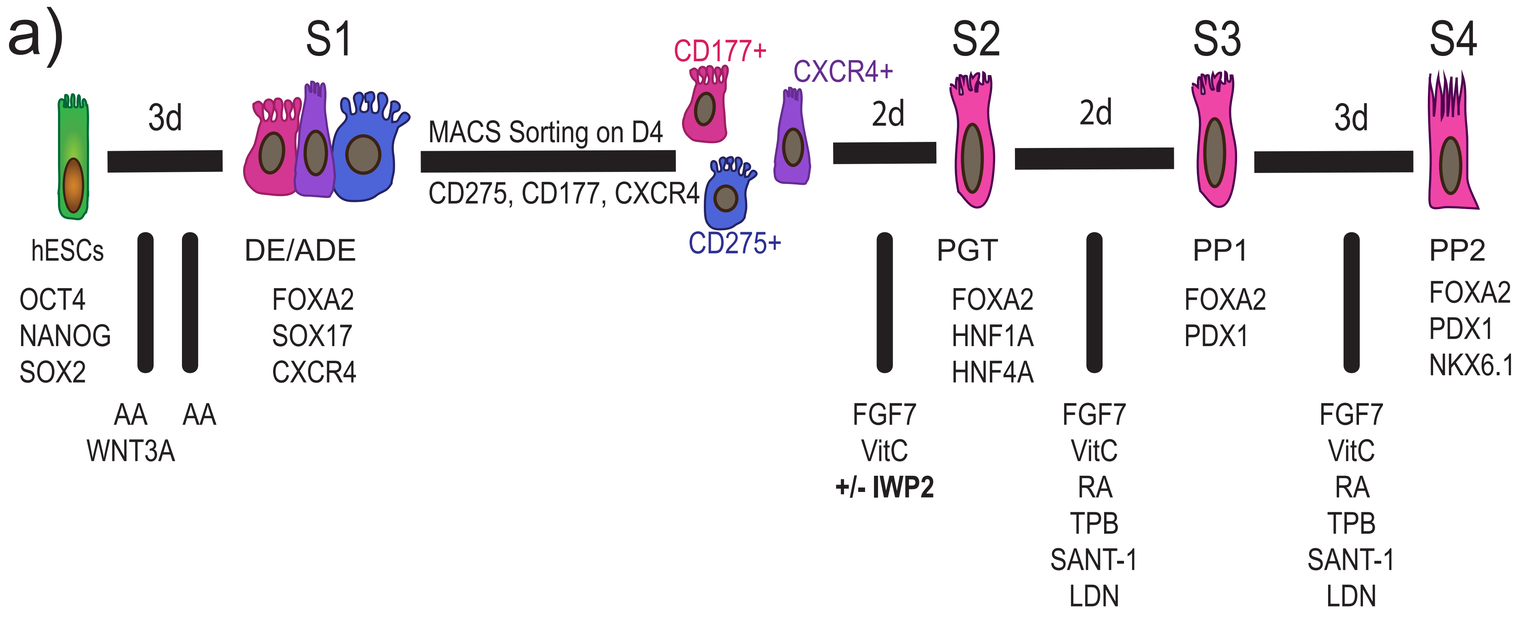
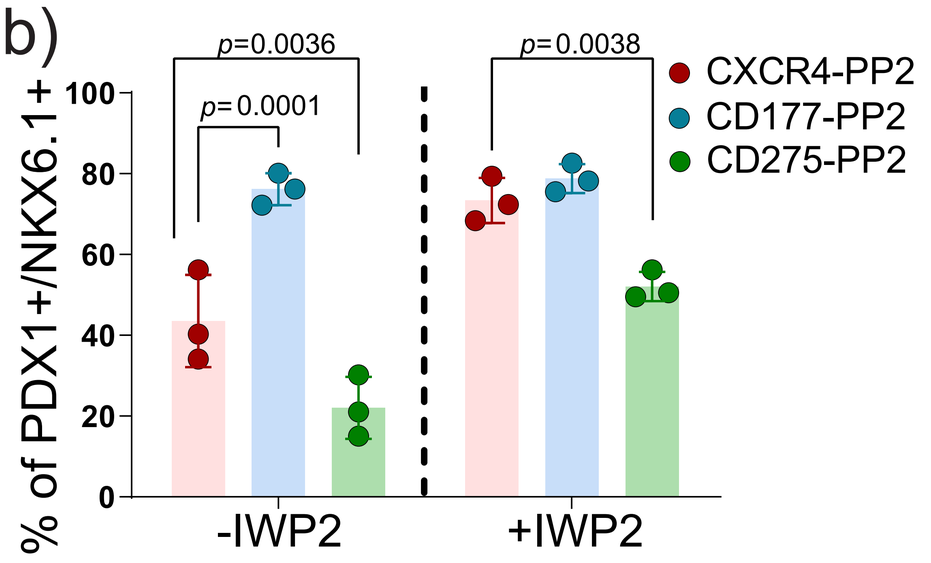
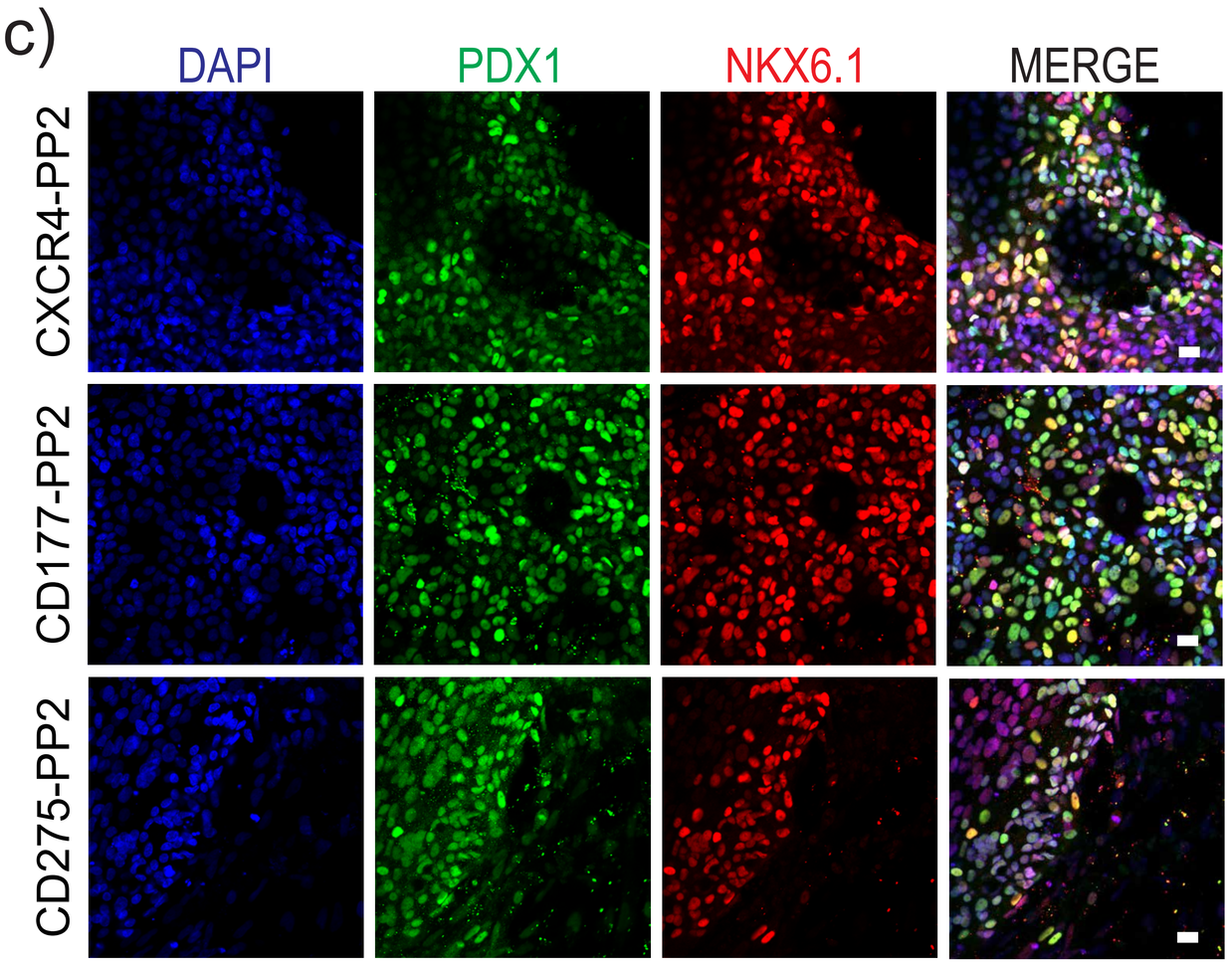
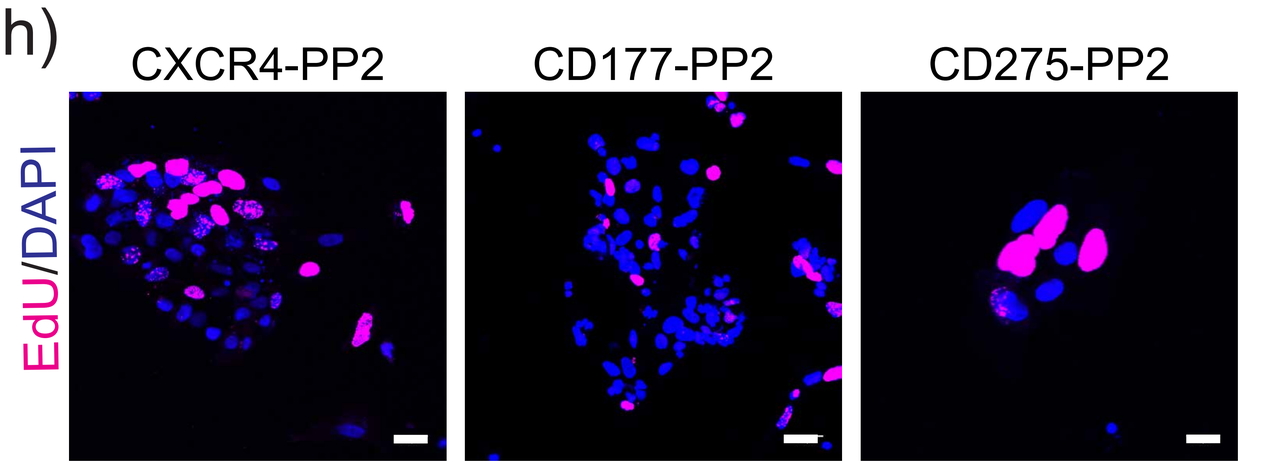


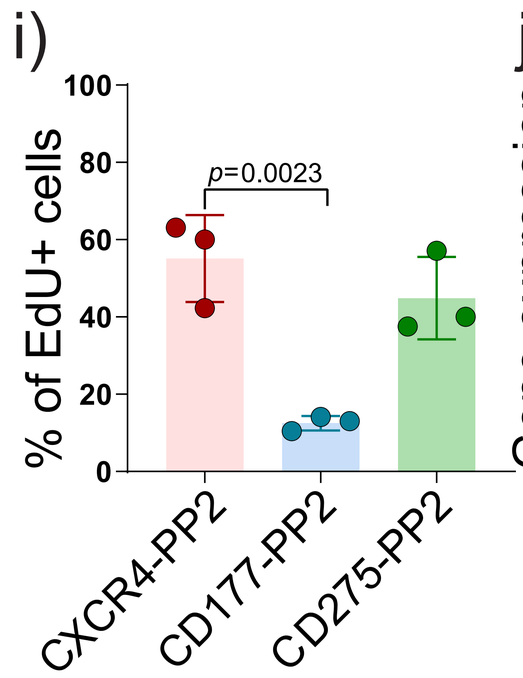
Figure 4. Inhibition of canonical WNT secretion promotes pancreatic differentiation

