Studies on removal of off-target cells for clinical application of pluripotent stem cell-derived pancreatic islet cells

(多能性幹細胞由来の膵島様細胞の臨床応用に向けた
目的外細胞除去に関する研究)

Hideyuki Hiyoshi

Contents

Abbreviations

Cells

iPSCs	induced pluripotent stem cells
ESCs	embryonic stem cells
iPICs	iPSC-derived pancreatic islet cells
s6-iPICs	six-stage iPICs
s7-iPICs	seven-stage iPICs
MSCs	mesenchymal stem cells
SMCs	smooth muscle cells
PMSCs	proliferative MSC- and SMC-like cells

Proteins

AXL	anexelekto (a receptor tyrosine kinase)
BSA	bovine serum albumin
CHGA	chromogranin A
CK19	Cytokeratin 19
EGF	epidermal growth factor
FGFR	fibroblast growth factor receptor
GCG	glucagon
HER	human epidermal growth factor receptor
IGF1	insulin like growth factor 1
INS	insulin
KGF	keratinocyte growth factor
LDHB	lactate dehydrogenase B
NKX6.1	NK6 homeobox 1
PDX1	pancreatic and duodenal homeobox 1
PLK	polo-like kinase
PP	pancreatic polypeptide
SFRP4	secreted frizzled related protein 4
YAP	yes-associated protein
bFGF	basic fibroblast growth factor
α-SMA	α-smooth muscle actin

Cell culture-related

ALK5i	activin receptor-like kinase 5 inhibitor II
Act A	activin A
CHIR	CHIR99021
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetra-acetic acid
GSI	γ-secretase inhibitor
HEPES	piperazineethanesulfonic acid
ITS-X	insulin-transferrin-selenium-ethanolamine
K-Cyc	3-keto-N-aminoethyl-N'-aminocaproyldihydrocinnamoyl cyclopamine
LDN	LDN-193189
$MgCl_2$	magnesium chloride
N-Cys	N-acetyl cysteine
$NaHCO_3$	sodium hydrogen carbonate
Nic	nicotinamide
P/S	penicillin/streptomycin
RA	retinoic acid
ROCKi	Rho-associated coiled-coil kinase inhibitor
T3	L-3,3′,5-triiodothyronine
TTNPB	4-[(E)-2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2- naphthalenyl)-1-propenyl] benzoic acid
$ZnSO_4$	Zinc sulfate
iMEM	improved minimum essential medium
2-DG	2-deoxy-D-glucose
2D	two-dimensional
3D	three-dimensional

Others

AAAKAC	Association for Assessment and Accreditation of Laboratory Animal Care International
ARRIVE	Animal Research: Reporting of In Vivo Experiments
DEGs	differentially expressed genes
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EC	enterochromaffin
ELISA	enzyme-linked immuno-sorbent assay
GO	gene ontology
HE	hematoxylin and eosin
HuN	human nucleus
NOD-scid	NOD.CB17-Prkdc-scid/J
OCT	optimal cutting temperature
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
RCA	reference component analysis
RNA	ribonucleic acid
SD	standard deviation
STG	streptozotocin
T1DM	type 1 diabetes mellitus
TR-FRET	time-resolved fluorescence resonance energy transfer
UMAP	uniform manifold approximation and projection
UMI	unique molecular identifier
cDNA	complementary DNA
mRNA	messenger RNA
qPCR	quantitative PCR
scRNA-seq	single-cell RNA sequencing
t-SNE	t-distributed stochastic neighbor embedding

Notes

Contents of this thesis have been published as follows:

1. Hiyoshi H, Sakuma K, Tsubooka-Yamazoe N, Asano S, Mochida T, Yamaura J, Konagaya S, Fujii R, Matsumoto H, Ito R, Toyoda T. Characterization and reduction of non-endocrine cells accompanying islet-like endocrine cells differentiated from human iPSC. *Scientific reports*, 12 (1), 4740, 2022

2. Hiyoshi H, Sakuma K, Asano S, Napier S, Konagaya S, Mochida T, Ueno H, Watanabe T, Kassai Y, Matsumoto H, Ito R, Toyoda T. Identification and removal of unexpected proliferative off-target cells emerging after iPSC-derived pancreatic islet cell implantation. *Proceedings of the National Academy of Sciences*, 121 (16), e2320883121, 2024

General introduction

Type 1 diabetes mellitus is an autoimmune disorder in which insulin-secreting pancreatic β-cells are destroyed, leading to uncontrolled hyperglycemia and hypoglycemic event, along with multiple devastating complications (1). To cure patients with this disease, whole-organ pancreas transplantation and pancreatic islet transplantation from deceased donors are effective therapeutic options (2, 3). However, due to donor shortages and the gradual loss of graft function, there is an urgent need to develop alternative and unlimited cell sources. As such an alternative, the development of pancreatic islet-like endocrine cells differentiated from pluripotent stem cells, such as induced pluripotent stem cells (iPSC) and embryonic stem cells (ESC), is anticipated (4, 5).

Diabetes mellitus is also a common disease in veterinary medicine, particularly in certain breeds of dogs where the disease frequently occurs due to their genetic backgrounds (6). While conventional insulin therapy is the primary treatment, it poses significant challenges for long-term management, placing considerable burdens on both the affected animals and their owners (7). Therefore, when cell-based therapy, which has the potential to provide a curative therapy, can be established for use in veterinary medicine, it would become a promising treatment option. Since most of the therapeutic agents used in veterinary medicine are adapted from human drugs (8), advances in cell-based therapies for human diabetes could pave the way for future applications in veterinary care.

Multiple studies have demonstrated the *in vitro* generation of pancreatic endocrine cells from human iPSC and ESC by stepwise differentiation that recapitulates the developmental process (9-13). Implantation of the generated cells normalizes hyperglycemia associated with insulin secretion from grafts in rodent type 1 diabetes models. However, the induction efficiency at each differentiation stage was not 100%, and off-target cells, such as undifferentiated cells or cells that have differentiated in the wrong direction, are inevitably included in the generated cells as by-products. These off-target cells potentially cause unwanted outcomes. For instance, implanted cells containing 2-10% unintended cells (14, 15) proliferated to form cysts and enlarged grafts by >10-fold within 200 days after implantation in 8 out of 8 animals (16). Such post-implant proliferation of off-target cells carries not only the risk of mechanical organ damage, but also the risk of potential tumorigenesis. Therefore, approaches to reduce or exclude contaminating off-target cells are essential for clinical applications.

Several attempts, such as cell sorting and protocol optimization, have been made to reduce off-target cell contamination (11, 17-19). Although these approaches suppress graft enlargement by off-target cells, it remains unclear whether all types of off-target cells are effectively removed. Moreover, given that $>1\times10^8$ cells are used in current islet transplantation protocols (20), a similar number of

cells will likely be required for the clinical application of iPSC/ESC-derived pancreatic islet-like cells. Therefore, an efficient methodology is required to eliminate all types of off-target cells from a large number of cells. One feasible candidate approach is small molecule compound treatment based on properties specific to each off-target cells. Hence, the purpose of this work is to profile all types of off-target cells contaminated during the induction of iPSC-derived pancreatic endocrine cells and to identify several compounds that can efficiently remove off-target cells based on their characteristics.

To address this objective, in Chapter 1, non-endocrine cells, which are the major off-target cells contaminating in iPSC-derived pancreatic endocrine cells, were profiled at single-cell resolution. Most of non-endocrine cells were found to be pancreatic-lineage cells, such as pancreatic duct-like cells and pancreatic progenitor cells, with expression of pancreatic and duodenal homeobox 1 (PDX1). Based on characteristics of these pancreas-associated off-target cells, compounds that are effective in removing these unintended cells were investigated. In addition, while evaluating long-term safety of iPSC-derived pancreatic endocrine cells after implantation, novel off-target cells which are outside the pancreatic lineage were unexpectedly discovered. The unexpected off-target cells were found to continue proliferating and become apparent after implantation, even if only slightly contaminated. Therefore, in Chapter 2, the unexpected off-target cells were characterized in detail to determine effective detection and removal methods.

Chapter 1.

Characterization and reduction of non-endocrine cells accompanying islet-like endocrine cells differentiated from human iPSC

1. Introduction

The production of tissues and organs from human iPSCs and ESCs, which have an unlimited proliferative capacity, could solve the problem of donor shortage for organ transplantation (21). Particularly, in the treatment of type 1 diabetes mellitus (T1DM), an autoimmune disorder resulting in severe loss of pancreatic β-cells, cell therapy using pancreatic endocrine cells derived from iPSC and ESC is expected to be an alternative to pancreas or islet transplantation (4, 5). Multiple studies have reported the properties and functions of iPSC/ESC-derived pancreatic endocrine cells, indicating that differentiation methods are reliable and robust (9-11, 17, 22, 23). In addition, these induced pancreatic endocrine cells are comparable to native human islets, showing an apparent glucose-normalizing effect with insulin secretion when implanted into T1DM model animals (9, 10, 23).

iPSC/ESC-derived pancreatic endocrine cells are induced by stepwise differentiation protocols that mimic actual pancreatic development (9-11). As with the developmental process, undifferentiated iPSCs or ESCs are first induced into definitive endoderm, followed by primitive gut tube, posterior foregut, and then pancreatic endoderm. Further differentiation directs pancreatic endoderm cells to become pancreatic endocrine-like cells. This differentiation process typically consists of 6 or 7 stages and takes about 4 weeks to complete. Although the stepwise differentiation protocols can induce pancreatic endocrine-like cells, the induction efficiency at each stage is not 100%. As a result, off-target cells, such as undifferentiated cells or cells that have differentiated in the wrong direction, are inevitably included in the generated cells as by-products. Contamination of off-target cells, especially those with proliferative potential, is a major risk for clinical application. Therefore, methods to efficiently remove off-target cells are required.

In the context of previous studies in my laboratory, an induction protocol for seven-stage iPSC-derived pancreatic islet cells (s7-iPIC) was established (12, 13, 24). A notable feature of s7-iPIC is that more than 90% of the cells are endocrine cells, with only a few non-endocrine off-target cells. This feature of s7-iPIC seems to rely on the treatment with kinase inhibitors (PD-166866 and TR06141363), which are added during the induction process as own inducers. However, regarding s7-iPIC, the following were not fully evaluated: 1) the mechanism by which kinase inhibitor treatment can selectively reduce non-endocrine off-target cells, and 2) the *in vivo* anti-diabetic effects of s7-iPIC. Therefore, in this chapter, the anti-diabetic effects of s7-iPIC were evaluated using T1DM model mice. In addition, s7-iPIC were analyzed at single cell resolution to scrutinize the involvement of kinase inhibitors in the removal of non-endocrine off-target cells.