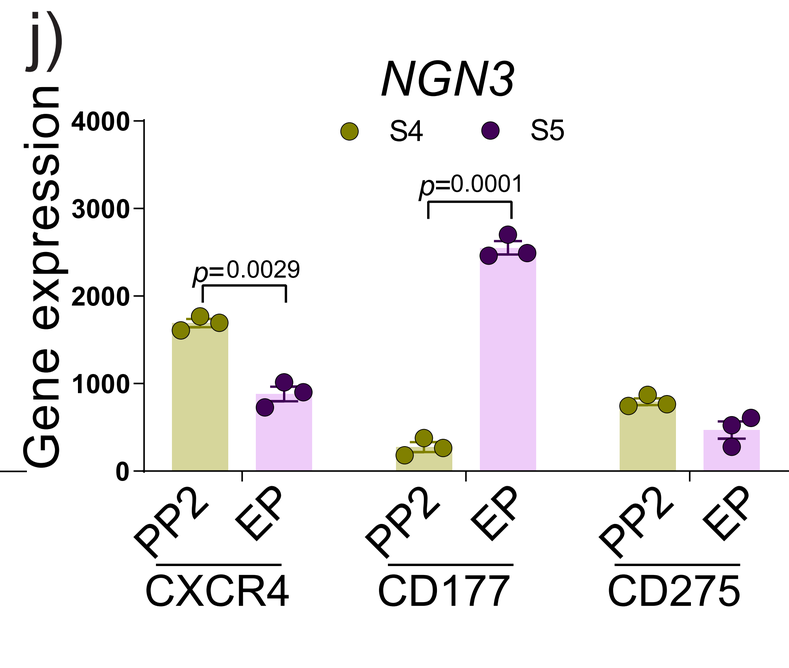


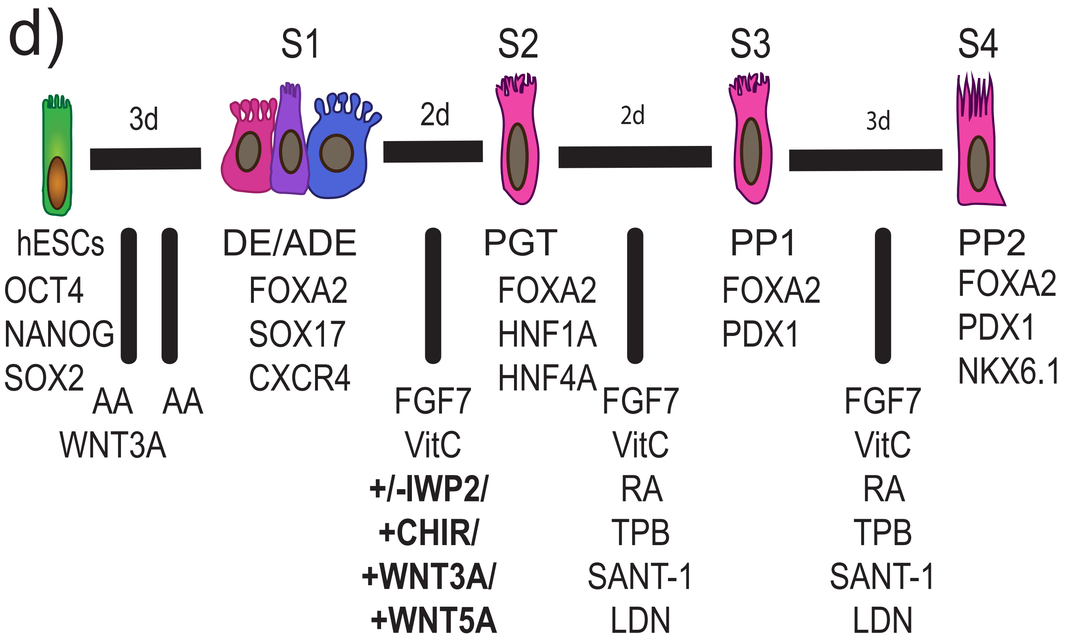


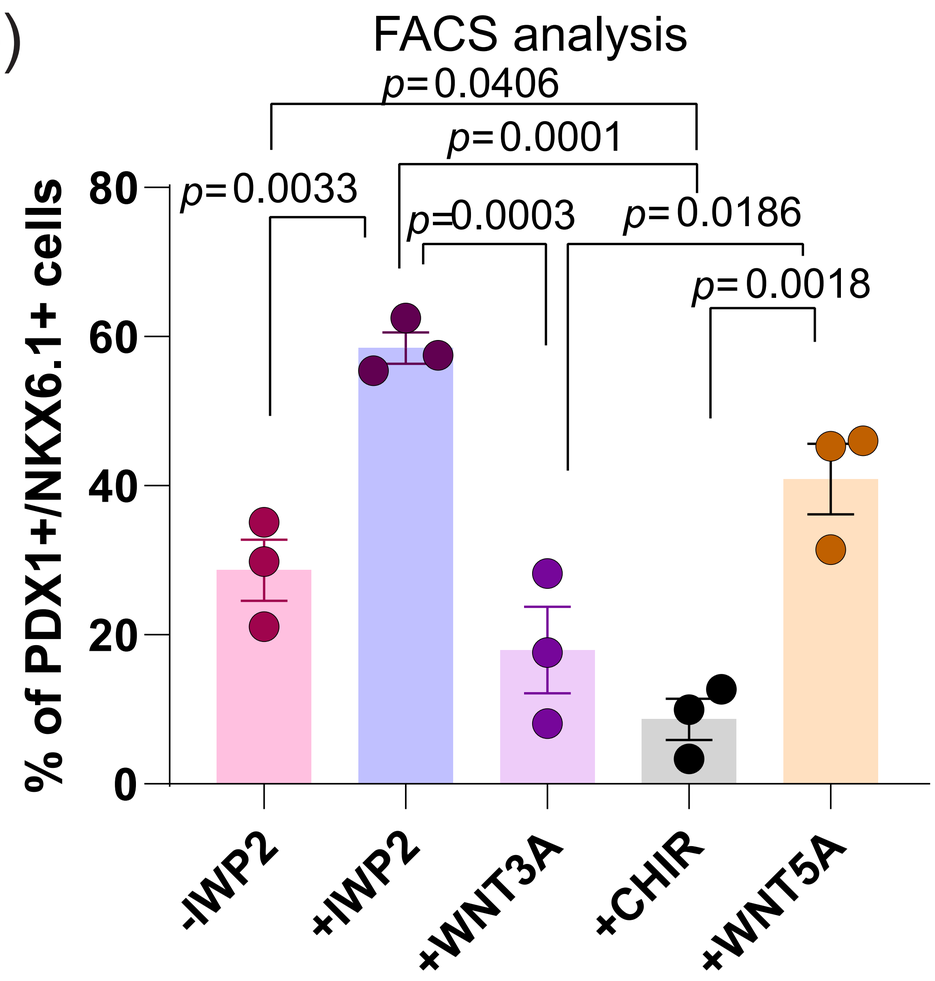


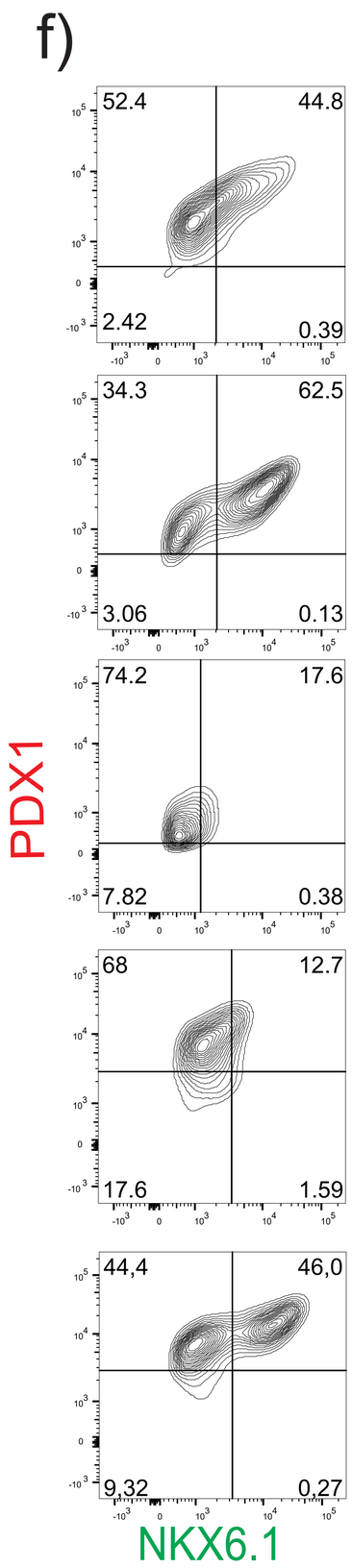












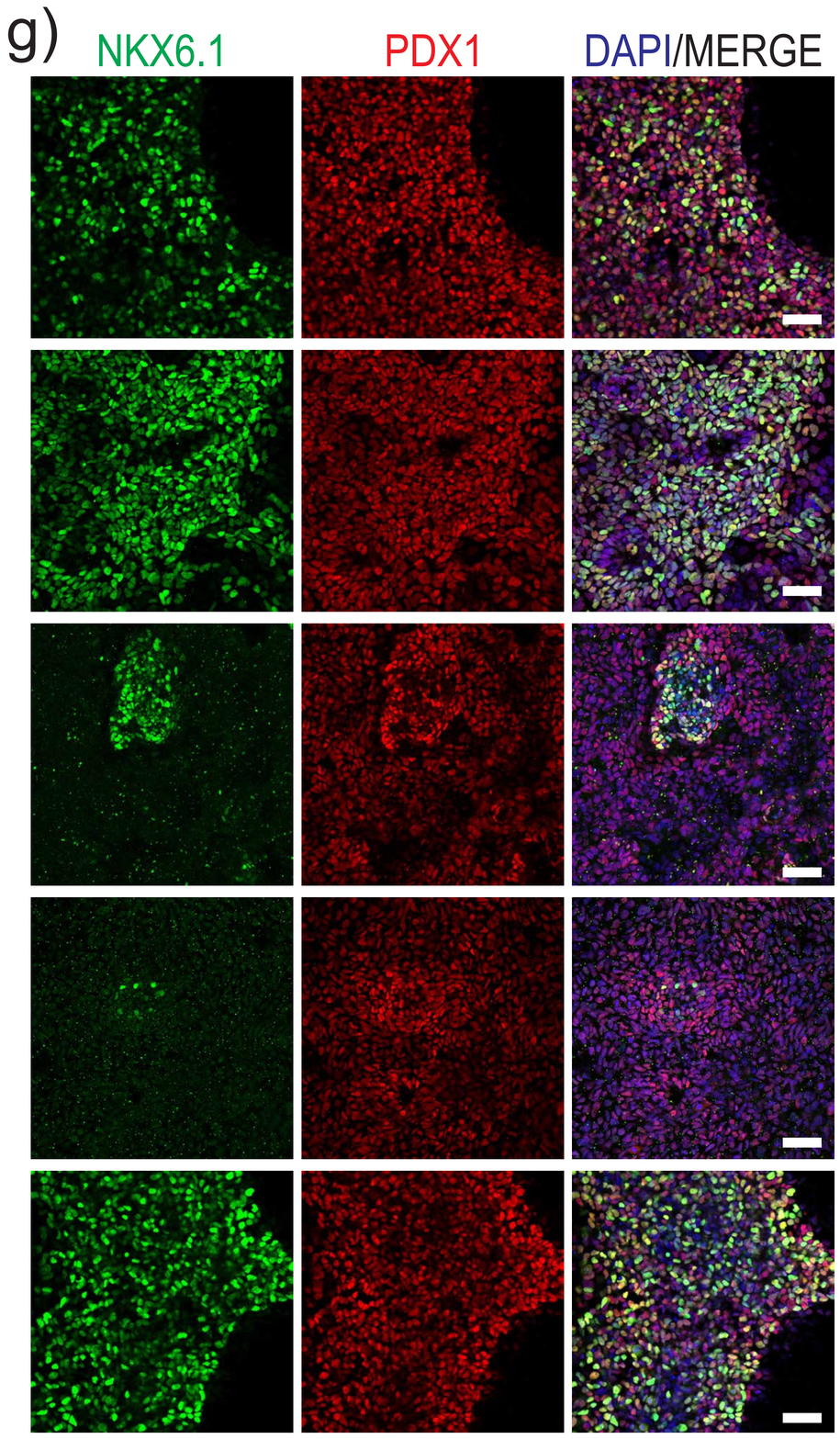
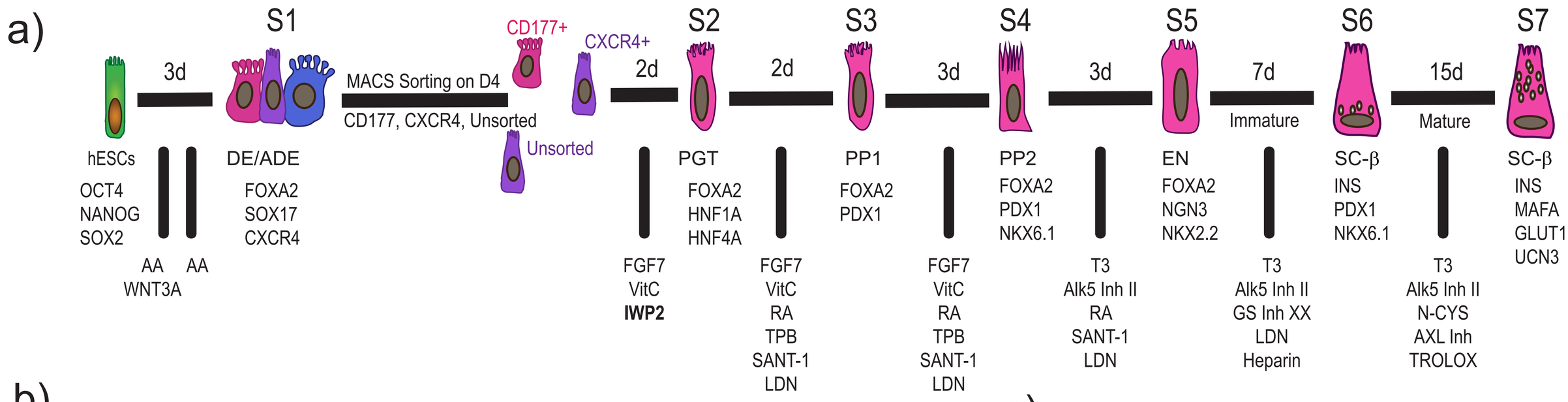
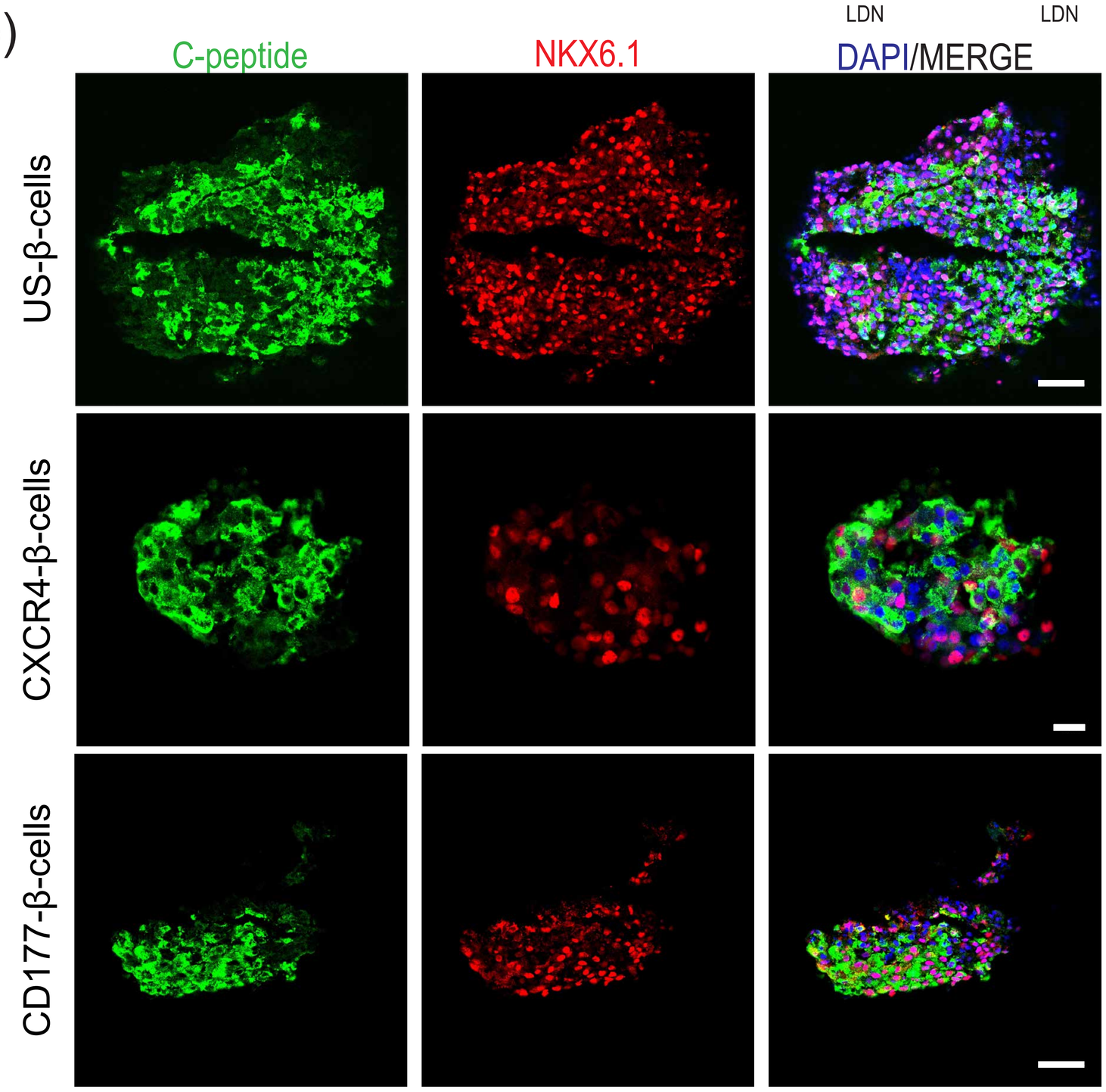
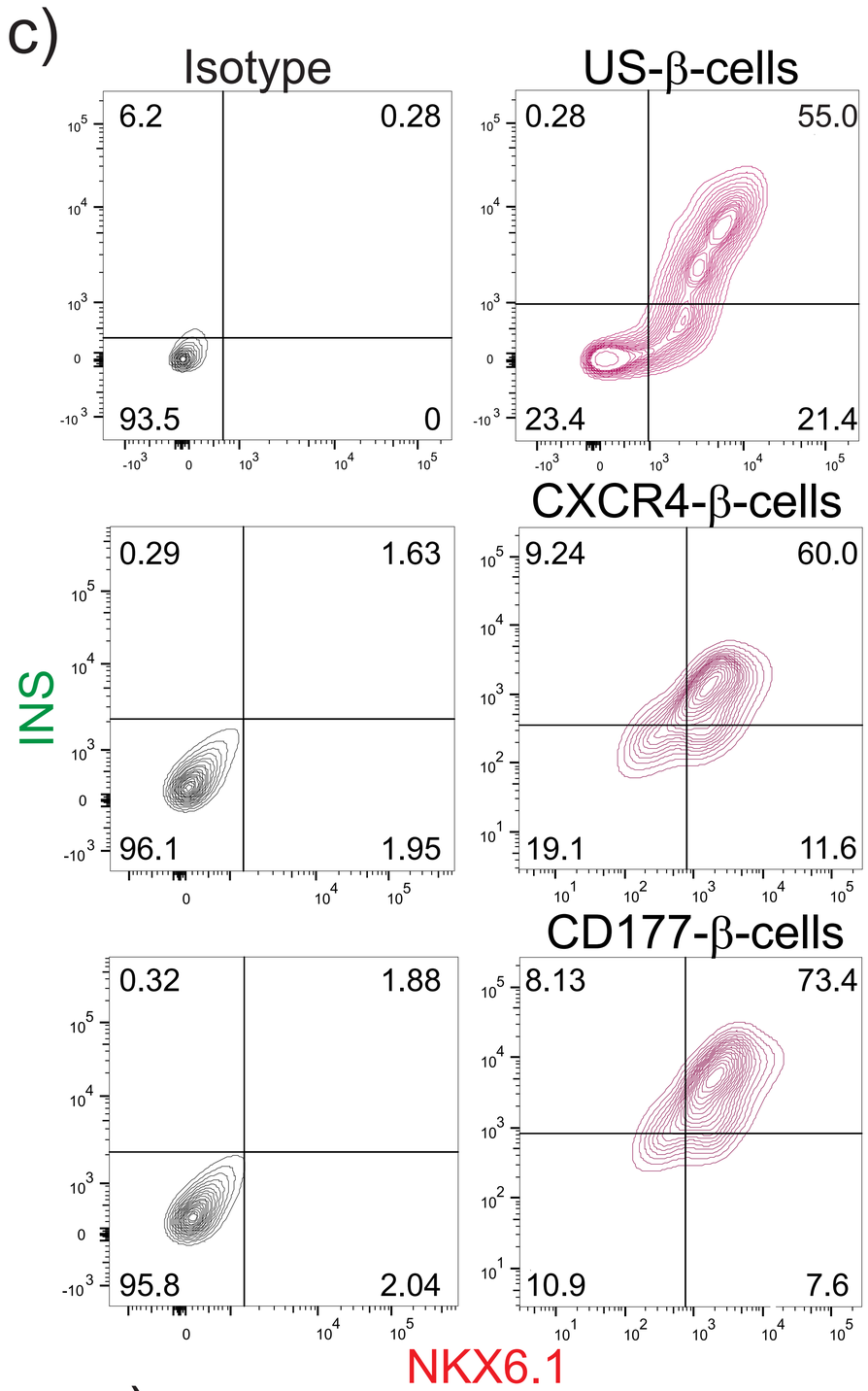
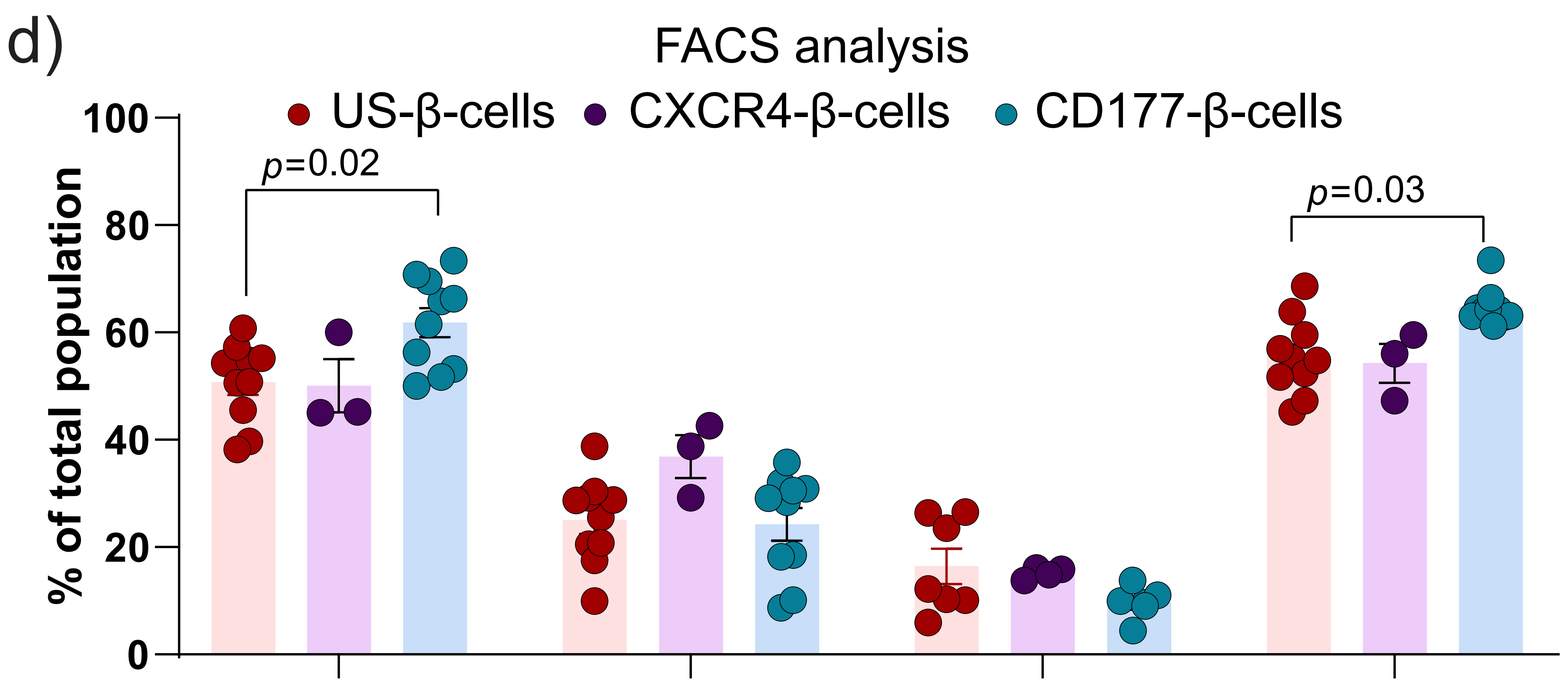


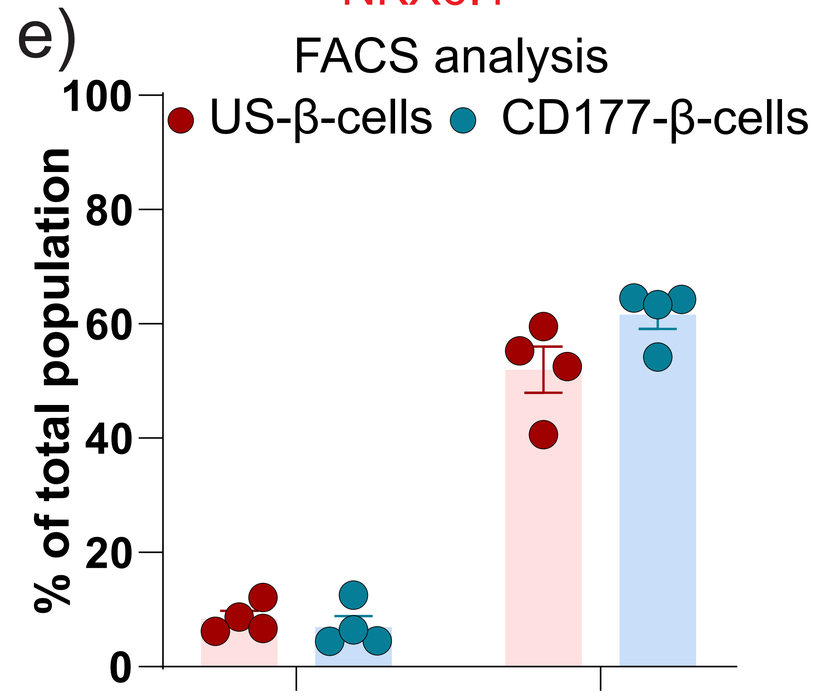
Figure 5. CD177⁺ ADE efficiently differentiates into β-like cells

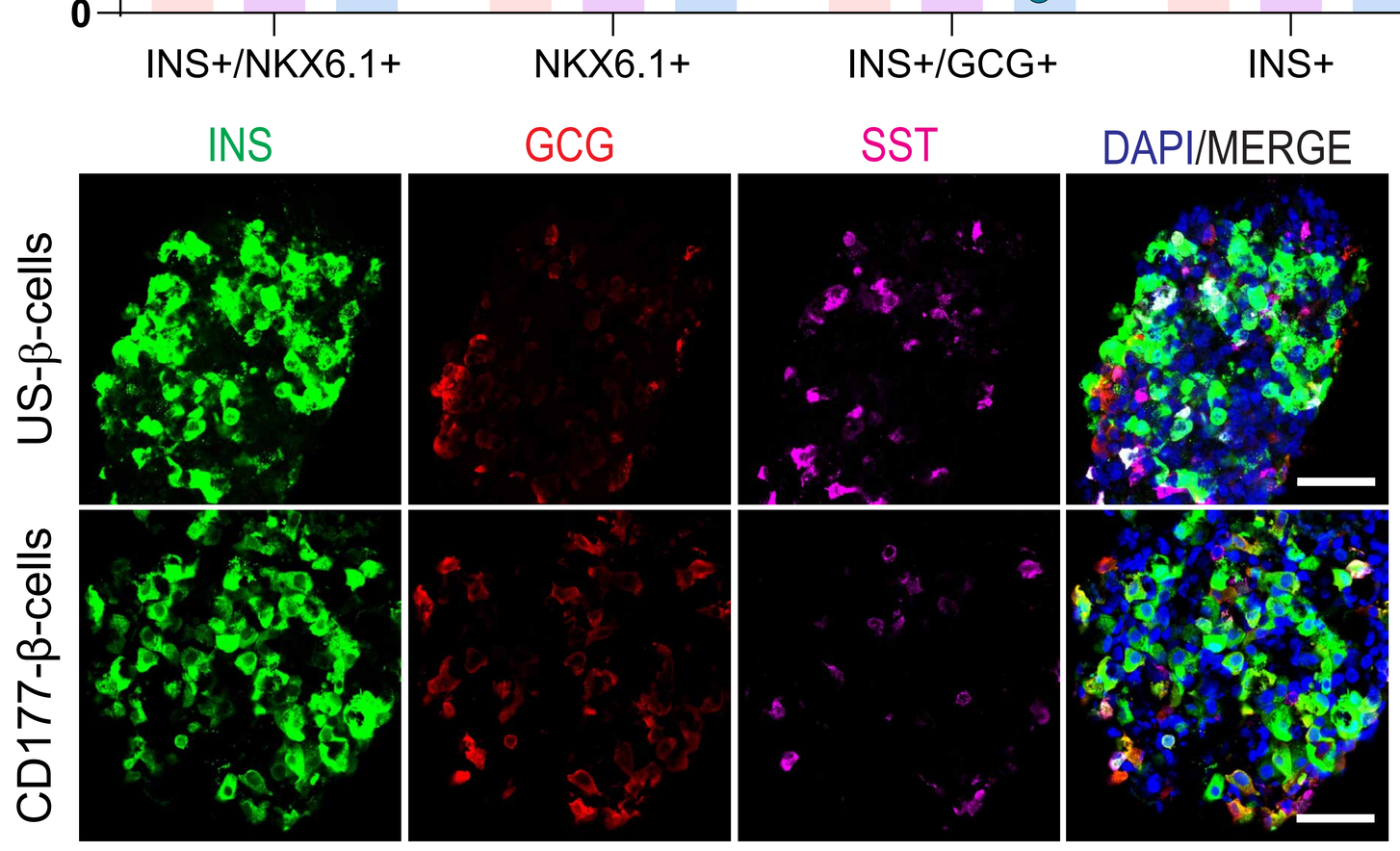












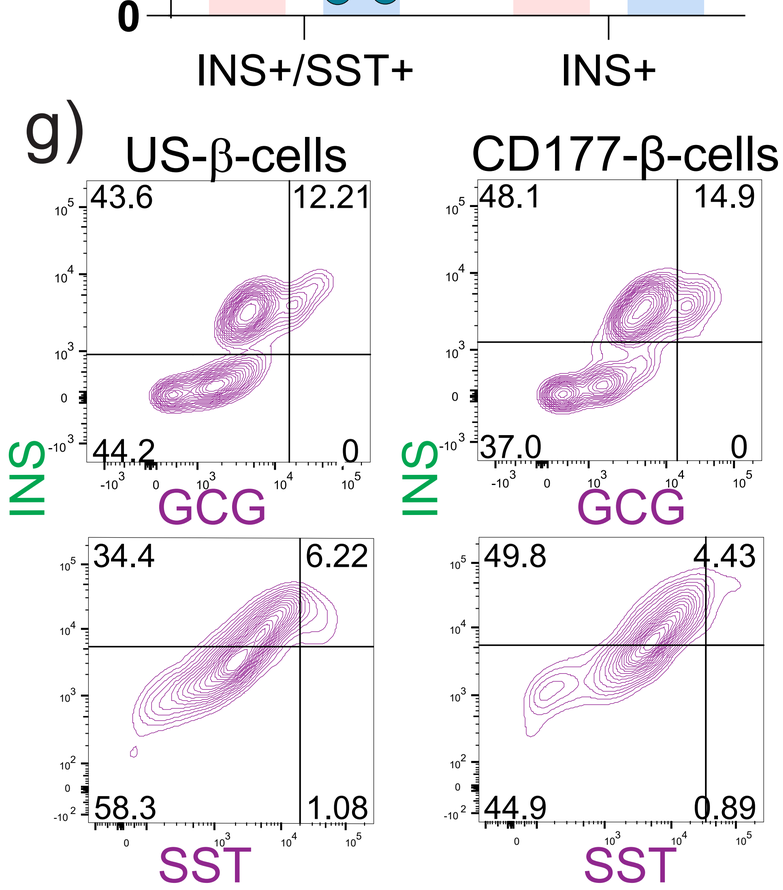
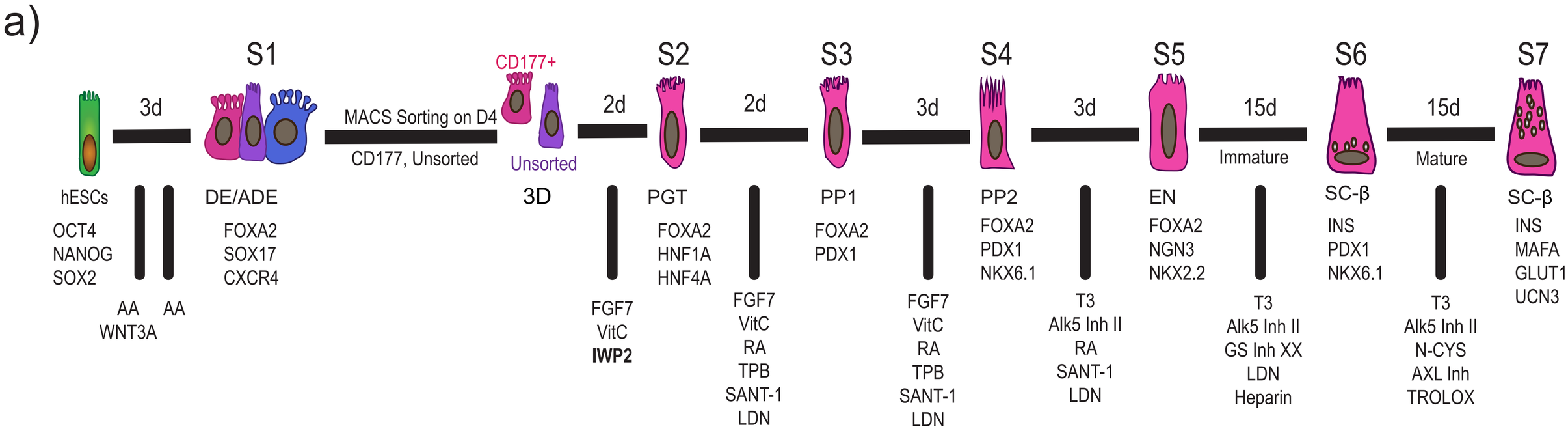
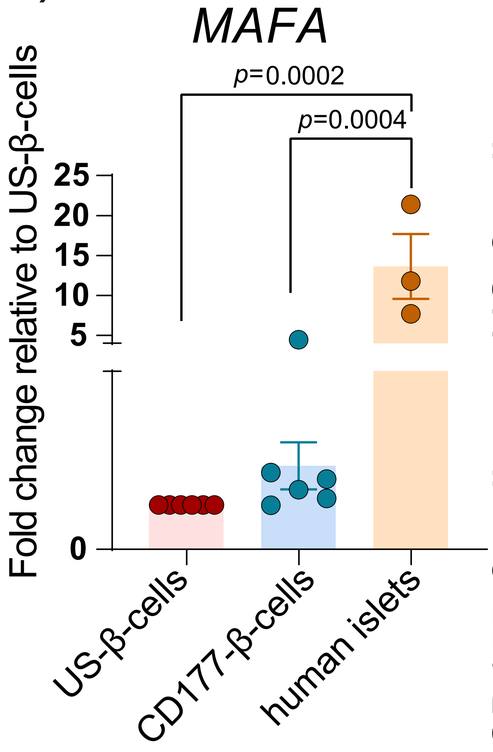
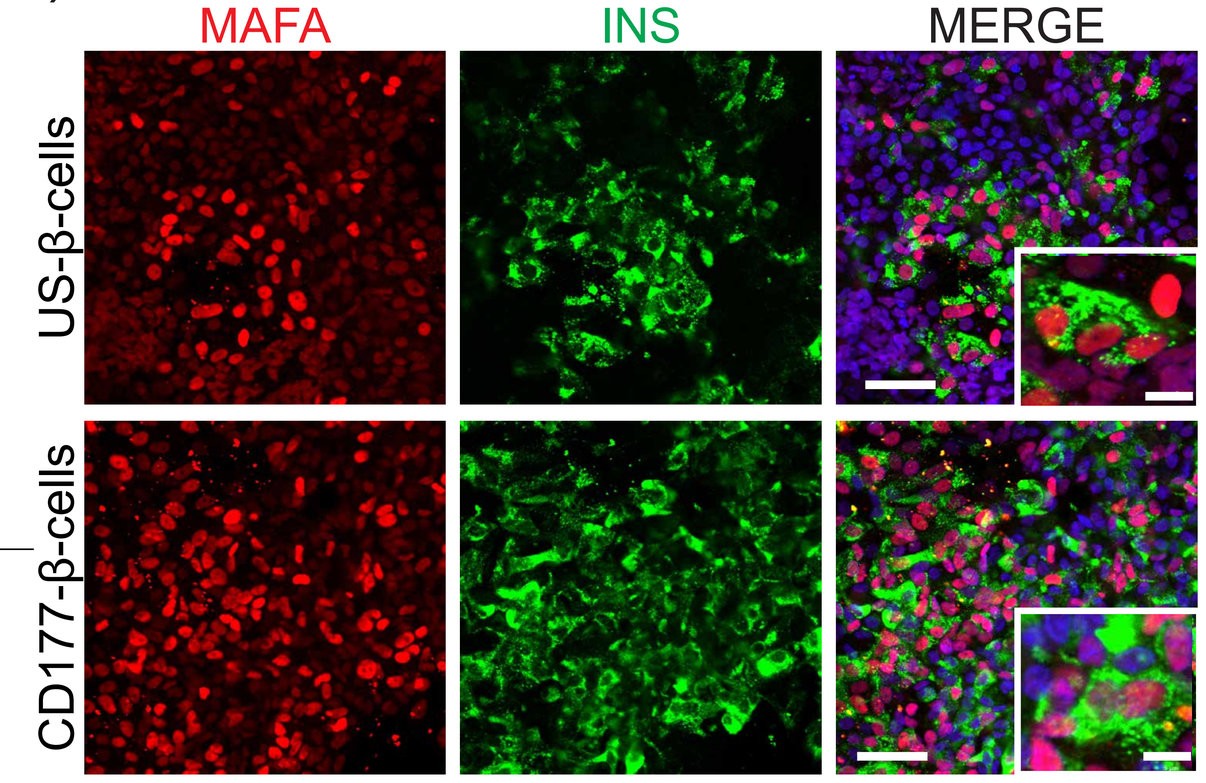
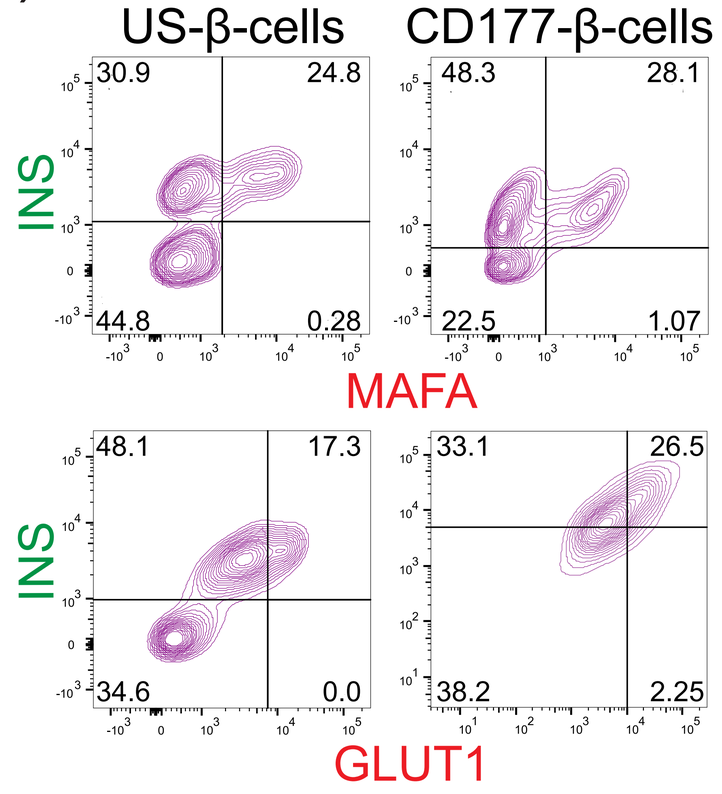


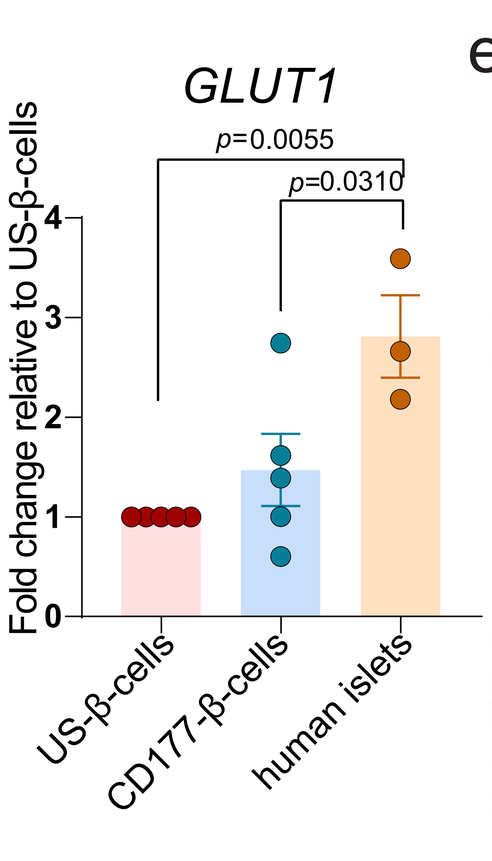
Figure 6. CD177⁺ ADE generates more mature β-like cells *in vitro*