2. Materials and methods

2.1 Cell culture and s7-iPIC differentiation

Ff-I14s04 and QHJI-14s04 were kindly provided by the Center for iPS Cell Research and Application (CiRA), Kyoto University. Ff-I14s04 is derived from the same clone as QHJI but cultured and stocked for non-clinical use. QHJI-14s04 is a stock for clinical use. Cells were maintained on iMatrix-511 (Nippi)-coated dishes in StemFit AK03N (Ajinomoto) at 37 °C in a humidified 5% carbon dioxide (CO_2) incubator. Cells were passaged every 3 or 4 days by non-enzymatic dissociation using 0.5 mM ethylenediaminetetraacetic acid (EDTA) (Thermo Fisher Scientific) and subjected to differentiation experiments, usually after over 2 weeks of running culture. The use of human iPSCs was approved by the ethical review committee of Kyoto University and Takeda Pharmaceutical Company Limited. For differentiation culture to generate s7-iPIC, two-dimensional (2D) monolayer to static aggregate culture based on previous report (13) and three-dimensional (3D) stirred-floating aggregate culture were performed. The details for 3D floating culture are provided below.

Stage 1

Dissociated undifferentiated iPSCs were resuspended at a density of 6×10^6 cells in a spinner type 30 mL bio-reactor (Biott) suspended in AK03N medium containing 10 μM Y-27632 (FUJIFILM Wako) and stirred at a speed of 70 rpm throughout culture. The next day, aggregated cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (high glucose, GlutaMAX Supplement, pyruvate; Thermo Fisher Scientific) or RPMI 1640 medium (Thermo Fisher Scientific) supplemented with 1% (v/v) penicillin/streptomycin (P/S, FUJIFILM Wako), 1× B-27 supplements (Thermo Fisher Scientific), 1% Pluronic® F-68 (Poloxamer 188, Merck Millipore) to reduce fluid mechanical damage, 5-10 ng/ml (25) activin A (Act A) (PeproTech), 3 μM CHIR99021 (CHIR) (Axon Medchem), and 1% dimethyl sulfoxide (DMSO) (FUJIFILM Wako). The following day, CHIR was removed from the medium, and culture was continued for another 2 days.

Stage 2

Cells were cultured with MCDB 131 medium (Thermo Fisher Scientific) supplemented with 1% P/S, 0.5× B27, 1% Pluronic® F-68 and 50 ng/ml keratinocyte growth factor (KGF) (R&D Systems), 4.44 mM glucose (added to yield a final concentration of 10 mM, FUJIFILM Wako), 1.5 g/L sodium hydrogen carbonate ($NaHCO_3$) (FUJIFILM Wako), and 1% GlutaMAX (Thermo Fisher Scientific) for 4 days.

Stage 3

Continuing culture was performed with improved minimum essential medium (iMEM) (Thermo

Fisher Scientific) containing 1% P/S, 0.5× B27, 1% Pluronic® F-68, 50 ng/ml KGF, 100 ng/ml Noggin (FUJIFILM Wako), 0.5 μM 3-keto-N-aminoethyl-N'-aminocaproyldihydrocinnamoyl cyclopamine (K-Cyc) (Toronto Research Chemicals), and 10 nM 4-[(E)-2- (5,6,7,8-tetrahydro- 5,5,8,8-tetramethyl-2-naphthalenyl)-1-propenyl] benzoic acid (TTNPB) (Santa Cruz Biotechnology) for 3 days.

Stage 4

Cells were exposed to iMEM containing 1% P/S, 0.5× B27, 1% Pluronic® F-68, 100 ng/ml KGF, 50 ng/ml epidermal growth factor (EGF) (R&D Systems), 10 mM nicotinamide (Nic) (STEMCELL Technologies), 0.1 μM TR05991851 (Takeda original Rho-associated coiled-coil kinase inhibitor, ROCKi), 0.5 μM phorbol 12,13-dibutyrate (PDBU, Merck Millipore), and 5 ng/mL Act A for 4 days.

Stage 5

Cells were treated with iMEM with 1% P/S, 0.5× B27, 1% Pluronic® F-68, 0.25 μM SANT-1 (Merck), 50 nM retinoic acid (RA) (Merck), 10 μM activin receptor-like kinase 5 inhibitor II (ALK5i) (Santa Cruz Biotechnology), 100 nM LDN-193189 (LDN) (MedChemExpress), 1 μM L-3,3′,5-triiodothyronine (T3) (Merck Millipore), 50 ng/ml basic fibroblast growth factor (bFGF) (PeproTech), 1 μM XAV939 (Merck Millipore), and 10 μM Y-27632 (as a ROCKi) for 2 days.

Stage 6

Cells were cultured with iMEM containing 1% P/S, 0.5× B27, 1% Pluronic® F-68, 1 μM γ-secretase inhibitor (GSI) (RO4929097, Chem Scene), 10 μM ALK5i, 100 nM LDN, and 1 μM T3 for 4 days. To generate prototype iPIC, cells were treated with the same medium for another 7 days. In the case of six-stage iPIC (s6-iPIC) and s7-iPIC, 1–10 μM PD-166866 (Merck Millipore) was added starting on the fourth day. In experiments to evaluate the effect of compound treatments on non-endocrine cell populations (Fig. 11c-h), 3 μM GSK 461364 (Cayman), CFI-400945 (Apexbio), 1 μM TR06141363 (Takeda original multi-kinase inhibitor), and 2-deoxy-D-glucose (2-DG) were added as alternatives to PD-166866 in the s6-iPIC protocol. To generate s7-iPIC, cells were dissociated on the seventh day at stage 6, seeded into Elplasia multi-well plates (4440, Corning) and cultured for the remaining 4 days in Stage 7 medium.

Stage 7

Stage 7 medium is based on Rezania et al. (9) with some modifications. Cells were exposed to MCDB 131 medium with 1% P/S, 2% fat-free bovine serum albumin (BSA) (FUJIFILM Wako), 14.44 mM glucose (to generate a final concentration of 20 mM), 1.5 g/L $NaHCO_3$, 1% GlutaMAX, 0.5% insulin-transferrin-selenium-ethanolamine (ITS-X) (Thermo Fisher Scientific), 10 μM ALK5i, 1 μM T3, 10 μM Zinc sulfate ($ZnSO_4$) (Merck Millipore), 1.4 IU/ml heparin sodium salt (Nacalai Tesque), 1 mM

N-acetyl cysteine (N-Cys) (Merck Millipore), 10 μM Trolox (FUJIFILM Wako), 2 μM R428 (Selleck), 1–10 μM PD-166866, 3 μM TR06141363, and 10 μM Y-27632, for 4 days.

2.2 Flow cytometry

Differentiation efficacy and quality at individual stages based on the developmental markers were analyzed with immunostaining methods and LSRFortessa X20 flow cytometry equipment (BD), as described previously (13). Data were processed with FlowJo software. The primary antibodies are listed in Table 1. Secondary antibodies of the appropriate species were conjugated to AlexaFluor 488, 546, 568 and 647 of appropriate species (Thermo Fisher Scientific or Jackson).

Table 1. List of primary antibodies used in immunofluorescence staining.

Antigen	Species	Manufacturer	Catalog #	Dilution for flow cytometry	Dilution for immunohistochemistry
PDX1	Goat	R&D systems	AF2419	1 : 200	1 : 200
NKX6.1	Mouse	DSHB	F55A12	1 : 380	-
CHGA	Rabbit	Abcam	ab68271	1 : 500	1 : 100
CHGA	Mouse	Thermo	MAS-13096	1 : 100	-
Ki67	Mouse	BD Biosciences	556003	1 : 100	-
Ki67	Rabbit	CST	9129S	-	1 : 100
INS	Rat	DSHB	GN-ID4	1 : 600	-
INS	Guinea pig	Dako	A0564	-	1 : 200
INS	Rabbit	CST	CST3014	-	1 : 200
GCG	Mouse	Sigma	G2654	1 : 200	1 : 200
HuN	Mouse	Takara Bio	Y40400	-	1 : 100
mouse CD31	Rat	dianova	DIA-310	-	1 : 100

2.3 Type 1 diabetes mouse model

NOD.CB17-Prkdc-scid/J (NOD-scid) mice were obtained from Charles River. Male mice between the ages of 8 and 9 weeks were intraperitoneally injected with multiple low doses of streptozotocin (STZ, 50 mg/kg/day for 5 days, Sigma). Within 2-3 weeks after STZ injection, almost all mice developed diabetic pathology with blood glucose levels above 300 mg/dL. These type1 diabetes model mice (STZ-NOD-scid mice) were subjected to implantation experiments. All animal studies were conducted at Shonan iPark, one of the international accreditation facilities by Association

for Assessment and Accreditation of Laboratory Animal Care International (AAALAC), and approved by the iPark institutional animal care and use committee. Animal experiment approval numbers: AU-00011518, AU-00011622, AU-00011795, AU-00020231, AU-00020664, and AU-00021121. All experiments were performed in accordance with the relevant guidelines and regulations, including the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines.

2.4 Implantation and in vivo assessment

Differentiated s7-iPIC aggregates were mixed with 100 μL of fibrinogen/50 μL of thrombin solution, incubated at 37°C for 5 min and then implanted in the subcutaneous space of STZ-NOD-scid mice (3-4 × 10^6 cells/mouse) under anesthesia induced by inhalation of 2-3% isoflurane. To serve as a control, animals in the sham operation group underwent the same surgical procedures without the cell implantation. Fibrinogen from human plasma (Merck Millipore) and thrombin (Sigma) were reconstituted in iMEM and in phosphate-buffered saline (PBS) to make 10 mg/mL and 50 IU/mL solutions, respectively, and stored at -80°C until use. For the kidney capsule implantation study, s6-iPIC (1.4 × 10^6 cells/mouse) or prototype (3.6 × 10^6 cells/mouse) was implanted directly into the kidney capsule without fibrin gel. The blood glucose levels of implanted animals were monitored using an Accu-Chek Aviva system (Roche DC Japan), and plasma samples were collected from the tail vein on the indicated days. For the oral glucose tolerance test, the mice were fasted overnight and orally injected with a 2 g/kg glucose solution (Otsuka), and plasma samples were collected from the tail vein before and 15, 30, 60 and 120 min after injection.

2.5 Plasma glucose and hormone measurements

Plasma glucose, human C-peptide and mouse C-peptide levels were measured using the glucose test C-II Wako (Fujifilm Wako), Mercodia Ultrasensitive C-peptide enzyme-linked immuno-sorbent assay (ELISA) kit (Mercodia) and mouse C-peptide measurement kit (Morinaga) according to the manufacturer's instructions.

2.6 Tissue processing and immunostaining

Implanted grafts, which had engrafted subcutaneously, were excised under anesthesia induced by inhalation of 2-3% isoflurane. The grafts were easily retrieved from the host tissue by blunt dissection. Following this procedure, the animals were euthanized by cervical dislocation under anesthesia. The collected grafts were then fixed with 4% paraformaldehyde (FUJIFILM Wako) for

over 24 h at 4°C and embedded in paraffin or frozen in Optimal Cutting Temperature (OCT) compound. Paraffin blocks were sectioned at 5 μm and used for hematoxylin and eosin staining, Masson trichrome staining and immunostaining. Frozen blocks were sectioned at 10 μm and used for immunofluorescence staining. The primary antibodies are listed in Table 1. Secondary antibodies were conjugated to AlexaFluor 488, 546 or 568 (Thermo Fisher Scientific or Jackson). Frozen sections were also counterstained with Hoechst (Thermo Fisher Scientific) to label the nucleus.