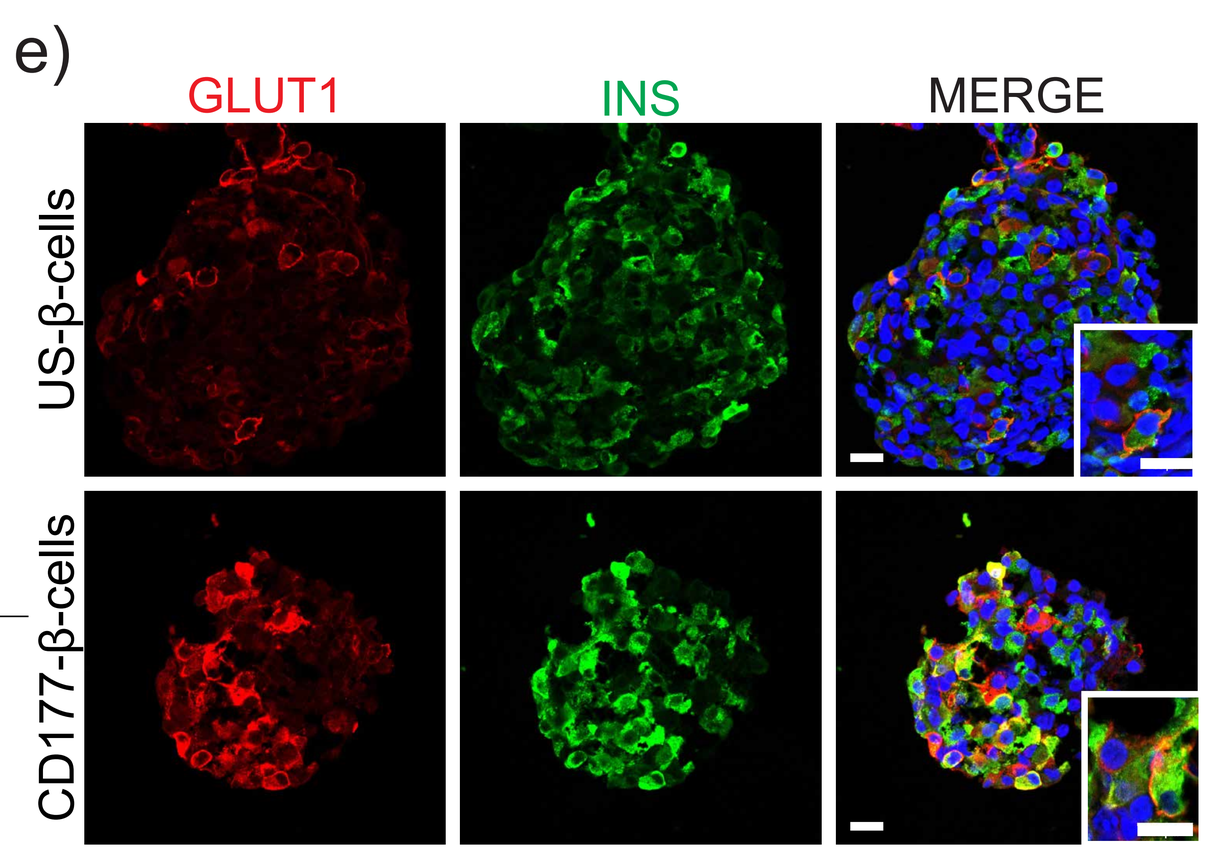












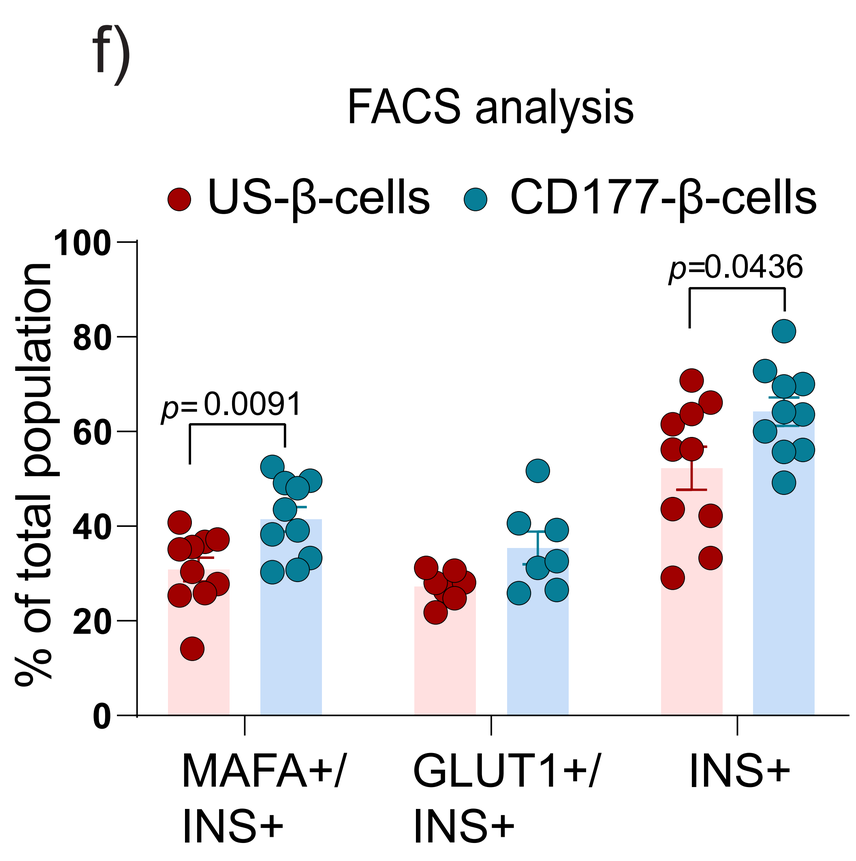
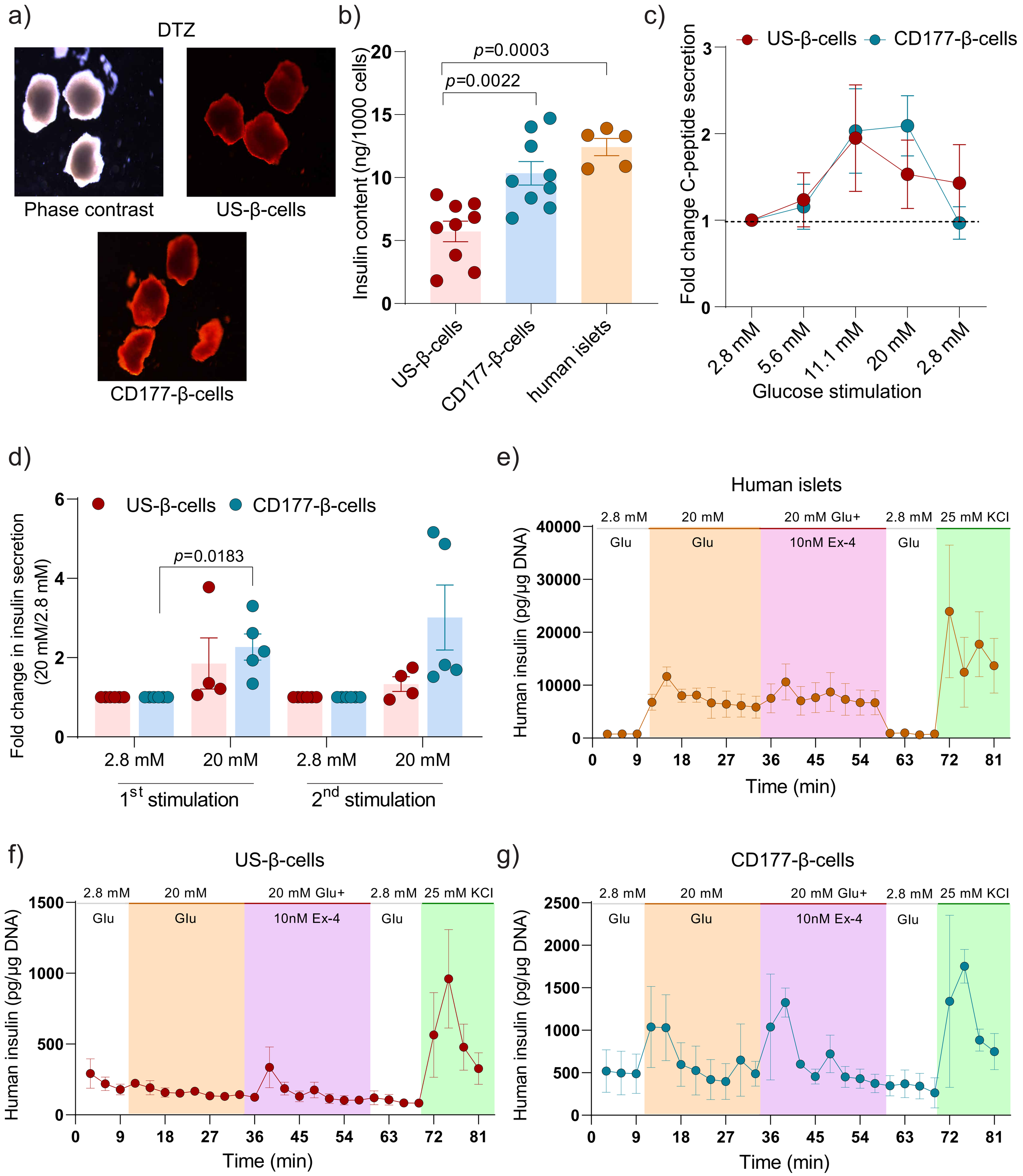
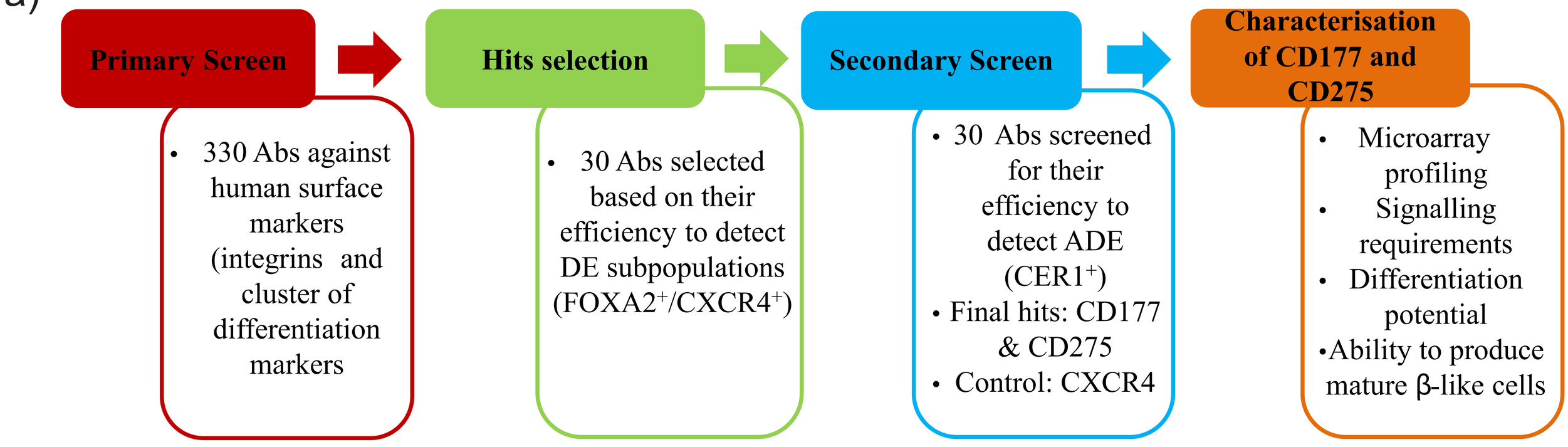
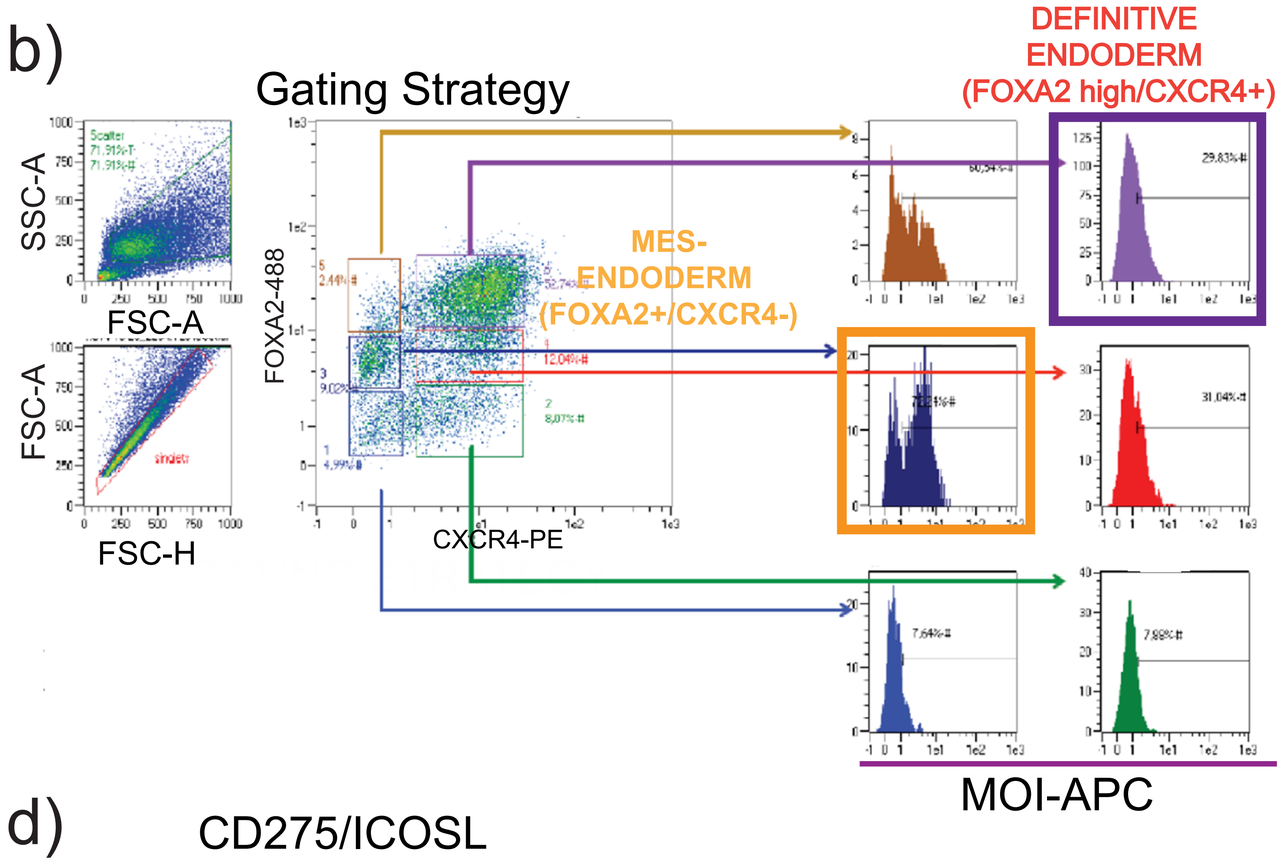


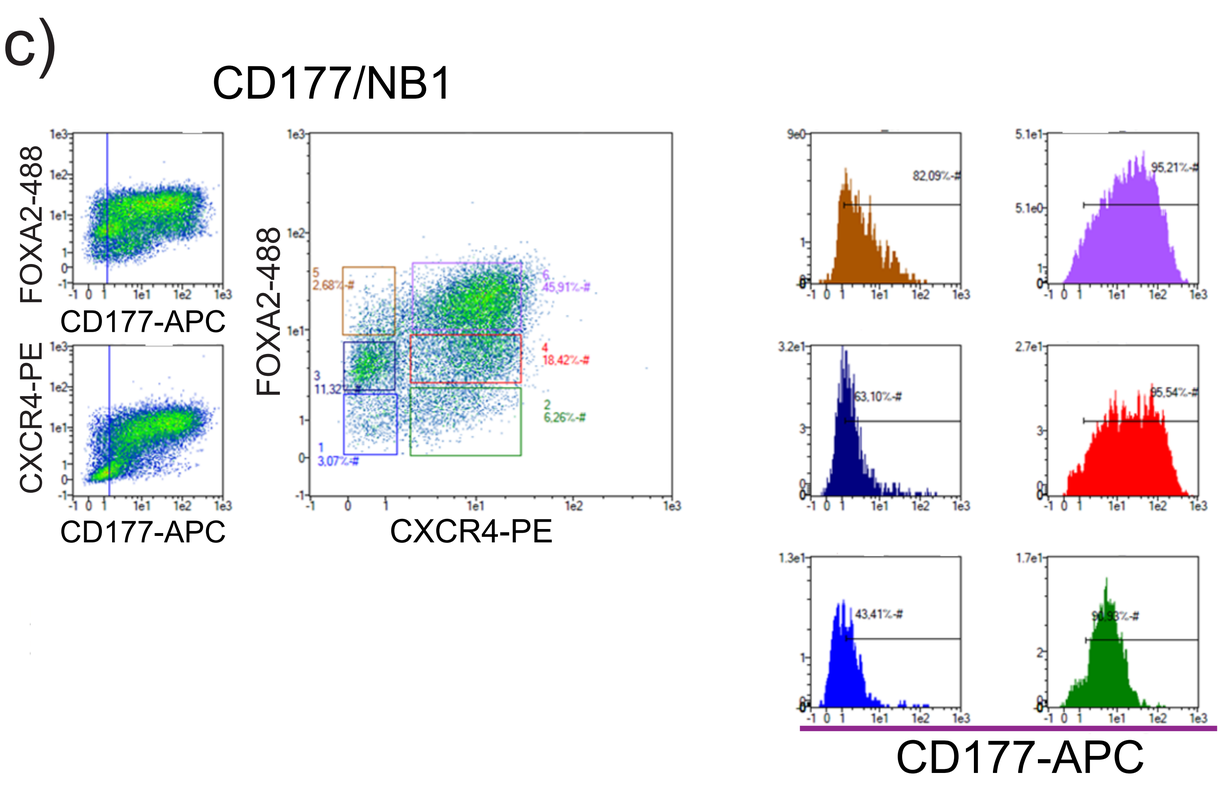
Figure 7. CD177⁺ ADE generates more functional β-like cells *in vitro*

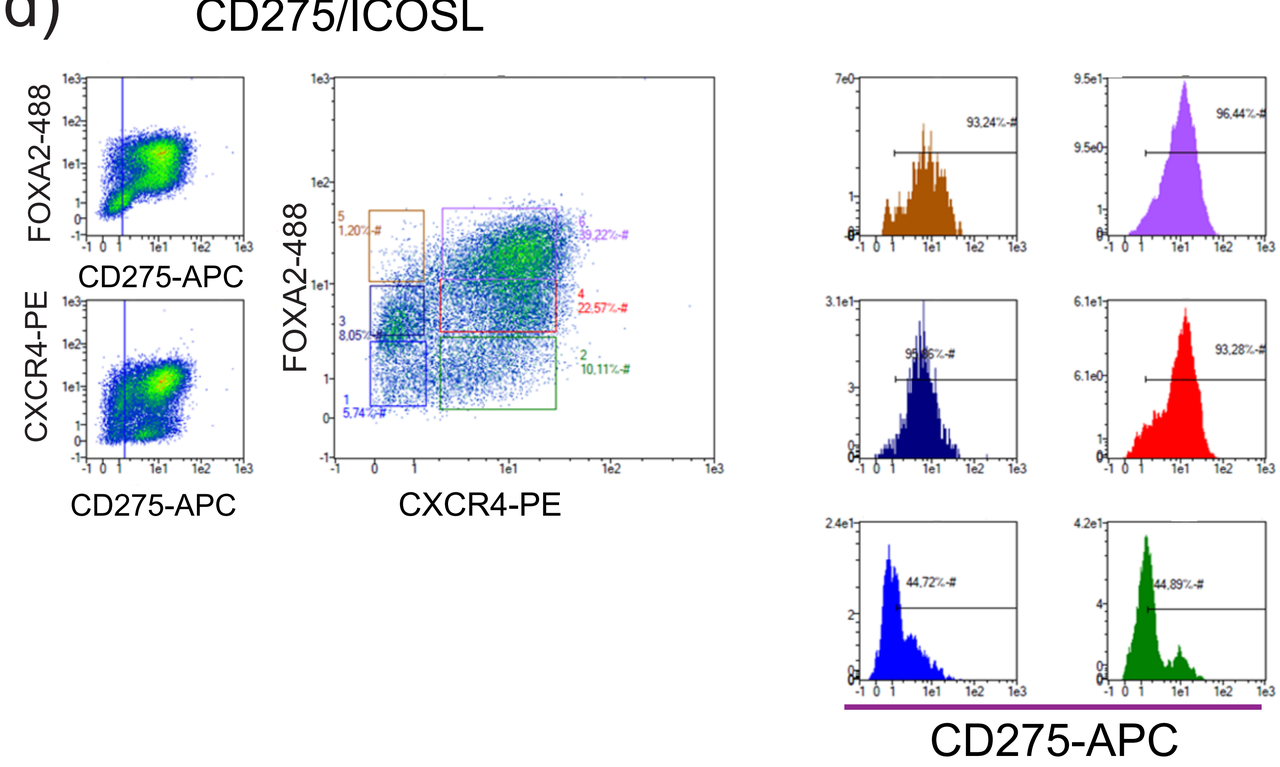


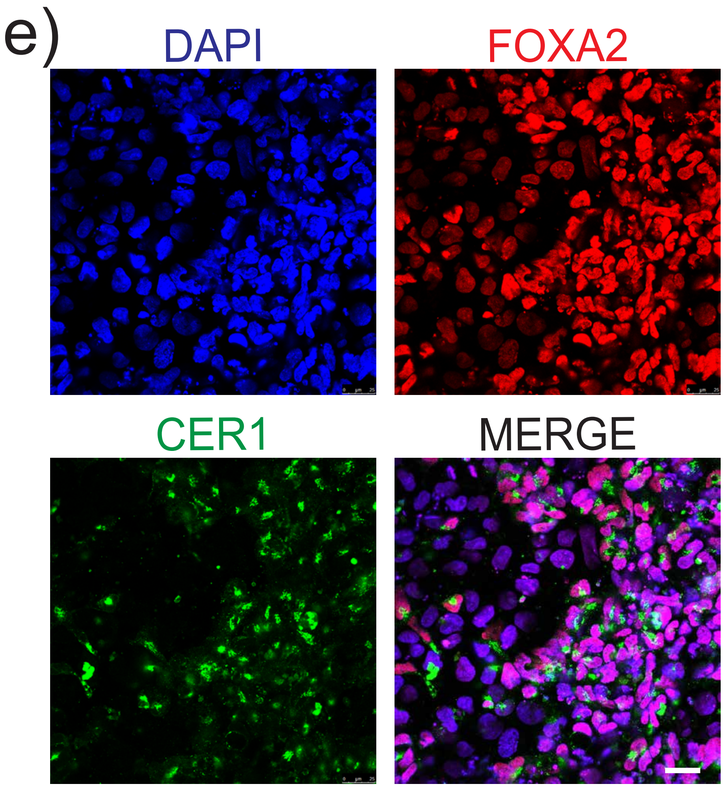
Supplementary Fig. 1. Screening strategy for the identification of endoderm subpopulations

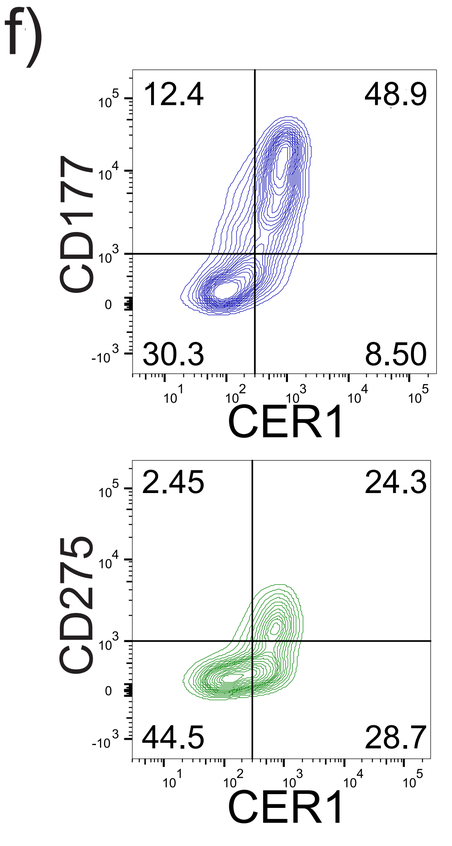


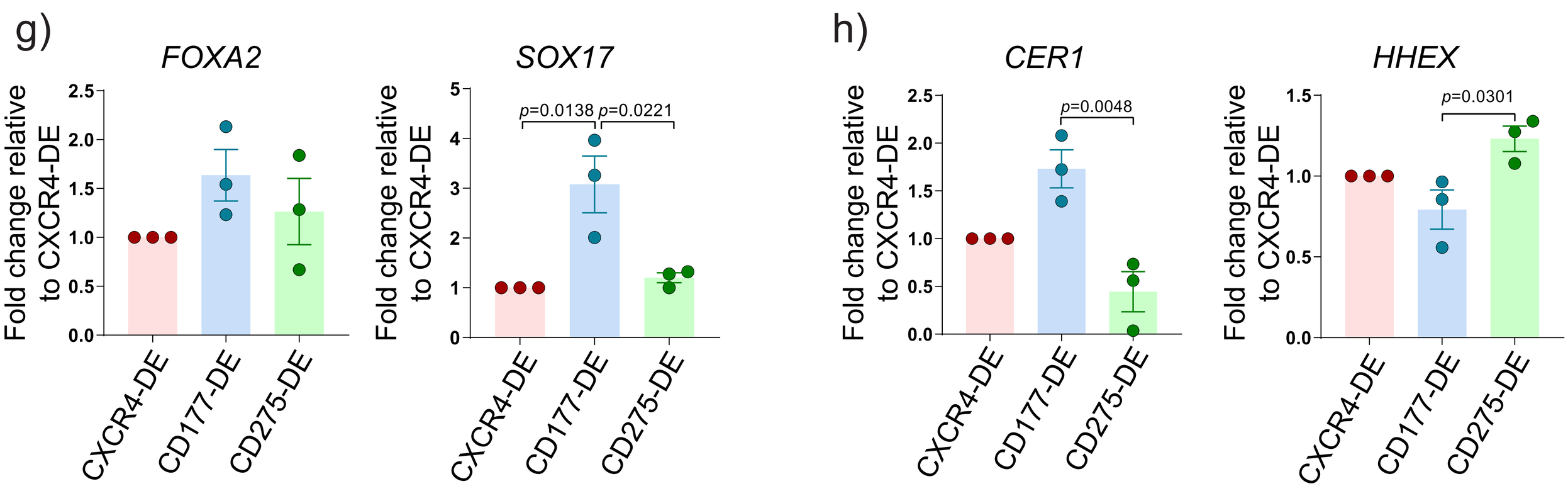




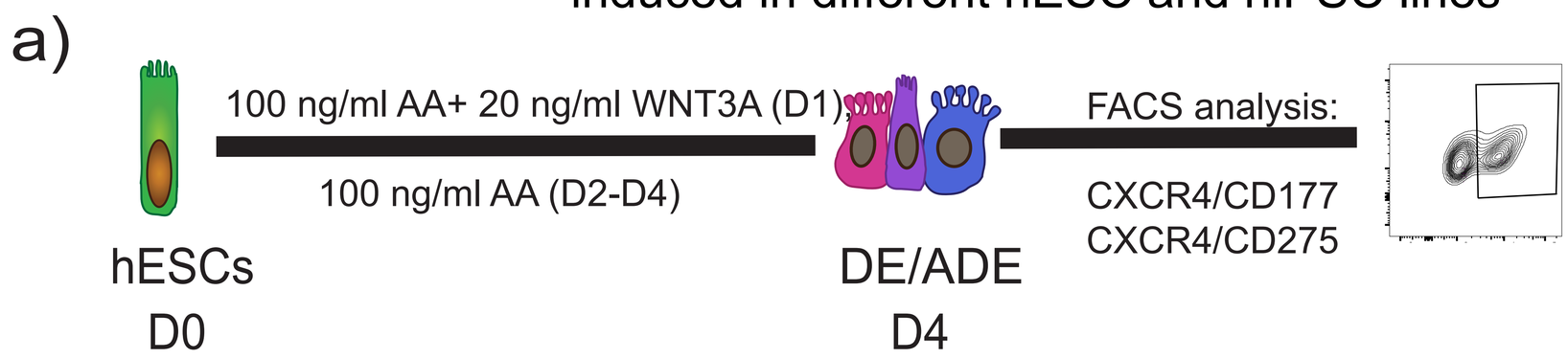


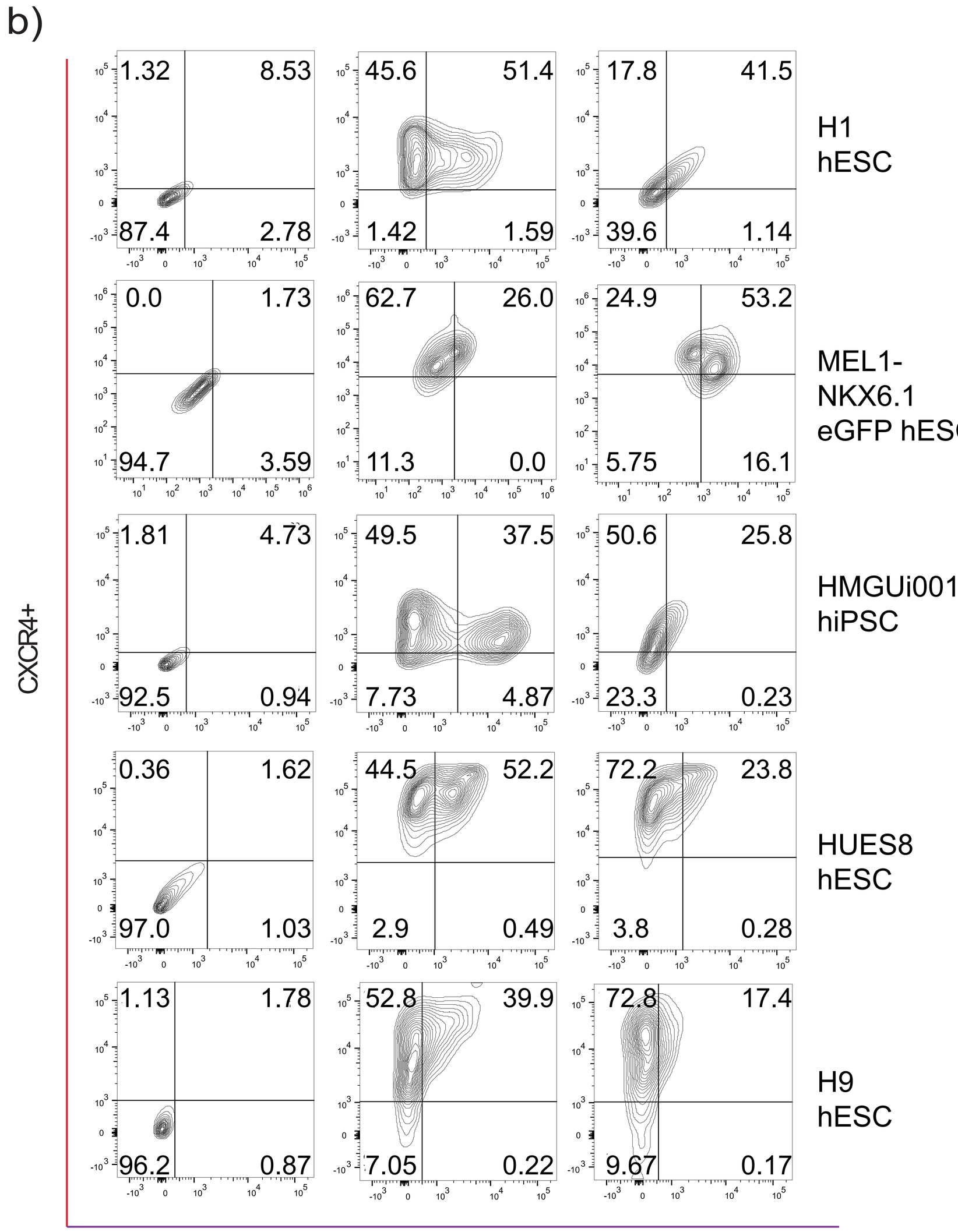


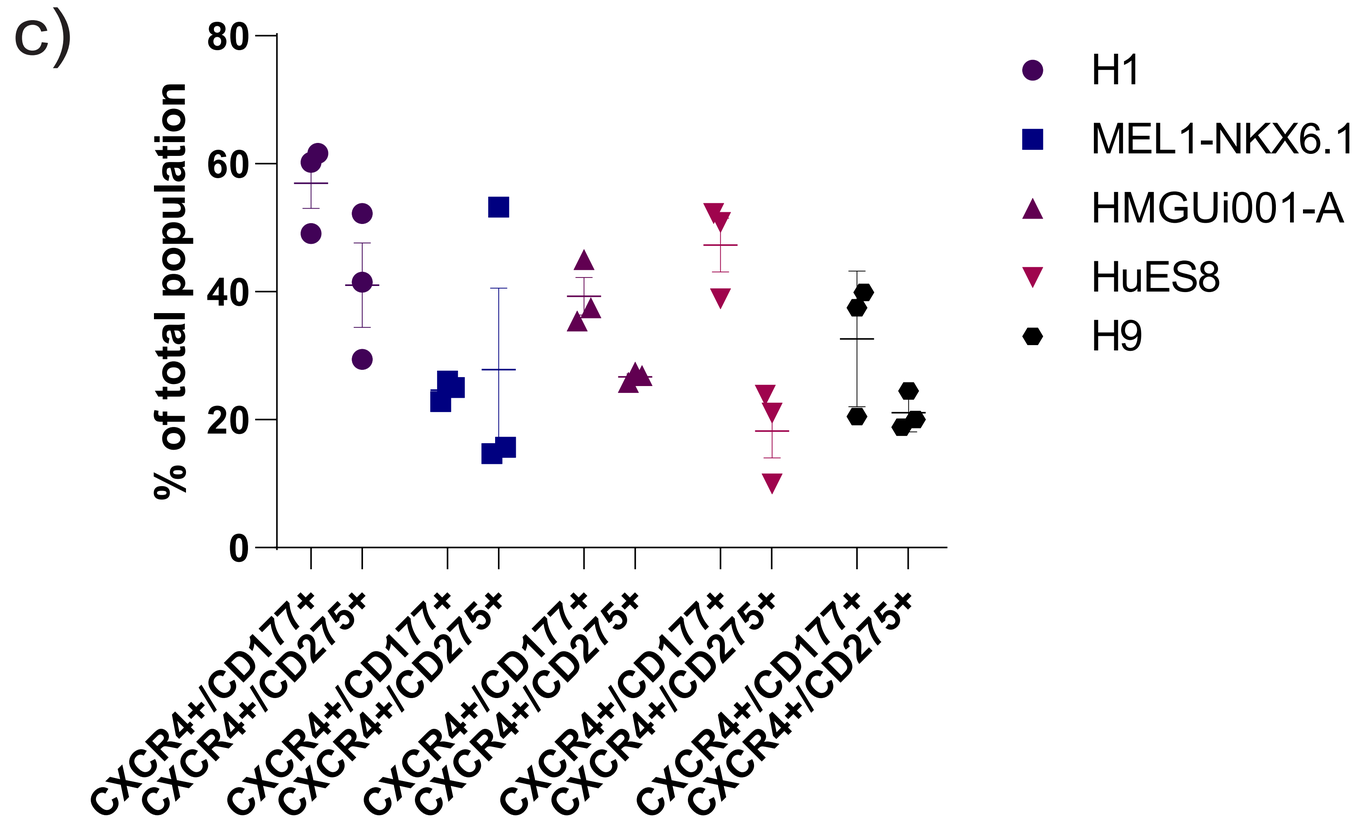




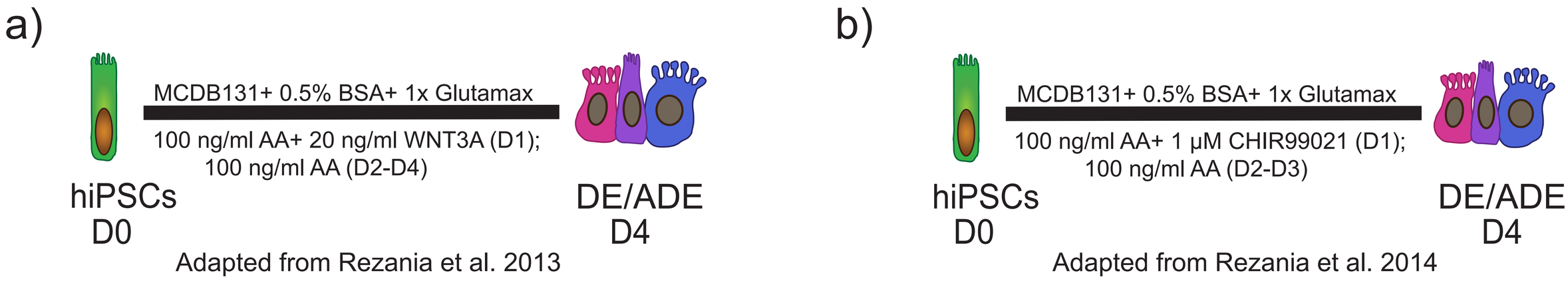
Supplementary Fig. 2. Percentage of CD177⁺ and CD275⁺ ADE subpopulations induced in different hESC and hiPSC lines