2.7 Single-cell ribonucleic acid sequencing (scRNA-seq) library preparation, sequencing and data processing

A total of 6 samples (1 sample of s7-iPIC, 3 samples of s6-iPIC, 1 sample of prototype, and 1 sample of reference human islet) underwent scRNA-seq. Human islets were purchased from Prodo Labs. All 5 samples of three versions of iPICs were prepared from Ff-I14s04 line according to the inducers of Fig. 2. iPICs were cultured from stage 1 with 3D stirred-floating culture with shear stress using bio-reactor (Biott). As an exception, only 1 of 3 samples of s6-iPIC was cultured in 2D monolayer from stages 1 to 4, and then in static aggregate culture without shear stress using non-adhesive V bottom 96 well plate (SUMITOMO BAKELITE) in accordance with previous methods (13). Libraries for scRNA-seq were generated using the 10x Genomics Chromium^{TM} controller and Chromium Single Cell 3' kits v2 (10x Genomics) according to the manufacturer's instructions. Successful complementary deoxyribonucleic acid (cDNA) amplification and library construction were ensured with High Sensitivity DNA kits on an Agilent 2100 Bioanalyzer (Agilent). The obtained libraries were sequenced using Hi-seq (Illumina) with 150 bp paired-end reads at a depth of >100,000 reads per cell. Sequencing reads were aligned to the human GRCh38 genome reference, and gene counts were quantified as unique molecular identifiers (UMIs) using Cell Ranger v2.0.1 (10x Genomics). UMI count matrices were imported into the R v3.3.1 software Seurat v2.0.1 package (26, 27), where normalization was performed according to the package's default setting. Cells with mitochondrial gene counts over 10% were regarded as dead or damaged cells and removed for further analyses. UMI count matrices were scaled by regressing out the number of total UMI counts per cell and the percentage of mitochondrial gene counts. Genes for dimensional reduction were selected by the average expression and dispersion of each gene, and principal component analysis was performed. Principal components were used for Seurat's shared nearest neighbor graph clustering and t-distributed stochastic neighbor embedding (*t*-SNE) dimensional reduction to create a visualization of data. The cell cycle was evaluated and scored using the expression of genes known as S-phase, G1, and G2M

markers. To estimate cell types and similarities within iPICs with reference to human organs, reference component analysis (RCA) was performed using the RCA v1.0.0 package (28). Differential gene expression analysis of each cluster compared with the others was performed using the likelihood-ratio test for single-cell gene expression in Seurat. For trajectory analysis, processed UMI count matrices were imported to a single-cell dataset for the Monocle v2.6.3 package (29-31). The genes for ordering cells with 'dpFeature' in monocle were selected, and the single-cell trajectories were contracted via the 'DDRTree' algorithm.

2.8 Measurement of inhibitory effects on fibroblast growth factor receptor (FGFR) and polo-like kinase (PLK) isoforms

The percent (%) inhibition of FGFR and PLK isoforms by PD-166866, TR06141363, GSK 461364, and CFI-400945 was evaluated with a time-resolved fluorescence resonance energy transfer (TR-FRET)-based competitive binding assay (32), performed using 1536-well, white, flat-bottomed plates (Greiner Bio-One). Assay buffer comprised 50 mmol/L 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 10 mmol/L magnesium chloride ($MgCl_2$), 1 mmol/L ethylene glycol tetra-acetic acid (EGTA), 0.1 mmol/L dithiothreitol (DTT), and 0.01% Brij-35. Test compounds (1 and 0.1 μmol/L) were mixed with recombinant kinase proteins with an epitope tag, a terbium-labelled anti GST-tag antibody, and fluorescent labelled ligands. After 1 h incubation at room temperature, TR-FRET signals were measured with EnVision (PerkinElmer). DMSO and 5 μmol/L control inhibitor (Staurosporine or other multi kinase inhibitors) were used as 0 and 100% controls, respectively. The percent (%) inhibition was calculated based on the signals of the 0% and 100% inhibition samples in the absence and presence of control inhibitors, respectively.

2.9 Analysis of messenger RNA (mRNA) expression by quantitative real-time polymerase chain reaction (PCR)

After iPIC differentiation, cDNA samples were synthesized from lysates using TaqMan Gene Expression Cells-to-Ct kits (Thermo Fisher Scientific) according to the manufacturer's instructions, followed by quantitative real-time polymerase chain reaction analysis using a Prism 7900HT sequence detector (Thermo Fisher Scientific). The thermal cycling parameters were 2 min at 50°C and 10 min at 95°C, followed by 40 cycles at 95°C for 15 sec and 60°C for 1 min. The mRNA levels were analyzed with the comparative Ct method (2-ΔΔCt) using *RPLP0* as the housekeeping gene. The TaqMan Gene Expression assays (Thermo Fisher Scientific) used herein were as follows: Hs00170586_m1 (*FGB*),

Hs00356521_m1 (*AGR2*), and Hs00420895_gH (*RPLP0*).

2.10 Statistical analysis

Data are expressed as the mean and SD values. Dunnett's multiple-comparison test was performed at a significance level of $P < 0.05$ to compare the mean values of experimental groups against the control group (Fig. 8c, 8d, 11c, and 11e). Additionally, the Aspin-Welch test or Student's t-test was used at a significance level of $P < 0.05$ for comparing two groups (Fig. 8c and 8d). The dose-response relationships in Fig. 11f and 11h were tested using Williams or Shirley-Williams tests with a one-tailed significance level of $P < 0.025$ based on the results of the homogeneity of variance test (Bartlett's test). All statistical analyses were performed using Statistical Analysis System version 9.3 (SAS Institute, NC, USA).

2.11 Data availability

The data that support the findings of this study are available from the author upon reasonable request. Data of scRNA-seq were deposited in the Gene Expression Omnibus database (GSE213290).

3. Results

3.1 Subcutaneously engrafted s7-iPIC showed long-term efficacy and morphology maintenance

With reference to landmark reports describing the differentiation protocol for human iPSC/ESC-derived endocrine cells (9, 10, 12, 13), seven-stage iPSC-derived pancreatic islet cells (s7-iPIC) were generated from the Ff-I14s04 and QHJI-14s04 iPSC lines using a 7-step *in vitro* differentiation protocol (Fig. 1a and 2). In s7-iPIC, ~95% of cells were positive for the endocrine cell marker Chromogranin A (CHGA), more than 95% of cells were positive for the pancreatic cell marker PDX1, and few $CHGA^-$ non-endocrine cells expressed the proliferative marker Ki67 (0.1%) (Fig. 1b and c). Approximately 30-35% of cells were either NK6 homeobox 1 $(NKX6.1)^+$/insulin $(INS)^+$ or $NKX6.1^-/INS^+$, and these cells were expected to mature into β and α cells, respectively, after implantation (10, 33, 34) (Fig. 1b and c).

To investigate the cell maturation and function of s7-iPIC *in vivo*, s7-iPIC embedded in fibrin gel, which is known to support the engraftment of porcine islets (35, 36), was implanted into the subcutaneous spaces of mice with streptozotocin-induced diabetes. The grafts of s7-iPIC formed granulation tissue and could be easily retrieved from the host 25-36 weeks after implantation (Fig. 1d). The graft size was less than 5 mm in diameter, which was not larger than the mixture of s7-iPIC and fibrin gel before implantation, in all mice. Histological analysis showed that the grafts were composed of compact endocrine cell clusters (50-500 μm) surrounded by host-derived fibrous tissues (Fig. 1e and f). The endocrine cell clusters were characterized by the arrangements of INS^+ cells in the core and glucagon $(GCG)^+$ cells in the periphery, similar to the structure of prenatal fetal human islets and adult rodent islets (37, 38). While the proliferation of non-target cells was not observed throughout the graft, endocrine clusters contained a few $Ki67^+$/human nucleus $(HuN)^+$ proliferative cells (Fig. 1f). This is within the normal range of cell turnover, given that a few $Ki67^+$ cells have been reported in endogenous islets *in vivo* (39, 40). Infiltration of host-derived blood vessels into the endocrine clusters was evidenced by the presence of mouse $CD31^+$ cells and red blood cells (Fig. 1e and f). Consistent with histological observations, eight weeks after implantation, human C-peptide secretion from implanted s7-iPIC plateaued at a high level (6.3 ± 1.8 ng/mL), followed by return to a normoglycemic state (Fig. 1g and h). To evaluate insulin secretion in response to glucose levels, an oral glucose tolerance test was performed 23 weeks after implantation, showing that plasma human C-peptide levels were increased within 15 min after glucose loading (Fig. 1i and j). This rapid response was comparable or superior to that of endogenous insulin secretion from mouse islets (Fig. 1k). The blood glucose levels in s7-iPIC implanted mice were lower than those in non-STZ control mice (Fig. 1g), which is

possibly attributed to the lower blood glucose setpoint by human islets than by mouse islets in rodent (16, 41). Nevertheless, s7-iPIC-implanted mice showed body weight transition equivalent to that of non-STZ control mice (Fig. 3) with no abnormalities in physical conditions. These results suggested that s7-iPIC has sufficient potential to normalize blood glucose levels in T1DM model mice over several months without morphological abnormality.

3.2 Classification of cells composing iPICs to expose potential remaining non-endocrine off-target cells

Based on the results of the *in vivo* experiments, the risk of contaminating off-target cells was relatively low in s7-iPIC. However, considering the cell number ($>1\times10^8$) assumed to be needed for implantation in patients (20), it is reasonable to pursue an understanding of, and methods for reducing, potentially contaminating cells as much as possible, to minimize future clinical risk levels. To expose potentially contaminating off-target cells, the single-cell transcriptomic analysis of s7-iPIC was conducted. However, since the percentage of off-target cells in s7-iPIC was approximately 5% (Fig. 1b), it was assumed that some contaminating cells could fail to be detected by scRNA-seq. Therefore, to widen the variety of potentially contaminating cells, not only s7-iPIC but also other two types of iPICs were analyzed: s6-iPIC that was not cultured with stage 7 differentiation factors, and prototype iPIC that was not treated with s7 differentiation factors and PD-166866 (Fig. 2 and 4a). The proportions of $CHGA^-$ and $CHGA^-Ki67^+$ cells were comparable in s6-iPIC (5.2% and 0.2%) and increased in prototype iPIC (40.8% and 9.2%) compared to s7-iPIC (Fig. 1b and 4b).

Sequence data were collected from 19,969 cells, including the three versions of iPICs and the reference human islet sample. After quality control, unsupervised cell clustering and visualization were performed on the data from the remaining 17,100 cells (Fig. 4c and d). The cells in the different induction batches of s6-iPIC were all classified into the same cell clusters with a slight percentage variation, indicating that there was no prominent batch effect in the differentiation protocol (Fig. 5).