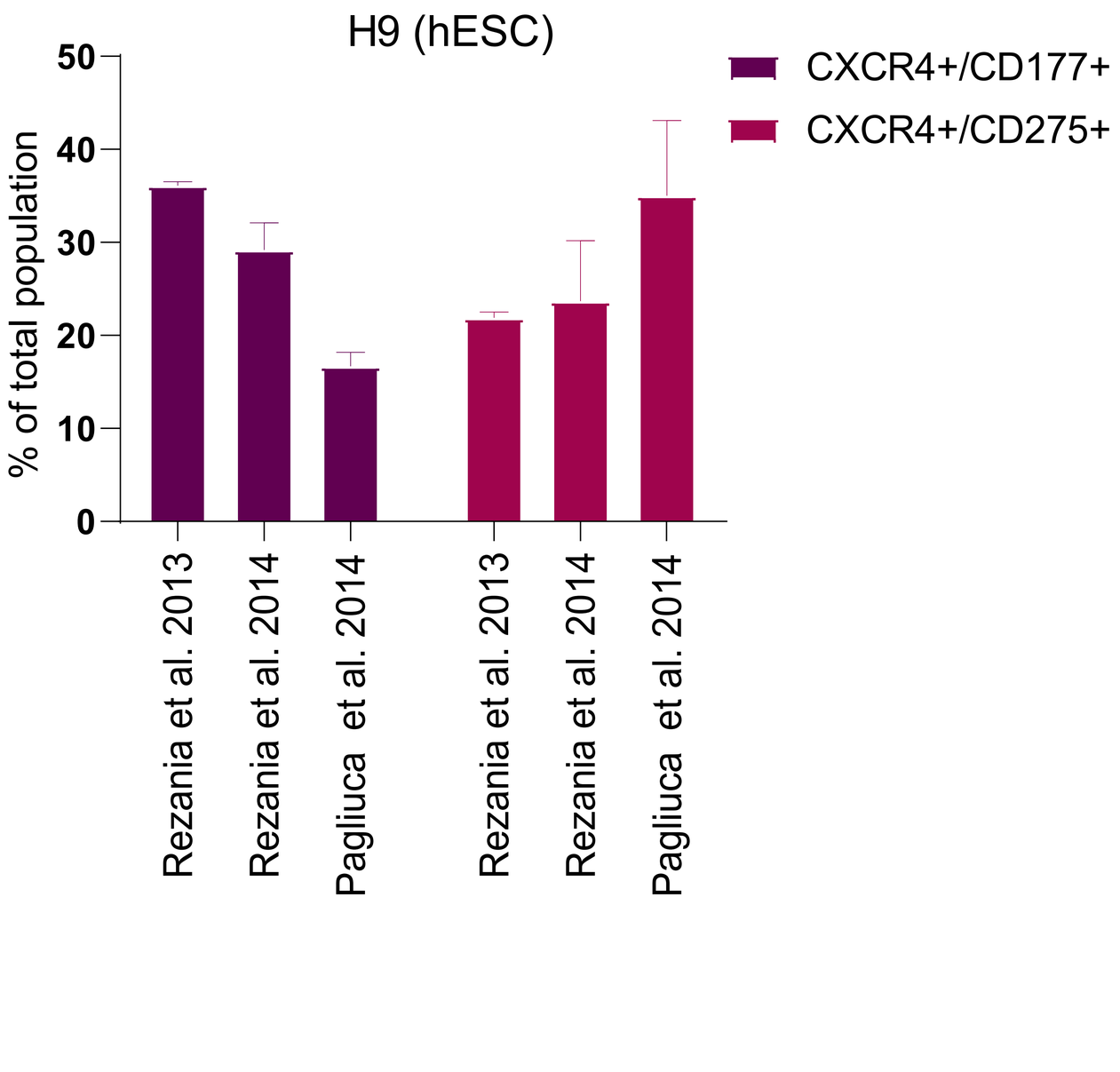


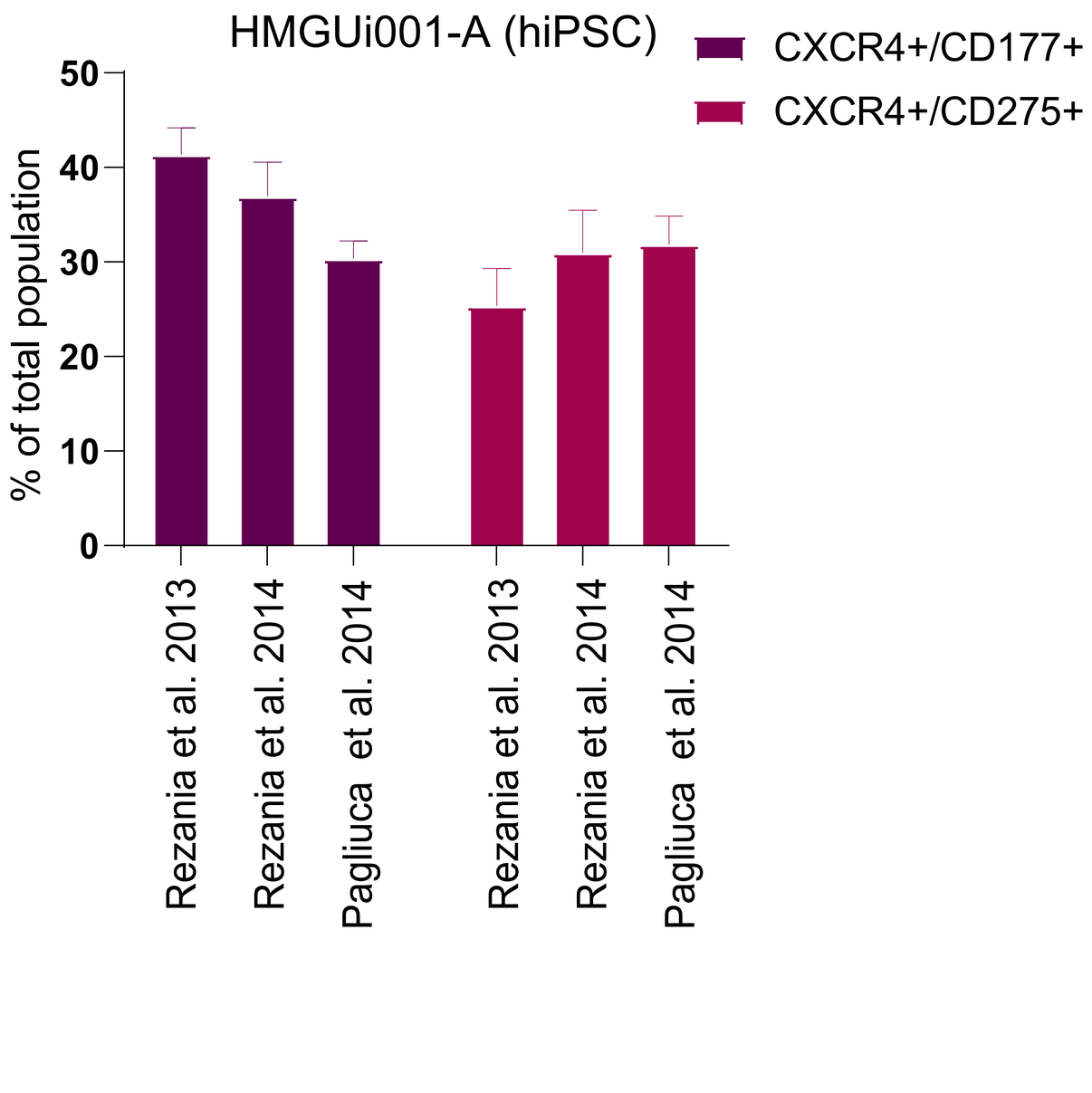




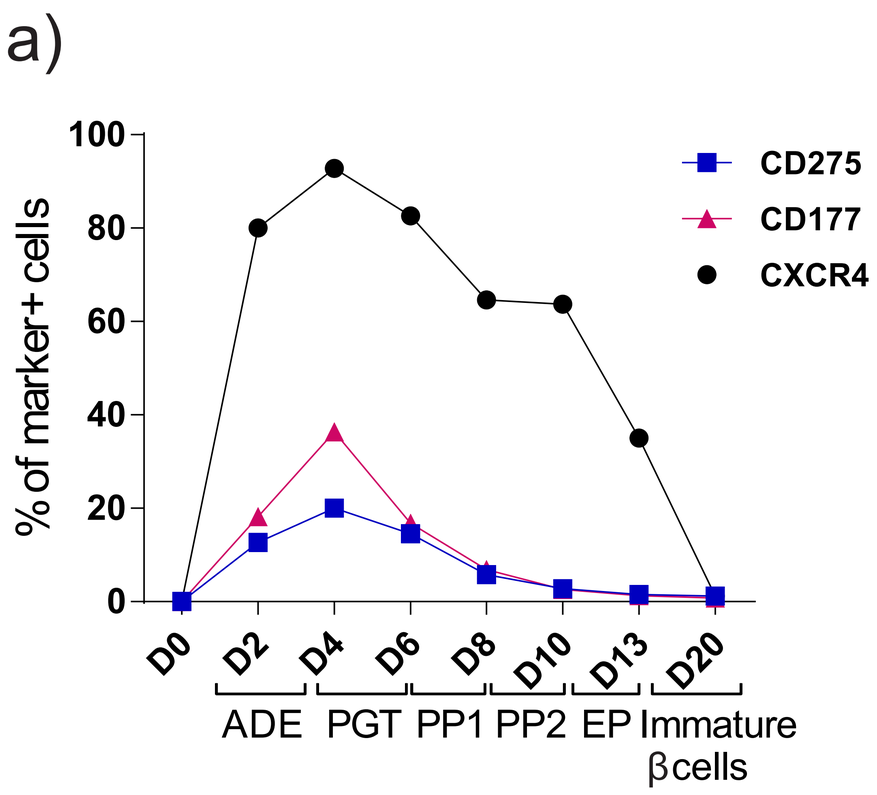
Supplementary Fig. 3. Induction efficiency of CD177⁺ and CD275⁺ ADE shows variation using different protocols