Next, the cell clusters were classified based on pancreas-related gene expression (Fig. 4d and 6a-b). In iPICs, three major $CHGA^{high}$ endocrine cell identities were observed, as previously reported (14, 17, 22): (i) β-cell fate cells (Clusters 2 and 3); (ii) a mixture of α-, δ-, pancreatic polypeptide (PP)- and gastrin-cell fate cells (42) (Cluster 0); and (iii) enterochromaffin (EC)-cell fate cells (Clusters 1, 4, 8 and 12). There were three other endocrine clusters: ε-cell fate cells (Cluster 17), endocrine progenitors that express *NGN3* (Cluster 9), and intermediate cells for β- or α- cells (Cluster 14). Conversely, Clusters 5, 13, and 16, and most of Cluster 10 were assigned to $CHGA^-$ non-endocrine

off-target cells with low expression of the endocrine markers *CHGA* and *NEUROD1* (inside the black dotted line in Fig. 4d and e). These non-endocrine cell populations expressed *PDX1*, *KRT19*, and *SOX9*, as did pancreatic duct cells (Cluster 6) in human islets (Fig. 4e), while *PRSS1* and *CPA1*, which are strongly expressed in acinar cells (Cluster 18), were not expressed (Fig. 6a). Considering that most $CHGA^-$ cells were $PDX1^+$ in flow cytometry analysis of iPICs (Fig. 4b), non-endocrine off-target cells were likely to be cells of the developing pancreatic duct or surrounding tissues. In terms of the residual degree of these $CHGA^-$ non-endocrine cell populations, no difference was detected between s6-iPIC and s7-iPIC in flow cytometry analysis (Fig. 4b), but in scRNA-seq, the residual cells of Cluster 5 and 13 were higher in s6-iPIC than in s7-iPIC (Fig. 4c and 5b). Thus, by employing iPICs with different degrees of non-endocrine cells, non-endocrine off-target cells were extracted at single-cell resolution, revealing that these cells were reduced gradually in the process leading to current best-practice s7-iPIC.

3.3 Characterization of non-endocrine cell populations and detection of non-endocrine subpopulations with qPCR

To evaluate non-endocrine off-target cell populations in multilateral ways, scRNA-seq data were analyzed using several methods. First, to define the developmental hierarchy of non-endocrine cells, reconstruction of cells consisting of iPICs was performed by trajectory analysis using Monocle 2 (29-31). Trajectory projections of all types of iPICs showed only one branching point (Fig. 7a). Given that endocrine cells are mature cell types, Clusters 5, 10, 13, and 16 were projected to be more immature progenitor-like populations. Next, to evaluate proliferative activity, cell cycle phase assignments were performed based on S-phase, G2 and M gene signatures, and observed high S-scores in Clusters 5, 10, 13, and 16 (Fig. 7b). In particular, Cluster 10 had a high G2M-score and expressed *MKI67* (Fig. 4e). Of note, the lower tip of Cluster 10 in Fig. 4d was actually $CHGA^+$ and $NEUROD1^+$ proliferative cells (Fig. 4e). Thus, Cluster 10 was a heterogeneous population of proliferating cells, including a few replicating endocrine cells in iPICs and human islets (0.8%, 19 out of 2,419 cells) (Fig. 5b). This proportion of proliferating cells in endogenous human islet is consistent with previous reports (39, 40).

Subsequently, to clarify the characteristics of each non-endocrine cell cluster, reference component analysis (RCA), which indicates transcriptome similarity to known tissues or cell lines (28), was performed. Non-endocrine cell populations (Clusters 5, 13, 16, and $CHGA^-$ population of Cluster 10) were again distinct from endocrine cells in RCA (Fig. 8a and 9a). Among non-endocrine cell

populations, Cluster 16 indicated in magenta scored higher for pancreas and pancreatic islets than Clusters 5, 13, and *CHGA*$^-$ population of Cluster 10 (brown, red, and green) (Fig. 8b and 9b). In addition, the expression intensity of *YAP1*, which is known to be repressed before endocrine specification (43), was lower in Cluster 16 than in the other non-endocrine clusters (Fig. 6a). Therefore, Cluster 16 might be the transitional state leading to the neighboring Cluster 9, an *NGN3*$^+$ endocrine progenitor population. The high score for tumor cells in *CHGA*$^-$ population of Cluster 10 (green) is consistent with the idea that Cluster 10 was a heterogeneous population of proliferating cells (Fig. 8b). Clusters 5 (brown) and 13 (red) scored high for the fetal liver and colon, respectively, in addition to the pancreas (Fig. 8b and 9c). In line with these results, Cluster 5 contained liver-related genes such as *FGB*, *GSTA1*, *AFP*, *APOC1* and *APOA*, and Cluster 13 contained intestine-associated *AGR* (44) among the top 10 differentially expressed genes (DEGs) (Fig. 6b). These results suggest that cells in Clusters 5 and 13 are basically pancreatic linage, but partially have liver or intestinal traits.

Although scRNA-seq revealed that cells in Clusters 5 and 13, a typical off-target non-endocrine cell population, were reduced in s7-iPIC compared to s6-iPIC (Fig. 4c and 5b), flow cytometry analysis did not distinguish s6-iPIC and s7-iPIC in terms of the remaining non-endocrine off-target cells (Fig. 8c). To detect the proportion of non-endocrine cells by means other than scRNA-seq that takes several weeks to assess cells, quantitative PCR (qPCR) was performed for *FGB* and *AGR2*, which were specific genes in the highest DEGs for Clusters 5 and 13, respectively (Fig. 6b). The expression of *FGB* and *AGR2* was significantly lower in s7-iPIC than in s6-iPIC (Fig. 8d). Of note, the expression of *FGB* and *AGR2* was not detected in two out of three independent qPCR measurements in s7-iPIC using 1-3 × 10^4 cells, demonstrating that Cluster 5 or 13 cells hardly remained in s7-iPIC. These results indicate that qPCR for *FGB* and *AGR2* is practical and sensitive for reliably detecting residual non-endocrine off-target cell presence instead of scRNA-seq.

3.4 Novel approaches to reducing the number of proliferative non-endocrine off-target cells

Finally, novel approaches were explored to reduce the number of non-endocrine off-target cells. First, the removal effect of PD-166866 on non-endocrine cells was scrutinized. PD-166866, which was used to induce s6-iPIC and s7-iPIC, not only reduced non-endocrine cells *in vitro* (Fig. 4b), but also suppressed graft hypertrophy *in vivo* while maintaining glycemic control activity (Fig. 10a-c). PD-166866 is a well-known selective inhibitor of FGFR1 (45, 46). To clarify the targets of PD-166866, the inhibition profile of PD-166866 for FGFR isoforms was assessed using a global kinase panel assay (32). PD-166866 inhibited all FGFR isoforms in the micromolar range (Fig. 11a), suggesting that PD-

166866 is a pan-FGFR inhibitor. Thus, the expression profile of *FGFR isoforms* in all types of iPICs was extracted from the scRNA-seq data. While *FGFR2* was dominantly expressed in non-endocrine cell subpopulations, *FGFR1*, *FGFR3* and *FGFR4* were ubiquitously expressed (Fig. 11b and 12a). In addition, PD-166866 reduced only the number of non-endocrine off-target cells without affecting endocrine cells (Fig. 11c and d). These results suggest that PD-166866 reduced the number of non-endocrine cells mainly via FGFR2 inhibition in non-endocrine populations.

Next, the investigation focused on whether inhibition of the cell cycle reduces the non-endocrine cell populations, as activation of the cell cycle was observed in these populations (Fig. 7b). As cell cycle inhibitors, PLK inhibitors were selected, for which clinical development as anticancer agents is progressing (47). By extracting the expression of *PLK isoforms* in all types of iPICs, it was found that *PLK1* and *PLK4* were highly expressed in Cluster 10, a non-endocrine cell cluster (Fig. 11b and 12a). In contrast, *PLK2* showed a tendency for higher expression in endocrine cell clusters, while *PLK3* exhibited no significant variation in expression across different cell clusters. Thus, the effects of the PLK1 inhibitor GSK 461364 and the PLK4 inhibitor CFI-400945 on the number of non-endocrine cells were evaluated. In a similar manner, the multi-kinase inhibitor TR06141363, which inhibits PLK4 and FGFR isoforms (Fig. 12b) and is added to the induction step of s7-iPIC generation, was tested. All the compounds with PLK-inhibitory activity reduced non-endocrine cells to the same extent as PD-166866 (Fig. 11c). Notably, these compounds selectively reduced non-endocrine cells with little reduction in endocrine cells (Fig. 11d). Accordingly, the proportion of $NKX6.1^{+}/INS^{+}$ cells increased (Fig. 11e). In addition, a global kinase panel assay validated that the PLK1 inhibitor GSK 461364 had almost no inhibitory activity against FGFR isoforms in the micromolar range (Fig. 12b). These results suggest that PLK inhibition selectively reduces the number of non-endocrine off-target cells via a mechanism different from FGFR inhibition.

For the other novel approach, attention was focused on *LDHB*, which was included in the top 10 significant DEGs of Clusters 5, 10 and 13 (Fig. 6b). Based on the high expression of *LDHB* in non-endocrine cells (Fig. 11b), the hypothesis was formed that a dependency on glycolysis as an energy production pathway is more prevalent in non-endocrine cells than in endocrine cells. Therefore, the investigation focused on whether 2-deoxy-D-glucose (2-DG), which can broadly inhibit glycolysis (48), reduces the number of non-endocrine cells. Treatment with 2-DG significantly reduced the number of non-endocrine cells and increased the proportion of $NKX6.1^{+}/INS^{+}$ cells in a dose-dependent manner (Fig. 11f, g and h). This treatment also decreased the number of endocrine cells, but the reduction rate of non-endocrine cells was higher (80.7 ± 2.6% at 10 mM) than that of endocrine

cells (37.4 ± 6.9% at 10 mM) (Fig. 11g). These results support the idea that glycolysis is more active in non-endocrine cells than in endocrine cells, and inhibition of glycolysis is a potential target for the reduction of unintended non-endocrine cells.