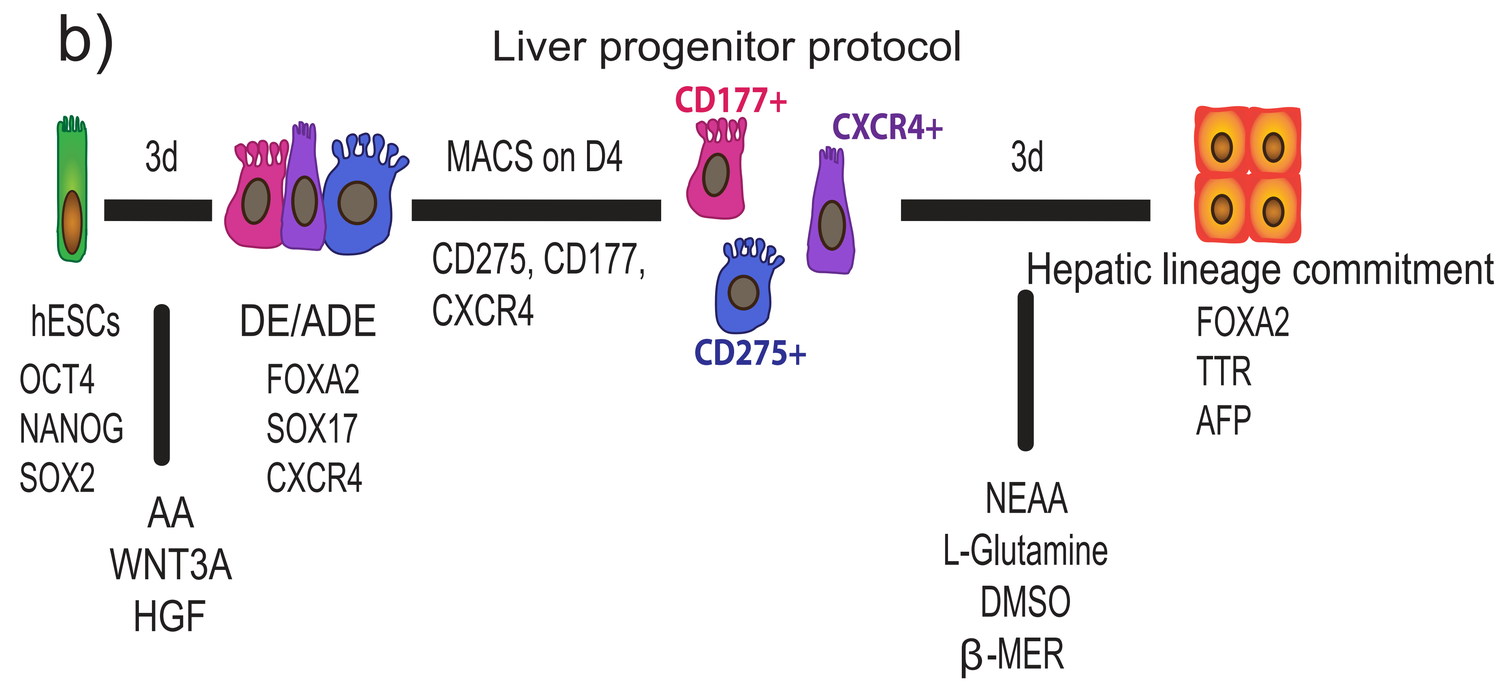


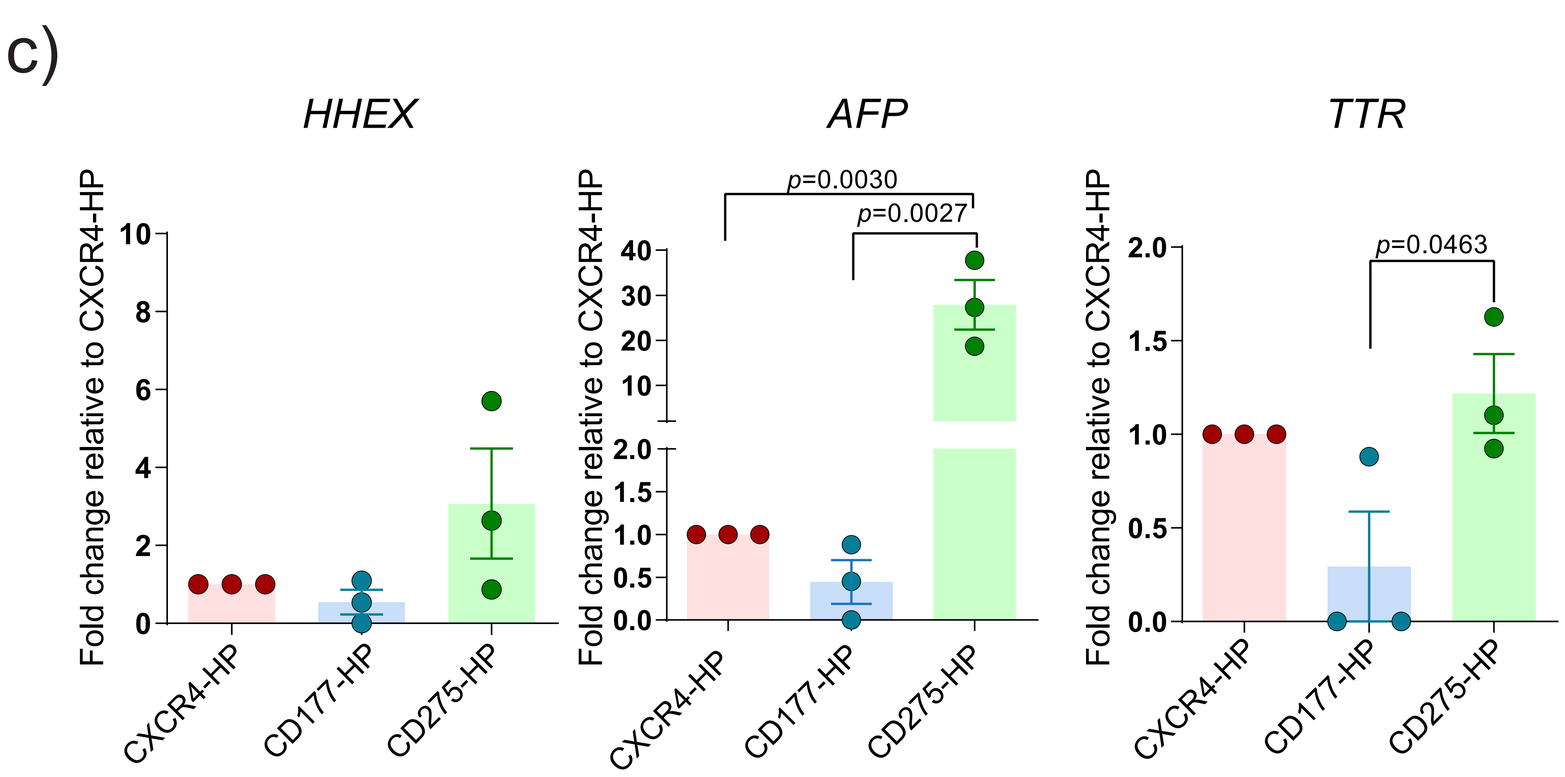


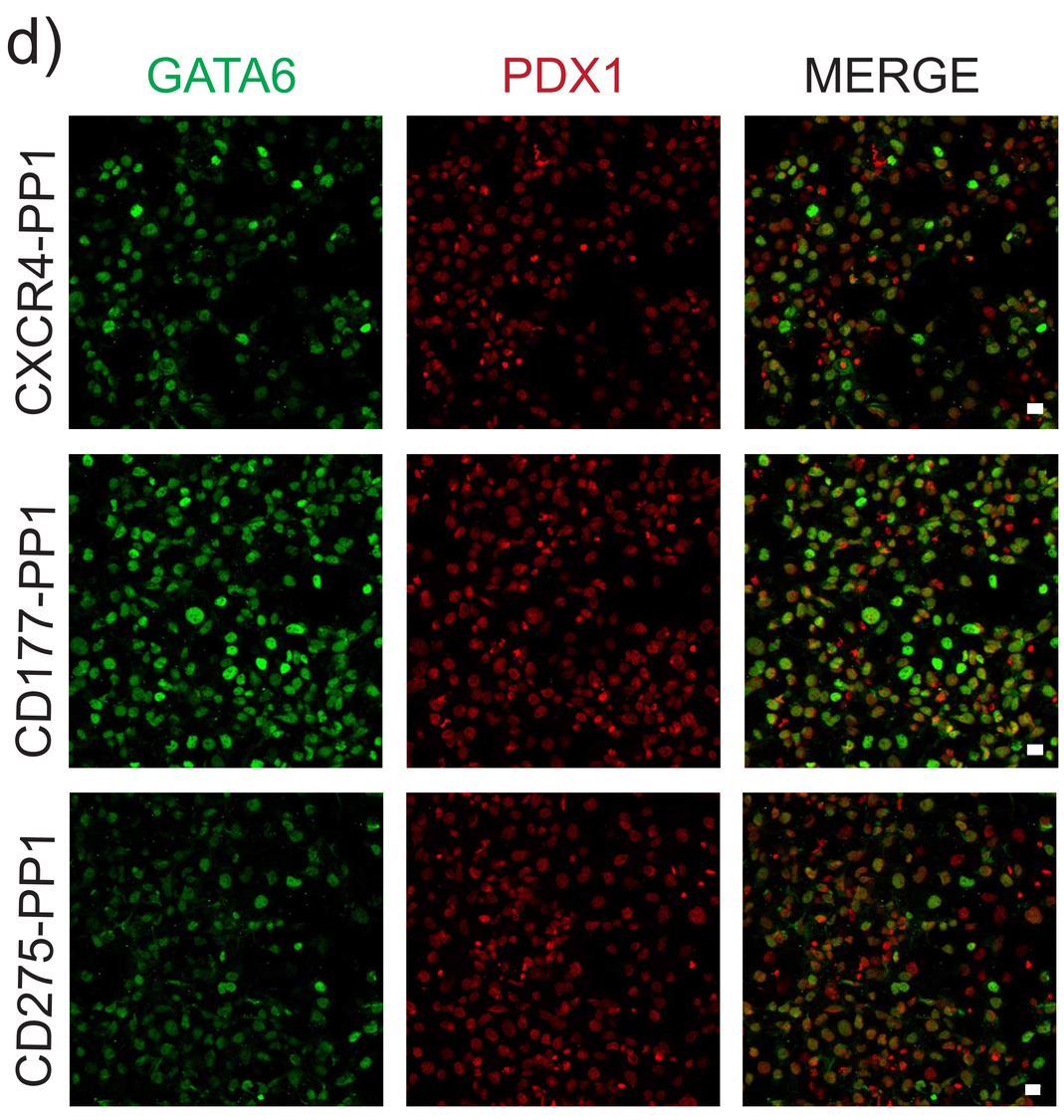


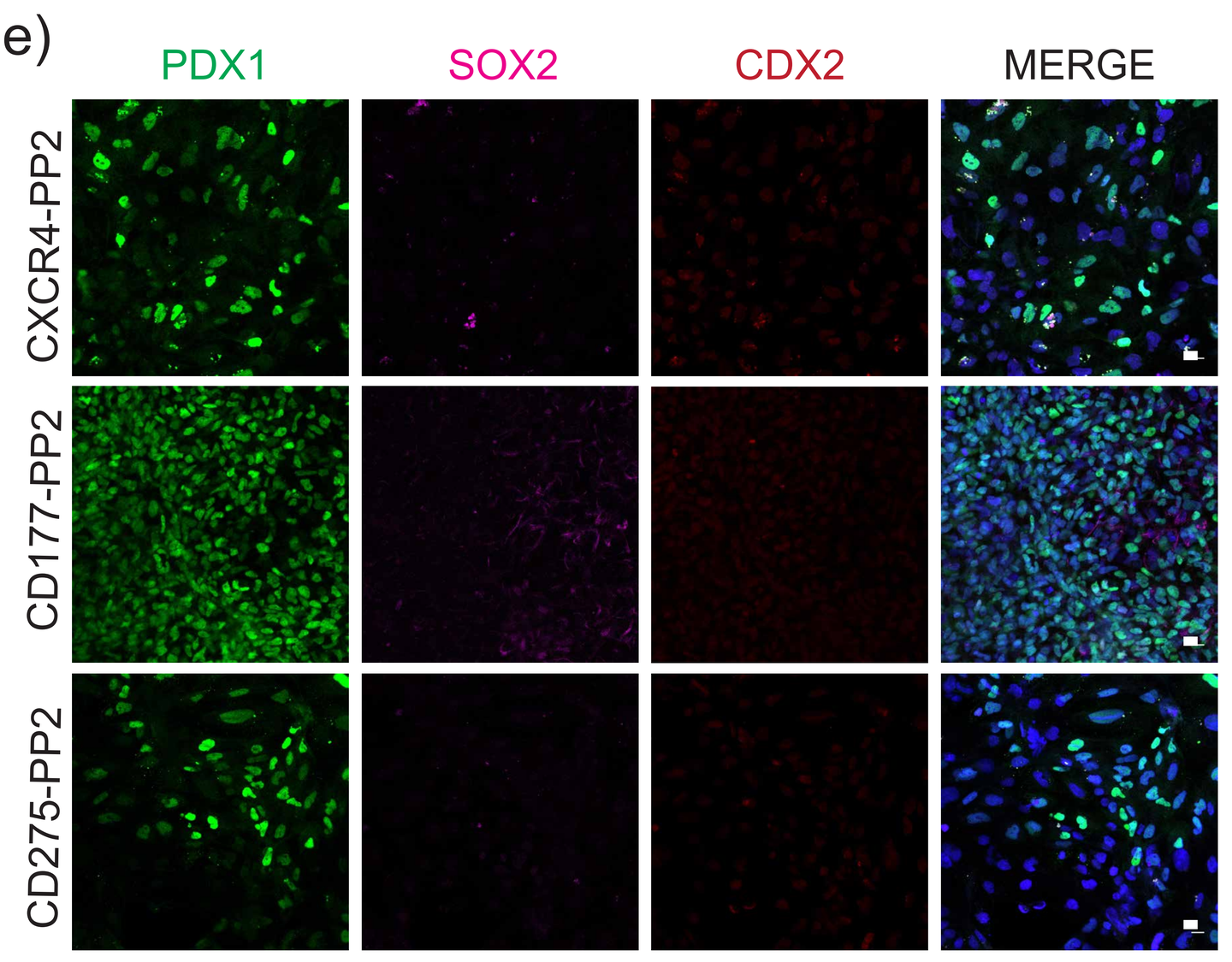
Supplementary Fig. 4. Differentiation of enriched CD177^{+} , CD275^{+} and CXCR4^{+} DE subpopulations towards liver fate



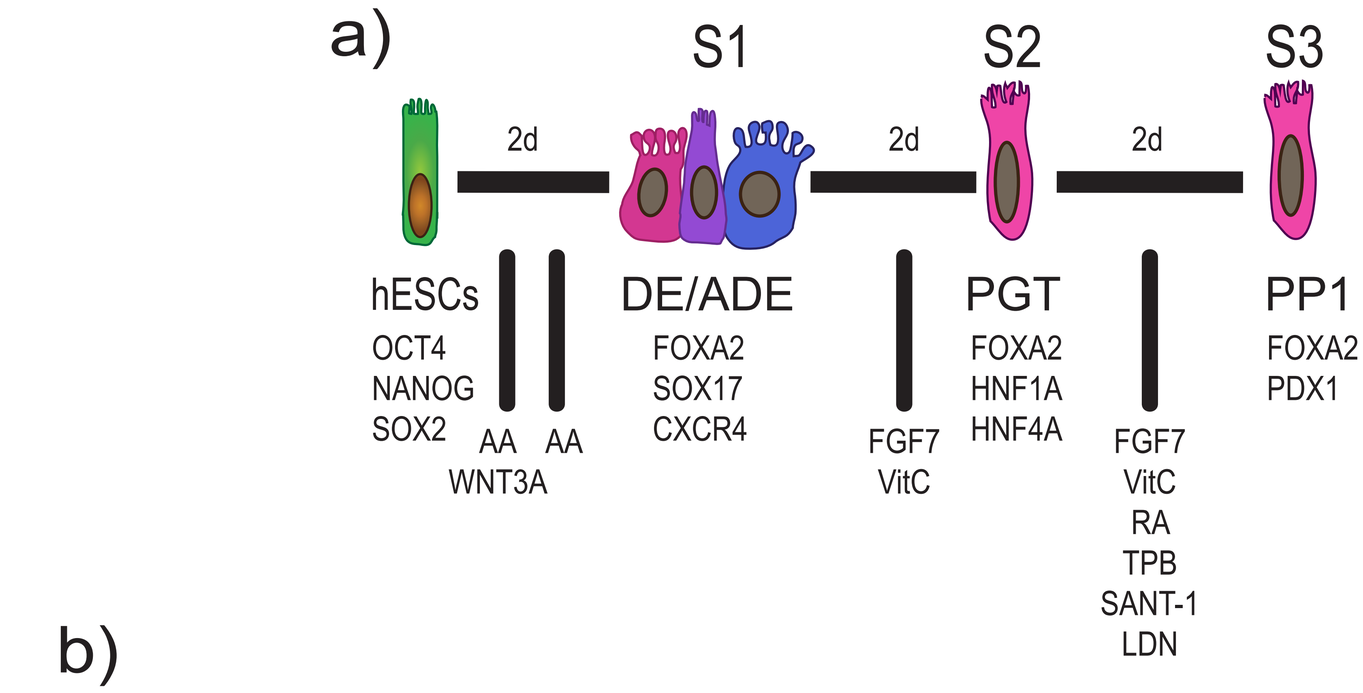


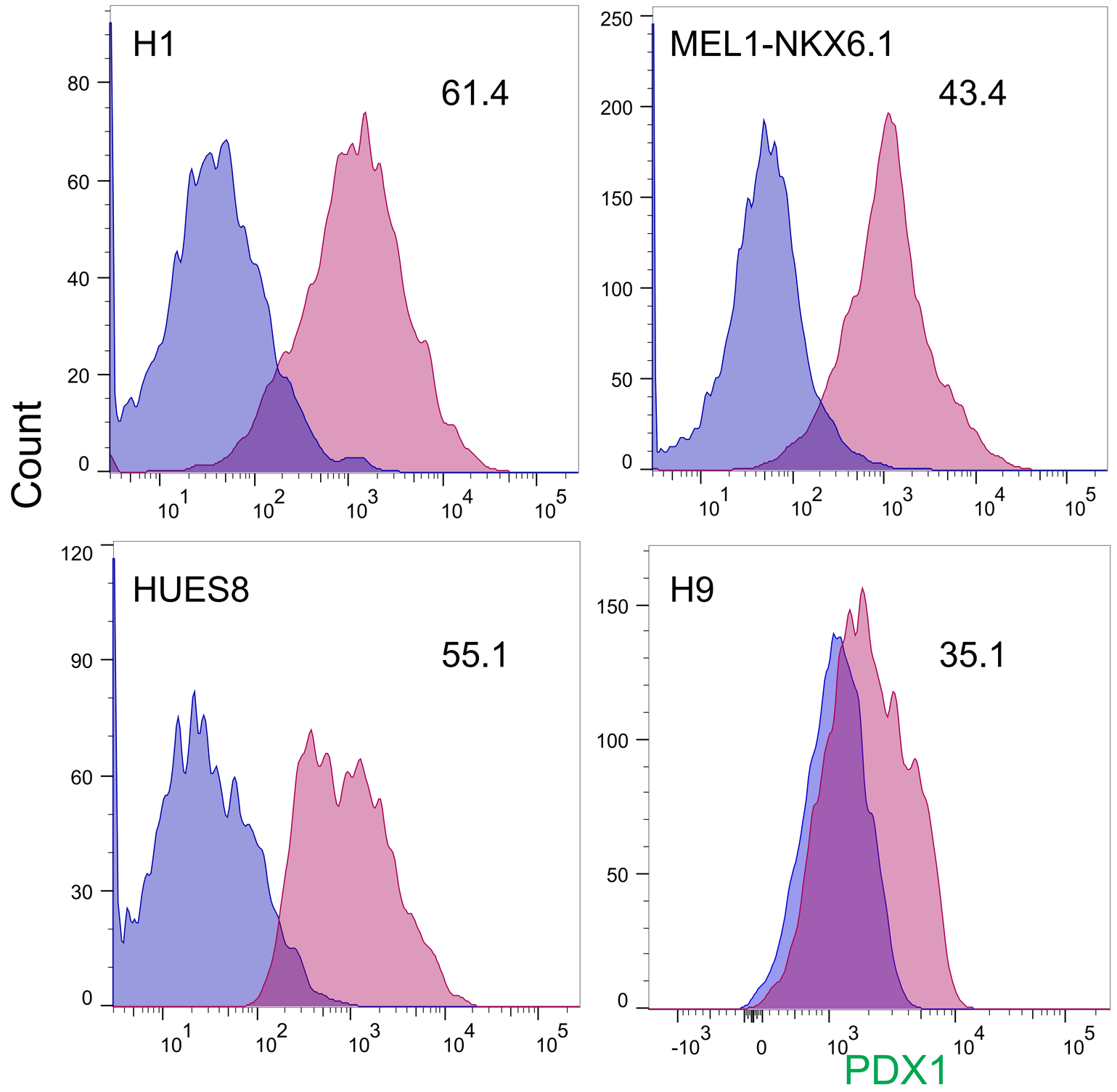


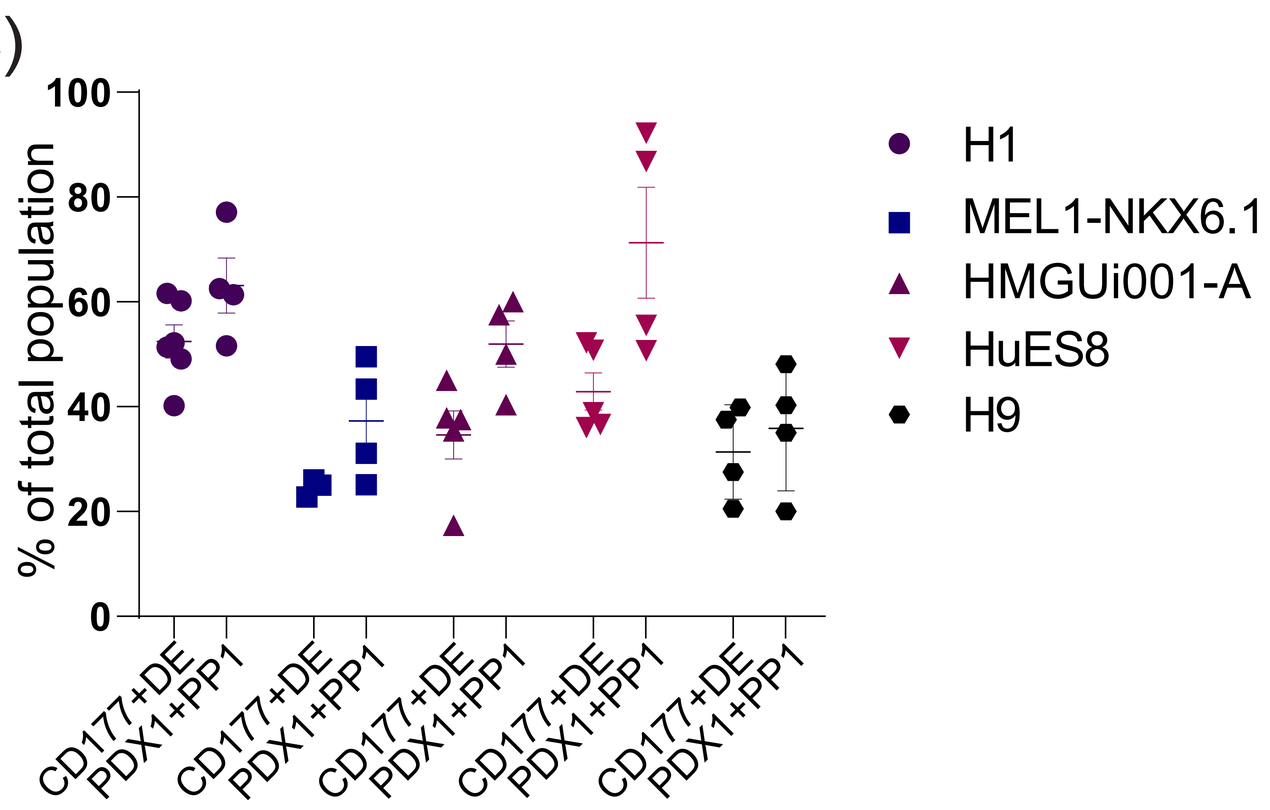




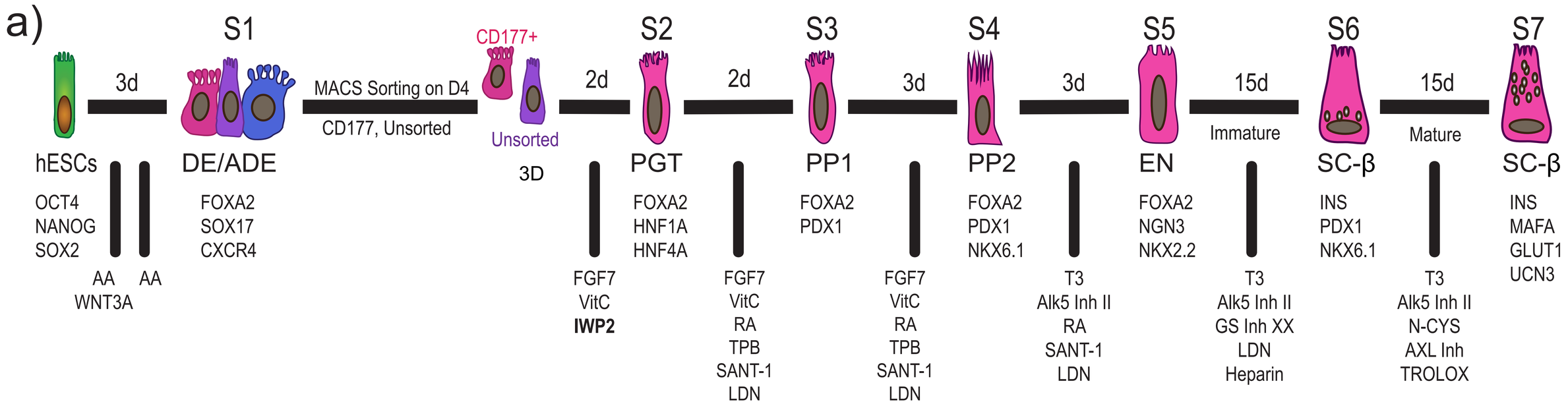
Supplementary Fig. 5.CD177^{+} ADE positively correlates with PP1 induction