4. Figures

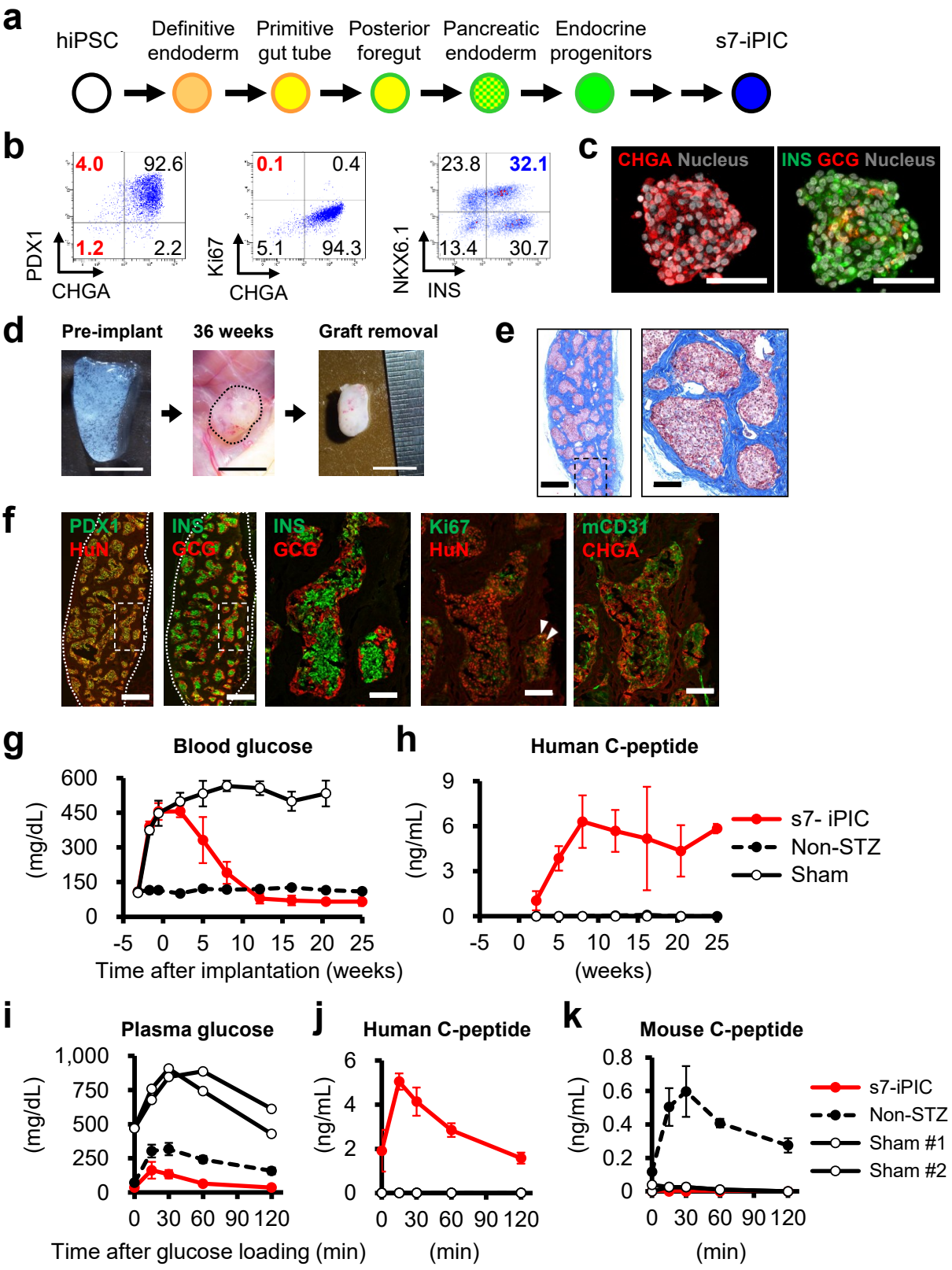

a
hiPSC
Definitive endoderm
Primitive gut tube
Posterior foregut
Pancreatic endoderm
Endocrine progenitors
s7-iPIC
b
4.0
92.6
1.2
2.2
PDX1
CHGA
0.1
0.4
5.1
94.3
Ki67
CHGA
23.8
32.1
13.4
30.7
NKX6.1
INS
c
CHGA Nucleus
INS GCG Nucleus
d
Pre-implant
36 weeks
Graft removal
e
f
PDX1 HuN
INS GCG
INS GCG
Ki67 HuN
mCD31 CHGA
g
Blood glucose
(mg/dL)
600
450
300
150
0
-5 0 5 10 15 20 25
Time after implantation (weeks)
h
Human C-peptide
(ng/mL)
9
6
3
0
-5 0 5 10 15 20 25
(weeks)
s7- iPIC
Non-STZ
Sham
i
Plasma glucose
(mg/dL)
1,000
750
500
250
0
0 30 60 90 120
Time after glucose loading (min)
j
Human C-peptide
(ng/mL)
6
4
2
0
0 30 60 90 120
(min)
k
Mouse C-peptide
(ng/mL)
0.8
0.6
0.4
0.2
0
0 30 60 90 120
(min)
s7-iPIC
Non-STZ
Sham #1
Sham #2

Figure 1 | Generation and subcutaneous implantation of s7-iPIC.

a, Schematic representation of s7-iPIC differentiation from human iPSC.

b, Representative flow cytometry plots illustrating the protein expression of s7-iPIC. The presented plots are those of samples for scRNA-seq.

c, Representative immunohistochemistry images of s7-iPIC aggregates before implantation. White bars indicate 50 μm.

d-k, Cell implantation experiments in STZ-NOD-scid mice. Mice were implanted with s7-iPIC (3-4 × 10^6 cells/mouse) embedded in fibrin gel into the subcutaneous space (s7-iPIC). Sham-operated STZ-NOD-scid mice (Sham) and non-STZ control mice (Non-STZ) were also prepared.

d, Macroscopic photographs of s7-iPIC embedded in fibrin gel before implantation and the s7-iPIC grafts subcutaneously engrafted or retrieved 36 weeks after implantation. White and black bars indicate 5 mm. Images are representative of six samples showing similar results.

e, Masson's trichrome-stained section 25 weeks after implantation. The collagen-rich fibrotic regions are stained blue. Black bars indicate 500 μm at low magnification and 100 μm at high magnification. Images are representative of six samples showing similar results.

f, Graft immunohistochemistry images 25 weeks after implantation. White bars indicate 500 μm at low magnification in the left two images and 100 μm at high magnification in the right three images. White arrowheads indicate $Ki67^+/HuN^+$ cells. Images are all taken from the same sample and are representative of six samples showing similar results.

g and h, Blood glucose (**g**) and plasma human C-peptide (**h**) levels after s7-iPIC (3 × 10^6 cells/mouse) implantation. Data are shown as the mean ± standard deviation (SD) (s7-iPIC; n = 5→3, Non-STZ; n = 4→3, sham; n = 4→3). The decrease in n number is due to unexpected death.

i-k, Plasma glucose (**i**), human C-peptide (**j**) and mouse C-peptide (**k**) levels during the oral glucose tolerance test at 23 weeks after implantation. Data are shown as the mean ± SD (s7-iPIC; n = 4, Non-STZ; n = 3, sham; individual data in n = 2).

	Stage 1 (3 days)	Stage 2 (4 days)	Stage 3 (3 days)	Stage 4 (4 days)	Stage 5 (2 days)	Stage 6 (7 days)	Stage 6 or 7 (4 days)
s7-iPIC	Act A CHIR	KGF	KGF Noggin K-Cyc TTNPB	KGF EGF Nic ROCKi PDBU Act A	SANT-1 RA ALK5i LDN T3 XAV939 bFGF ROCKi	GSI ALK5i LDN T3 PD-166866	ALK5i T3 N-Cys Trolox R428 PD-166866 TR06141363
s6-iPIC						GSI ALK5i LDN T3 PD-166866	GSI ALK5i LDN T3 PD-166866
prototype						GSI ALK5i LDN T3	GSI ALK5i LDN T3

Figure 2 | A Schematic diagram of the differentiation reagents for three types of iPICs.

Act A; activin A, CHIR; CHIR99021, KGF; keratinocyte growth factor, K-Cyc; 3-keto-N-aminoethyl-N'-aminocaproyldihydrocinnamoyl cyclopamine, TTNPB; 4-[(E)-2-(5,6,7,8-tetrahydro-5,5,8,8 tetramethyl-2-naphthalenyl)-1-propenyl] benzoic acid, EGF; epidermal growth factor, Nic; nicotinamide, ROCKi; Rho-associated coiled-coil kinase inhibitor, PDBU; phorbol 12,13-dibutyrate, RA; retinoic acid, ALK5i; activin receptor-like kinase 5 inhibitor II, LDN; LDN-193189, T3; L-3,3′,5-triiodothyronine, bFGF; basic fibroblast growth factor, GSI; γ-secretase inhibitor, N-Cys; *N*-acetyl cysteine.

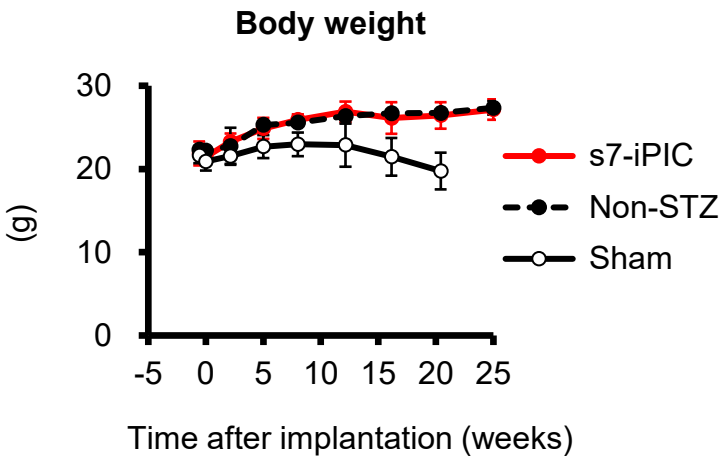


Figure 3 | Body weight change during implantation study in Figure 1.

Body weight change following s7-iPIC (3×10^6 cells/mouse) implantation. Data are shown as the mean ± SD (s7-iPIC; n = 5→3, Non-STZ; n = 4→3, sham; n = 4→3). The decrease in n number is due to unexpected death.

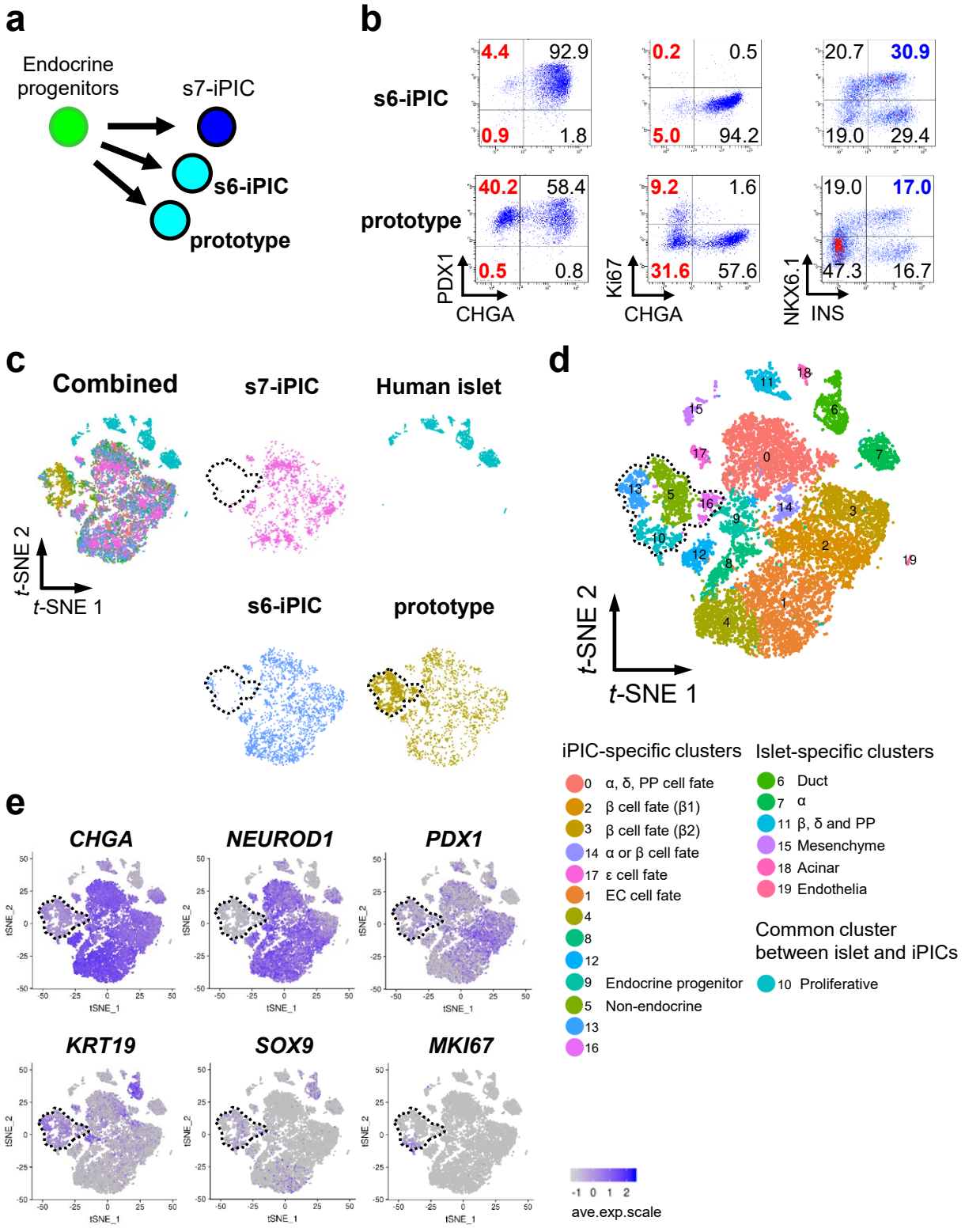


Figure 4 | Classification of the cells composing iPICs and extraction of non-endocrine off-target cells through scRNA-seq.

a, Schematic representation of three versions of iPICs differentiation subjected to scRNA-seq.

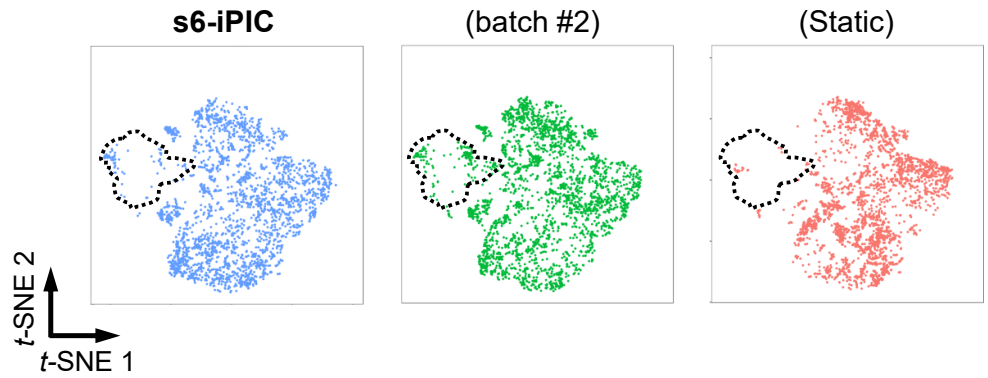
b, Representative flow-cytometry plots illustrating the protein expression of s6-iPIC (upper) and prototype (lower). The presented plots are those of samples for scRNA-seq.

c, A combined cell distribution from a total of six samples and individual cell distributions. The six samples are as follows: 1 sample of s7-iPIC, 3 samples of s6-iPIC, 1 sample of prototype, and 1 sample of reference human islet. For s6-iPIC, only a representative sample of the 3 samples is shown. For the remaining two samples of s6-iPIC, individual *t*-SNE projects are shown in Figure 5a. The dotted line in black represents the non-endocrine (*CHGA*$^{-}$*NEUROD1*$^{-}$) population from the iPICs, as shown in Figure 4e.

d, Shared nearest neighbors clustering identified fourteen cell clusters for *in vitro* differentiated iPICs (0-5, 8-10, 12-14, 16, and 17) and seven clusters for human islet samples (6, 7, 10, 11, 15, 18, and 19). Clusters 1, 4, 8, and 12 are all EC cell fate, while Clusters 5, 13, 16 are all non-endocrine cell. The dotted line in black represents the non-endocrine population (Clusters 5, 13, 16, and *CHGA*$^{-}$ population of Cluster 10) from the iPICs.

e, Single-cell gene expression of representative markers related to the pancreas, endocrine cells and proliferation in combined *t*-SNE projections. The dotted line in black represents the non-endocrine population from the iPICs.

a



b

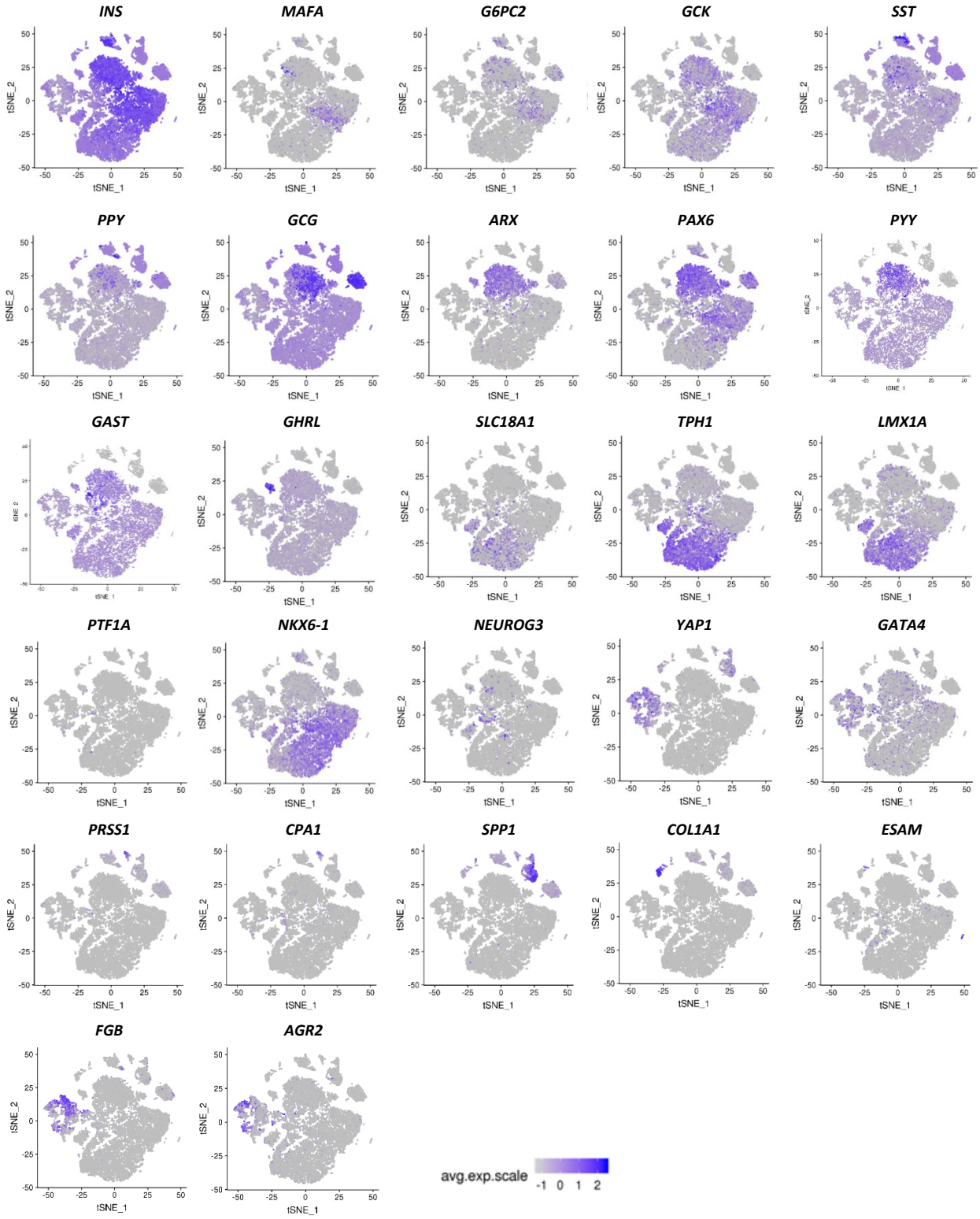
Cluster	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	% population																			
s7-iPIC	22.4	21.1	21.5	6.9	11.2	0.0	0.0	0.0	1.6	3.6	0.8	0.0	5.0	0.0	3.5	0.0	0.3	2.0	0.0	0.0
s6-iPIC	17.7	21.9	14.1	10.5	16.1	0.7	0.0	0.0	2.9	3.6	0.3	0.0	3.8	2.1	3.1	0.0	0.6	2.5	0.0	0.0
(batch #2)	27.1	14.1	16.8	8.7	11.3	2.8	0.0	0.0	4.2	3.7	0.6	0.0	4.4	1.7	2.1	0.0	1.0	1.6	0.0	0.0
(Static)	20.8	30.7	9.4	17.8	3.6	0.1	0.0	0.0	10.8	2.7	0.5	0.0	0.7	0.9	1.8	0.0	0.3	0.0	0.0	0.0
prototype	12.9	6.8	7.6	4.1	5.4	24.2	0.0	0.0	2.1	5.8	13.0	0.0	2.2	8.5	1.6	0.0	5.1	0.7	0.0	0.0
Human islet	0.0	0.0	0.0	0.0	0.0	0.0	30.1	26.8	0.0	0.0	0.8	21.7	0.0	0.0	0.0	12.2	0.0	0.0	5.5	2.9

Figure 5 | Additional information on single-cell RNA sequencing samples.

a, Individual *t*-SNE projections for the first batch of s6-iPIC (which is also shown in Fig. 4c as a representative sample) and two additional s6-iPIC samples (batch #2 and static culture). The dotted line in black represents the non-endocrine population from the iPICs.

b, Percentage of cells in the clusters classified in Fig. 4d. Of note, in s7-iPIC, cells were not detected in Clusters 5 and 13, while cells were detected in proliferative cluster 10 across all iPICs and islet samples.