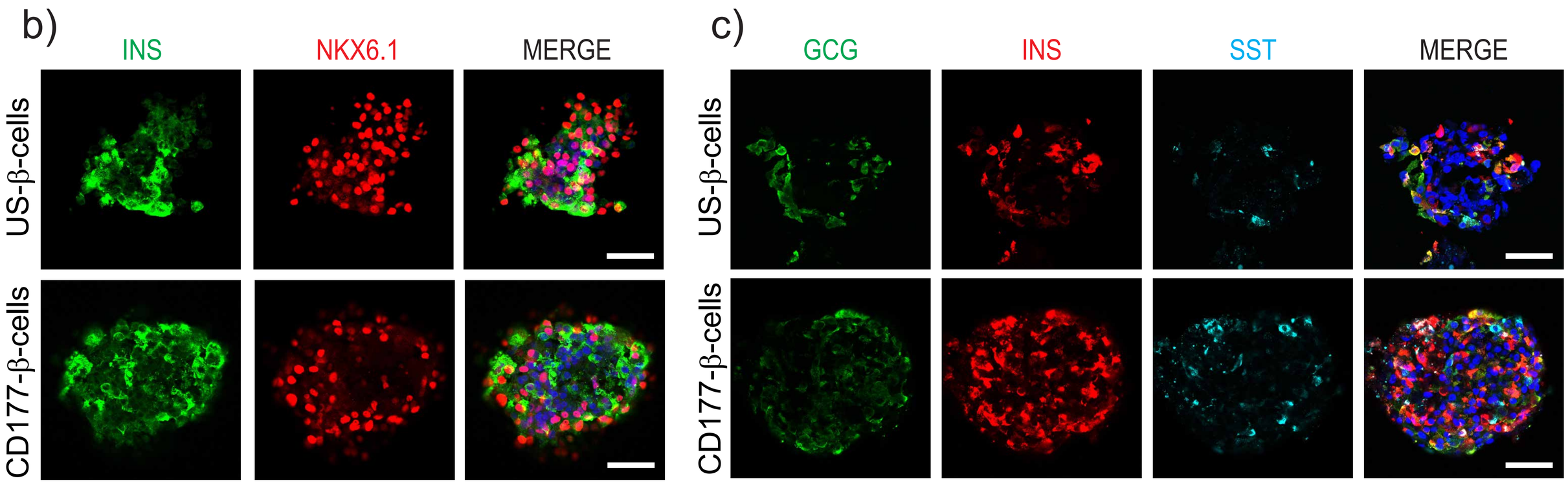


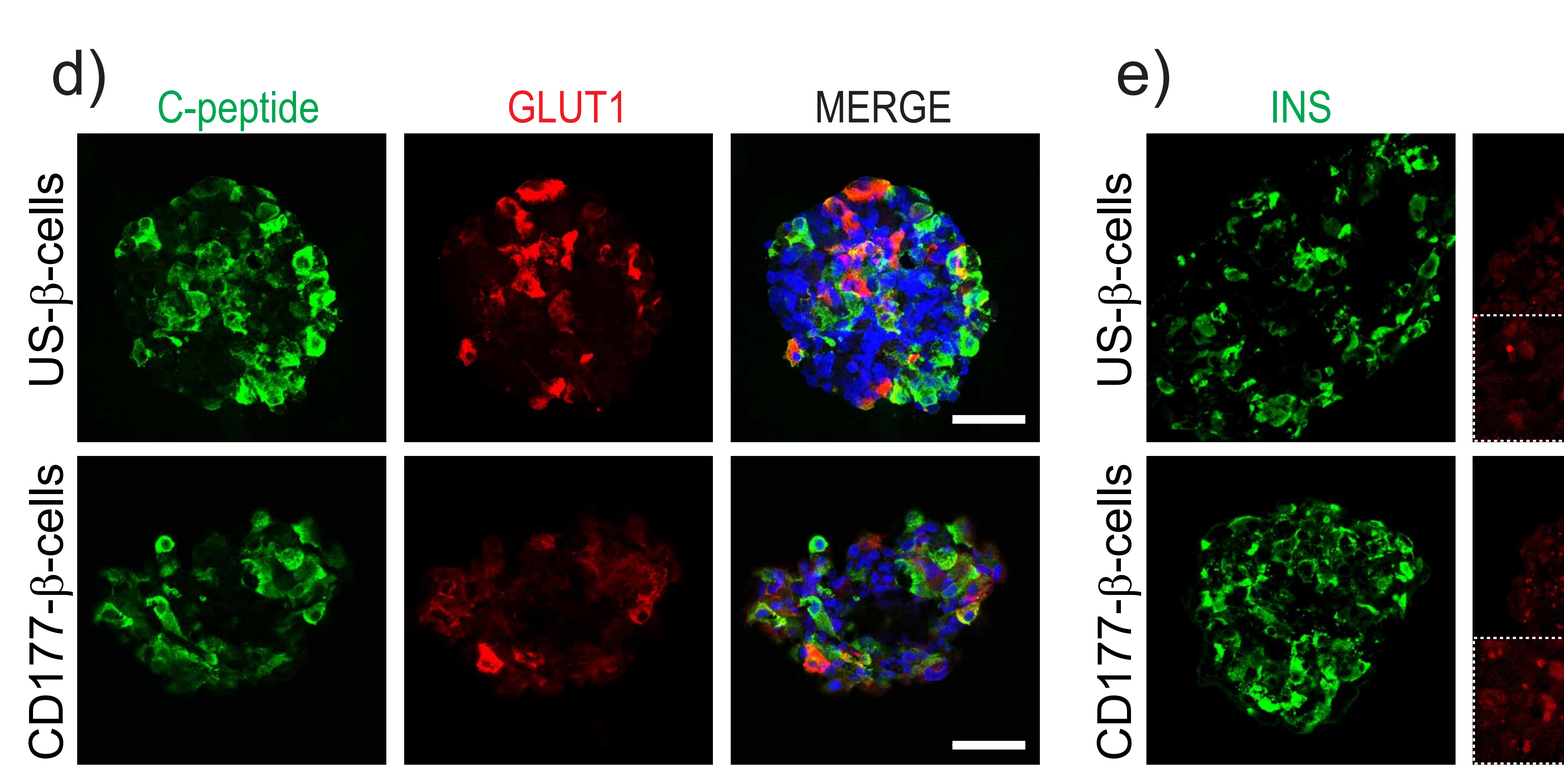


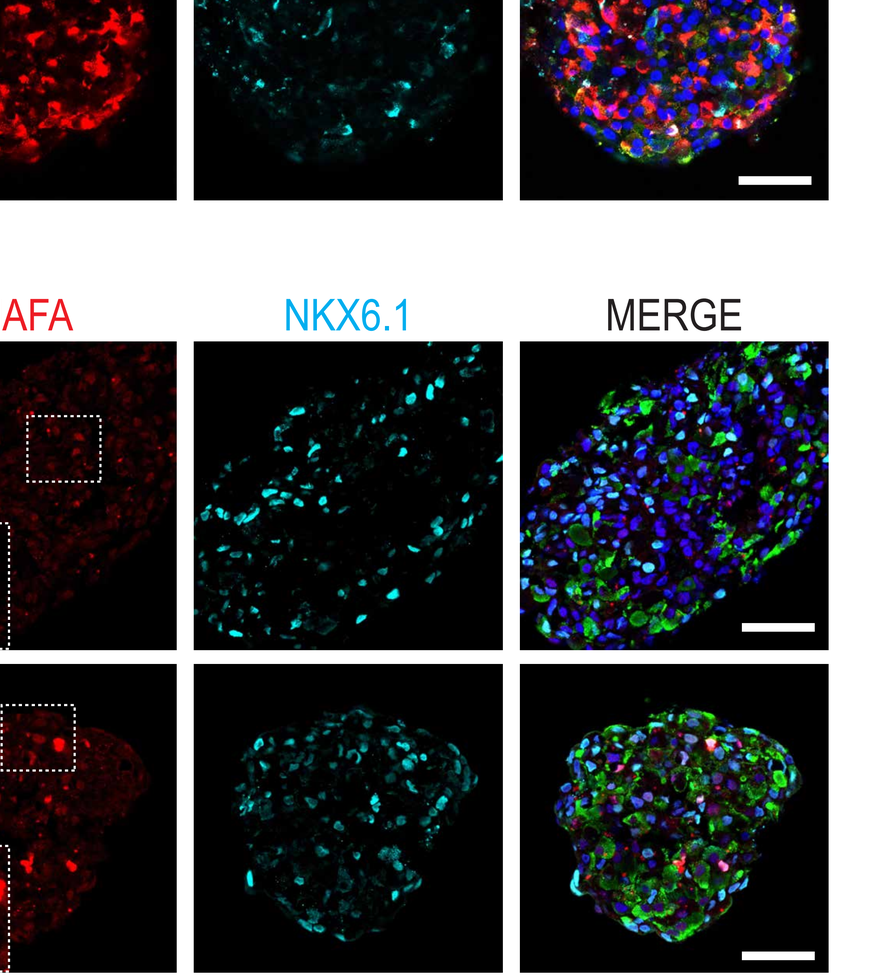


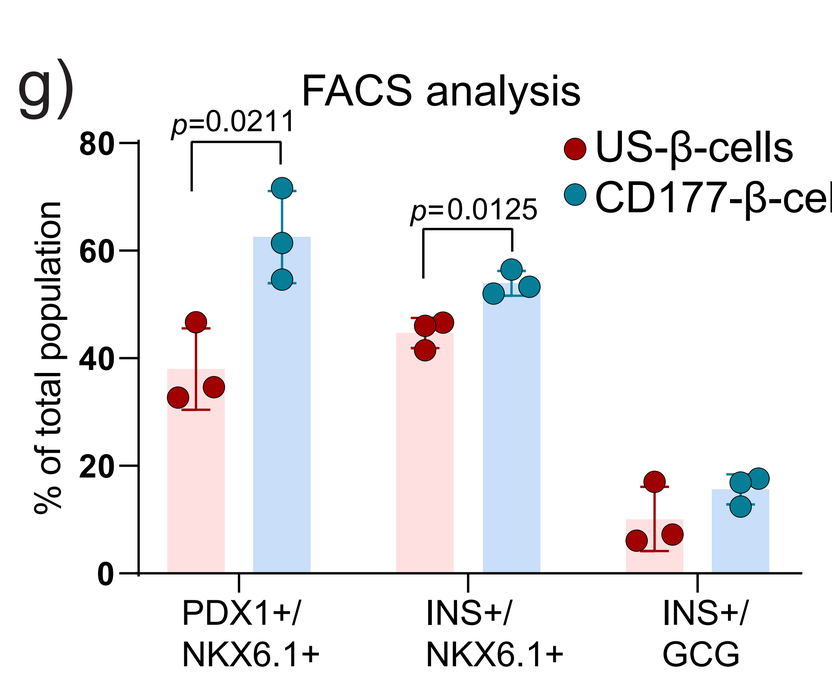
Supplementary Fig.6. H1 hESC pancreatic and endocrine differentiations in CD177⁺ ADE and US-DE



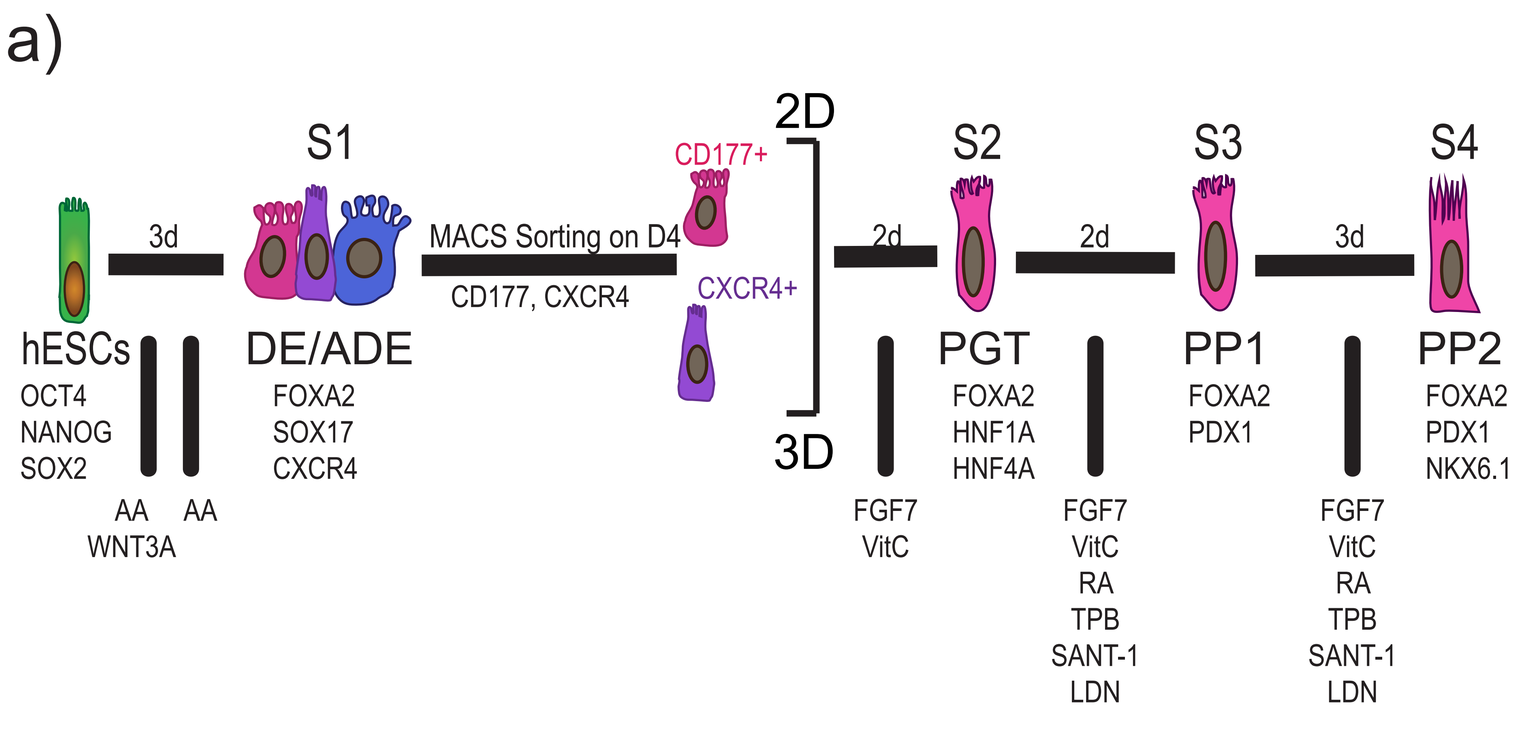


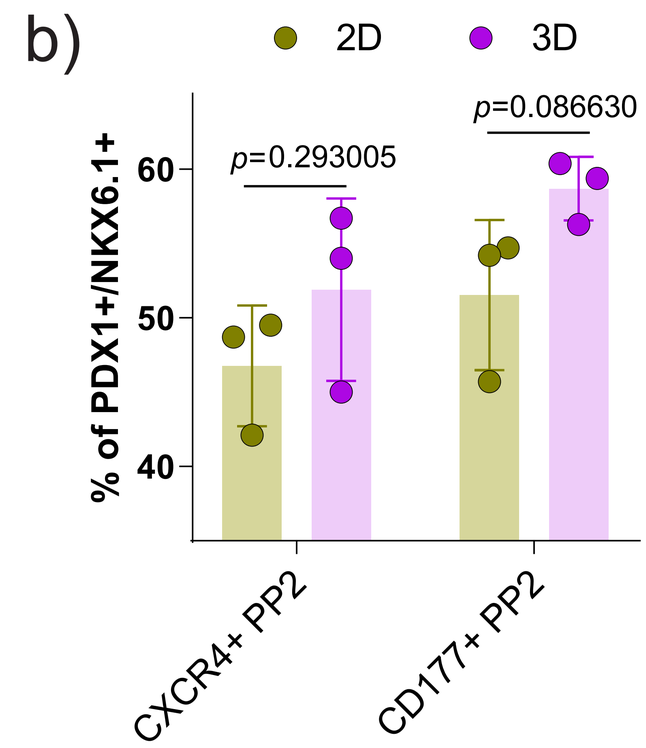


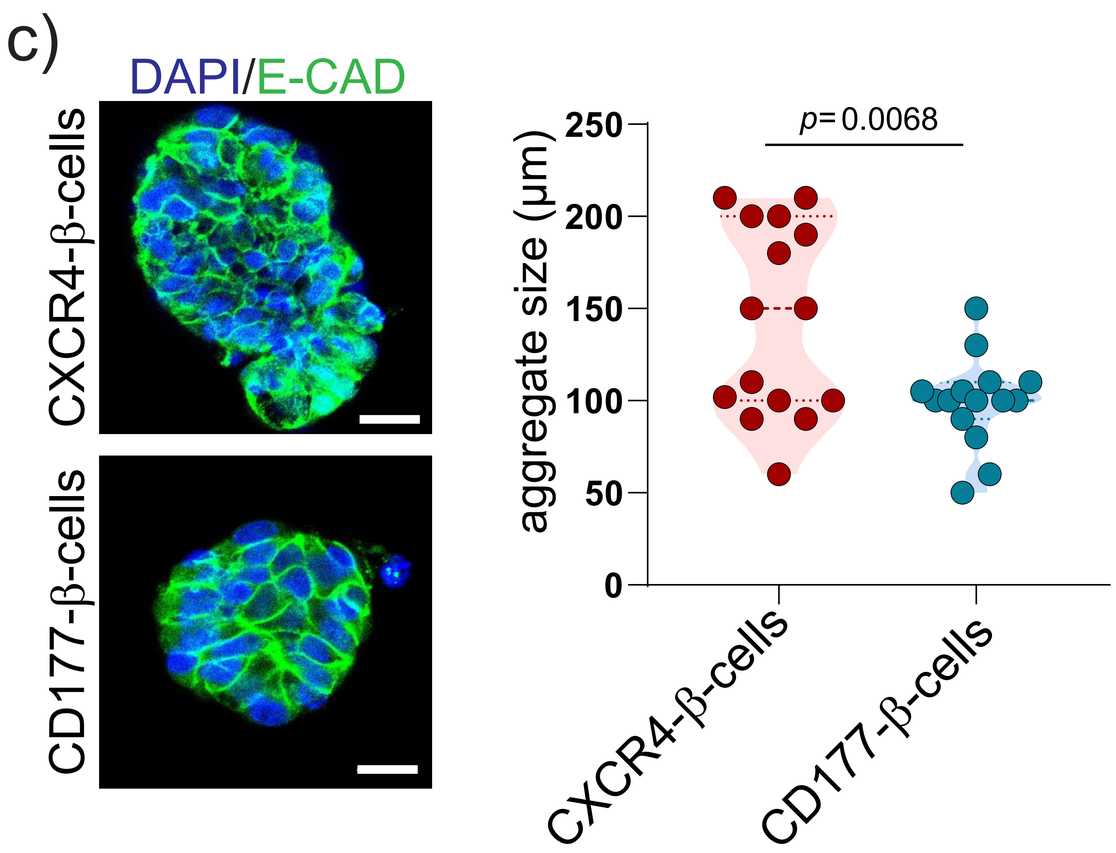




Supplementary Fig. 7. Comparison of 2D and 3D cell culture system on pancreatic differentiation







Supplementary Figure 8. H1 hESC CD177⁺ ADE generates more functional β-like cells *in vitro*