a


INS
MAFA
G6PC2
GCK
SST
PPY
GCG
ARX
PAX6
PYY
GAST
GHRL
SLC18A1
TPH1
LMX1A
PTF1A
NKX6-1
NEUROG3
YAP1
GATA4
PRSS1
CPA1
SPP1
COL1A1
ESAM
FGB
AGR2
tSNE_1
tSNE_2
avg.exp.scale
-1 0 1 2

b

Cluster	# of DEGs	Top 10 significant differentially expressed genes (p-value < 0.05 & FC > 2)
0	33	PYY, GAST, APOA2, GCG, DPP4, ETV1, SMC2, IGF2, UCN, ARX
1	30	STC1, TPH1, CBLN1, PRPS2, OVOS2, DDC, SYT13, ADRA2A, CRISPLD1, CRYBA2
2	14	SPEG, PCDH7, CALB2, ASIC1, CHRNA3, ACVR1C, HADH, CDK6, LBH, WSCD2
3	29	PCP4, NEFM, NEFL, LOXL4, DCX, PLXNA2, TMEM185A, CALB2, ACVR1C, DLK1
4	80	TAC1, PRAC1, FEV, KCTD12, TPH1, SSTR1, CBLN2, DDC, C9orf16, TMEFF2
5	102	FGB, APOE, GSTA1, AFP, GPC3, LDHB, APOC1, APOA1, PABPC1, RARRES2
6	233	SPP1, CCL20, LCN2, CCL2, CXCL1, MMP7, SOD2, CXCL8, CXCL6, CRP
7	64	GCG, CHGB, TM4SF4, VGF, HLA-A, FXYD5, B2M, TTR, CLU, GPX3
8	37	KLK1, PHOX2A, AP1S2, MLLT11, LINC00682, KLK4, STMN4, NOVA1, TAGLN3, PDZRN3
9	23	GCA, ZNF672, HES6, RBP1, GC, PGA5, MFAP4, TDO2, NEUROG3, CCK
10	102	HMGB2, HIST1H4C, TUBA1B, KIAA0101, TOP2A, UBE2C, LDHB, CKS1B, HMGA1, CENPF
11	35	SST, IAPP, PPY, RBP4, AQP3, MT1X, MT2A, CHGB, SAMD11, TIMP1
12	69	PITHD1, MESP1, ARHGAP18, ITIH1, SCG2, CHGA, HIPK2, AC004540.4, EFNA5, CA2
13	89	GSTA1, AGR2, APOE, S100A10, RARRES2, PDLIM1, KLF5, LDHB, SAT1, CLDN6
14	7	INS, CRYBA2, TTR, CALB2, PYY, GAST, GCG
15	220	COL1A1, TIMP1, COL3A1, COL1A2, LGALS1, MT2A, TIMP3, CRYAB, CCL26, SPARC
16	111	RPL7, EIF3E, RPS14, RPL34, RPS6, RPL26, RPL35A, RPS3A, RPL10A, RPLP0
17	84	GHRL, ACSL1, BHMT, GPR160, VSTM2L, C10orf10, CMTM8, VTN, B3GNT7, WNK3
18	156	REG1A, REG3A, REG1B, PRSS2, CTRB2, PRSS1, SPINK1, CTRB1, C15orf48, CPA1
19	311	RGCC, INSR, SLC9A3R2, CLIC4, RAMP2, TCF4, CCDC85B, PGF, FKBP1A, HSPG2

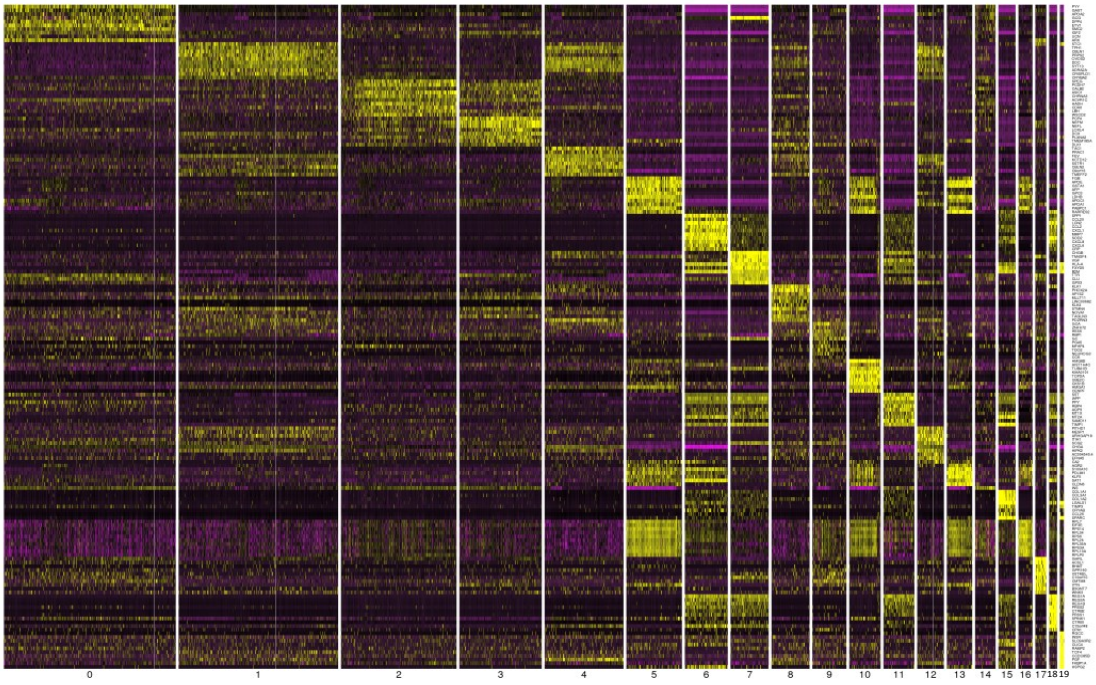


Figure 6 | Expression distribution of known cell type-specific markers on *t*-SNE projections and differentially expressed genes in each cluster.

a, Single-cell expression of remarkable cluster determinant markers in iPICs and human islets.

b, Most significant differentially expressed genes (DEGs), with a greater than 2-fold change and statistical significance of $p < 0.05$ as determined by the likelihood ratio test for single-cell expression.

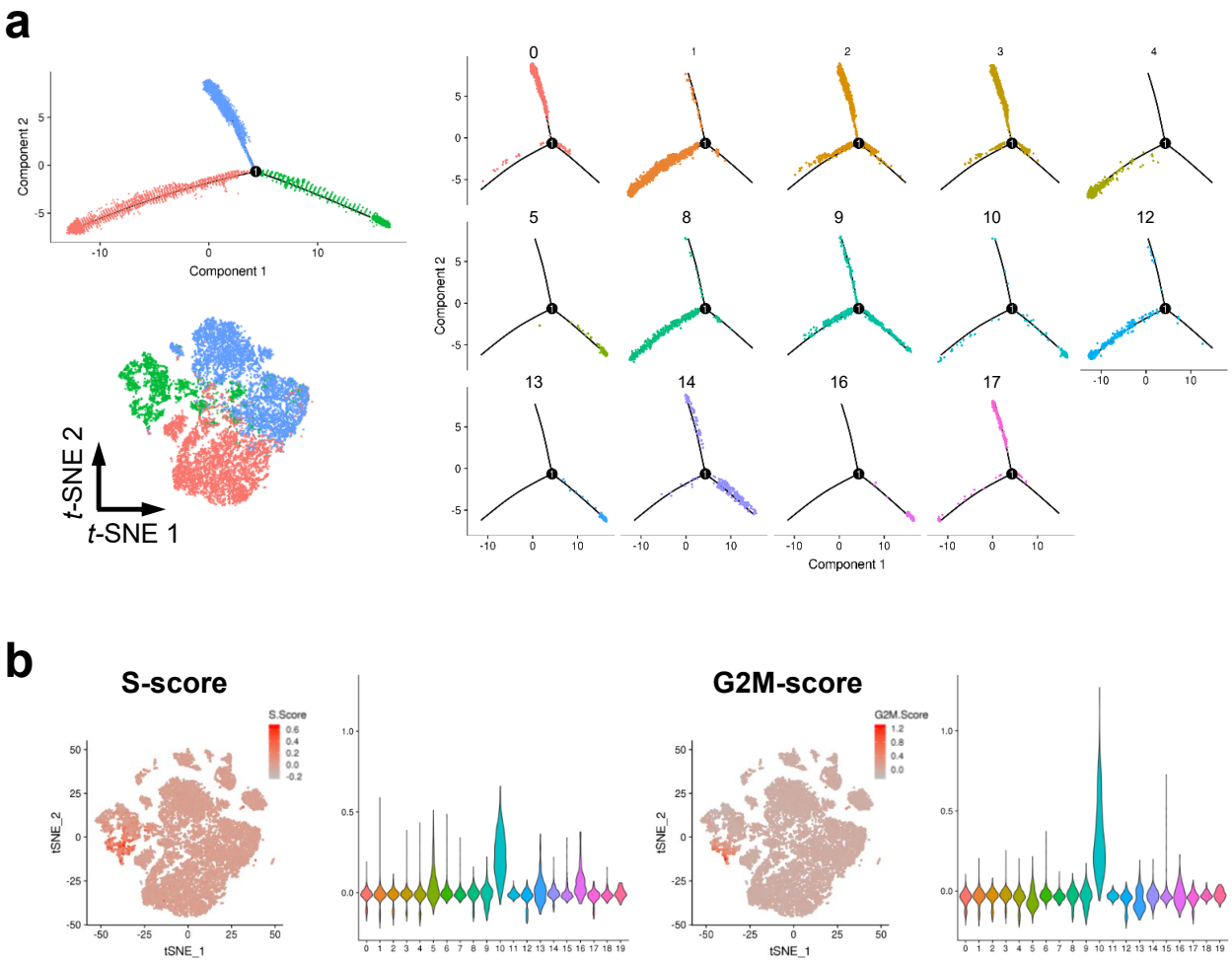


Figure 7 | Multifaceted analyses focusing on non-endocrine cells.

a, Single-cell trajectory for cells composing iPICs reconstructed with *Monocle 2* to describe the progression of the differentiation process. The upper left panel shows the trajectory projection of the reconstructed iPICs with one branching point. The right panels show the position of cells composing each cluster classified in Fig. 4d on the trajectory projection. The lower left panel shows the trajectory analysis results reflected on the *t*-SNE projection.

b, Cell cycle phase assignments based on S-phase, G2 and M gene signatures indicated a highly proliferative population in red within the combined *t*-SNE plot and upward shifts in the violin plot.

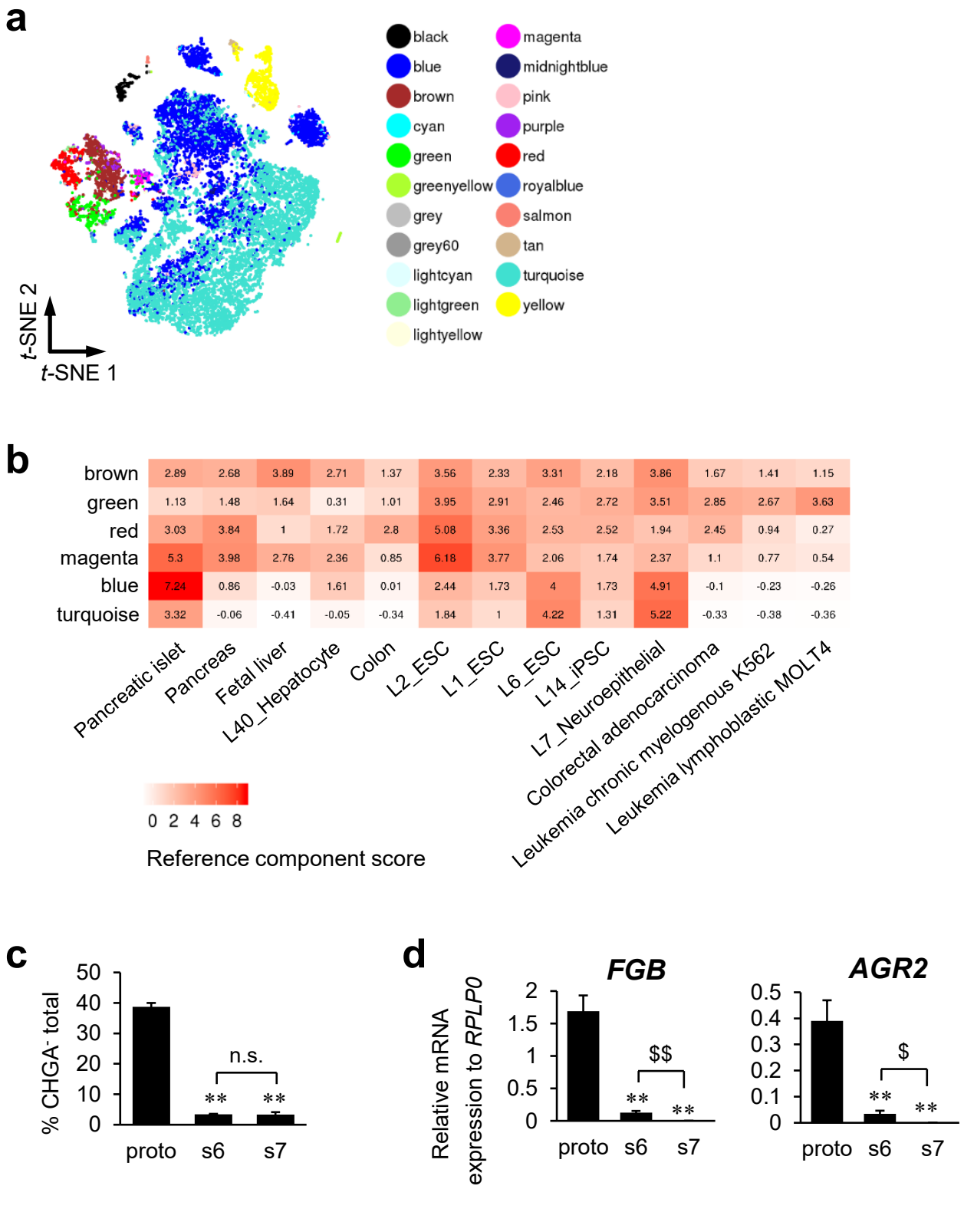


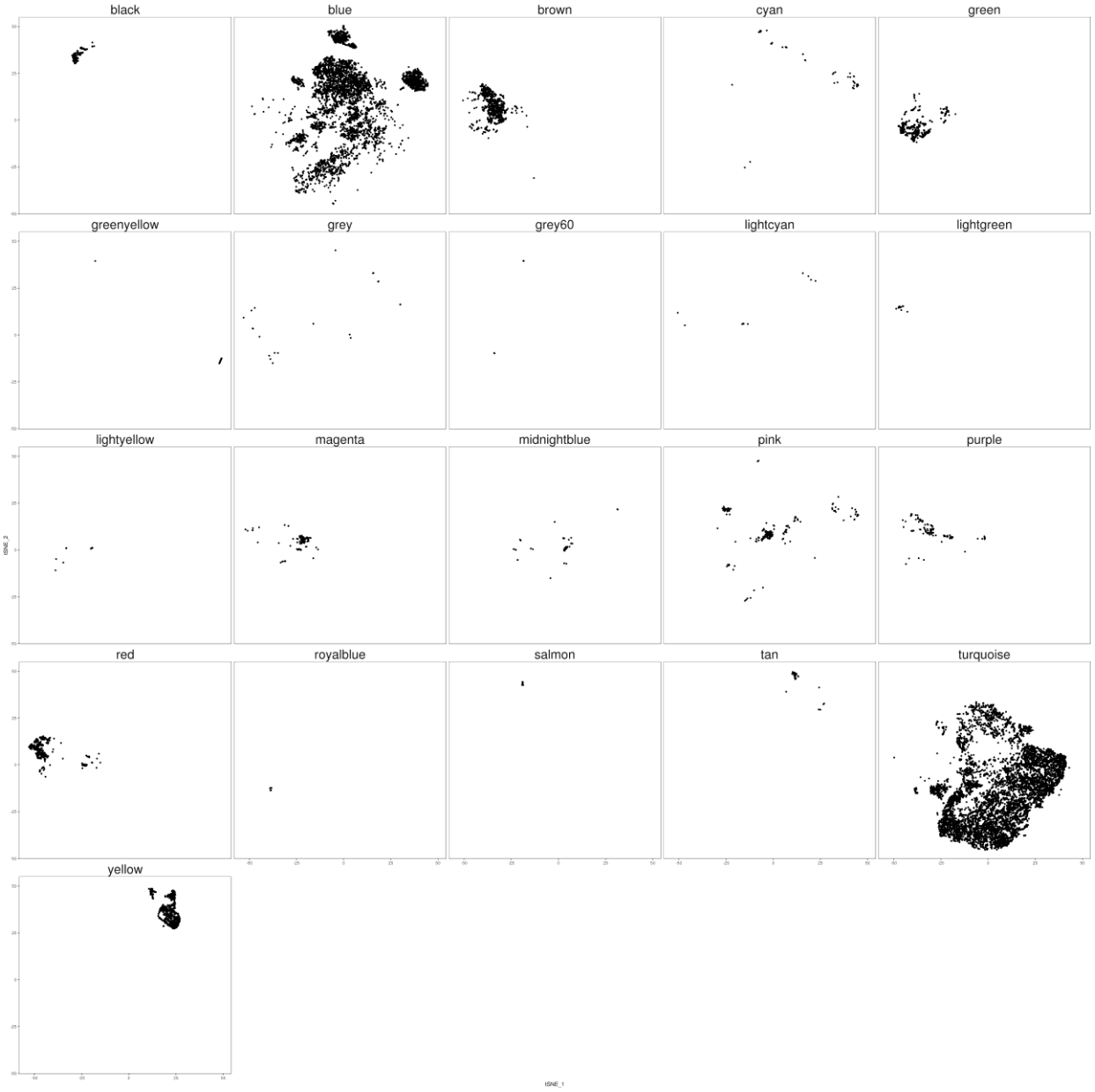
Figure 8 | Detailed analysis focusing on non-endocrine cells and validation of identified markers for detecting non-endocrine subpopulations.

a, Newly classified cell populations by reference component analysis (RCA) on the *t*-SNE projection. **b,** A heatmap of tissues and cell lines with high similarity (reference component score >2.5) to non-endocrine cells (brown, green, red and magenta). See Figure 9b for the full RCA heatmap.

c, The proportion of putative non-endocrine cells ($CHGA^-$) for differentiated prototype, s6-iPIC and s7-iPIC by flow cytometry analysis. Data are shown as the mean ± SD (n = 4, technical replicates). Reproducibility confirmed by 3 independent experiments. $^{**}P < 0.01$ versus s6-iPIC, Dunnett's test. n.s.; not significant, Aspin-Welch *t*-test.

d, Validation of the identified novel non-endocrine markers in quantitative real-time PCR experiments. Expression levels were normalized to *RPLP0* and are shown as the mean ± SD (n = 4, technical replicates). Reproducibility confirmed by 3 independent experiments. $^{**}P < 0.01$ versus s6-iPIC, Dunnett's test. $^{\$\$}P < 0.01$, $^{\$}P < 0.05$, Student t-test.

a


black
blue
brown
cyan
green
greenyellow
grey
grey60
lightcyan
lightgreen
lightyellow
magenta
midnightblue
pink
purple
red
royalblue
salmon
tan
turquoise
yellow
tSNE_2
tSNE_1

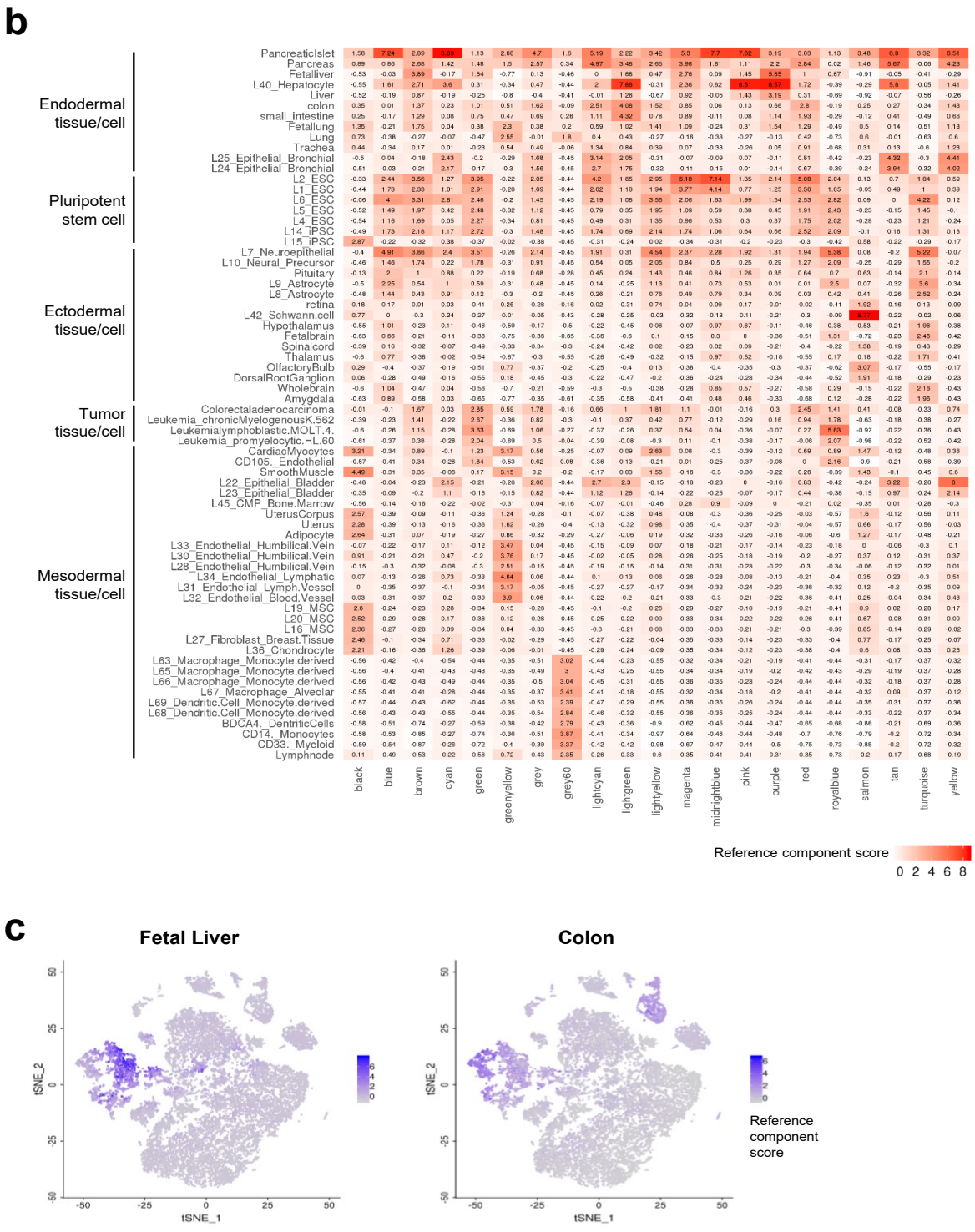


Figure 9 | Additional information for RCA.

a, Distribution of each color in RCA on the *t*-SNE projection.

b, Heatmap of similar tissues or cell lines for each cluster classified by RCA.

c, Cellular distribution of reference component scores for fetal liver and colon references.

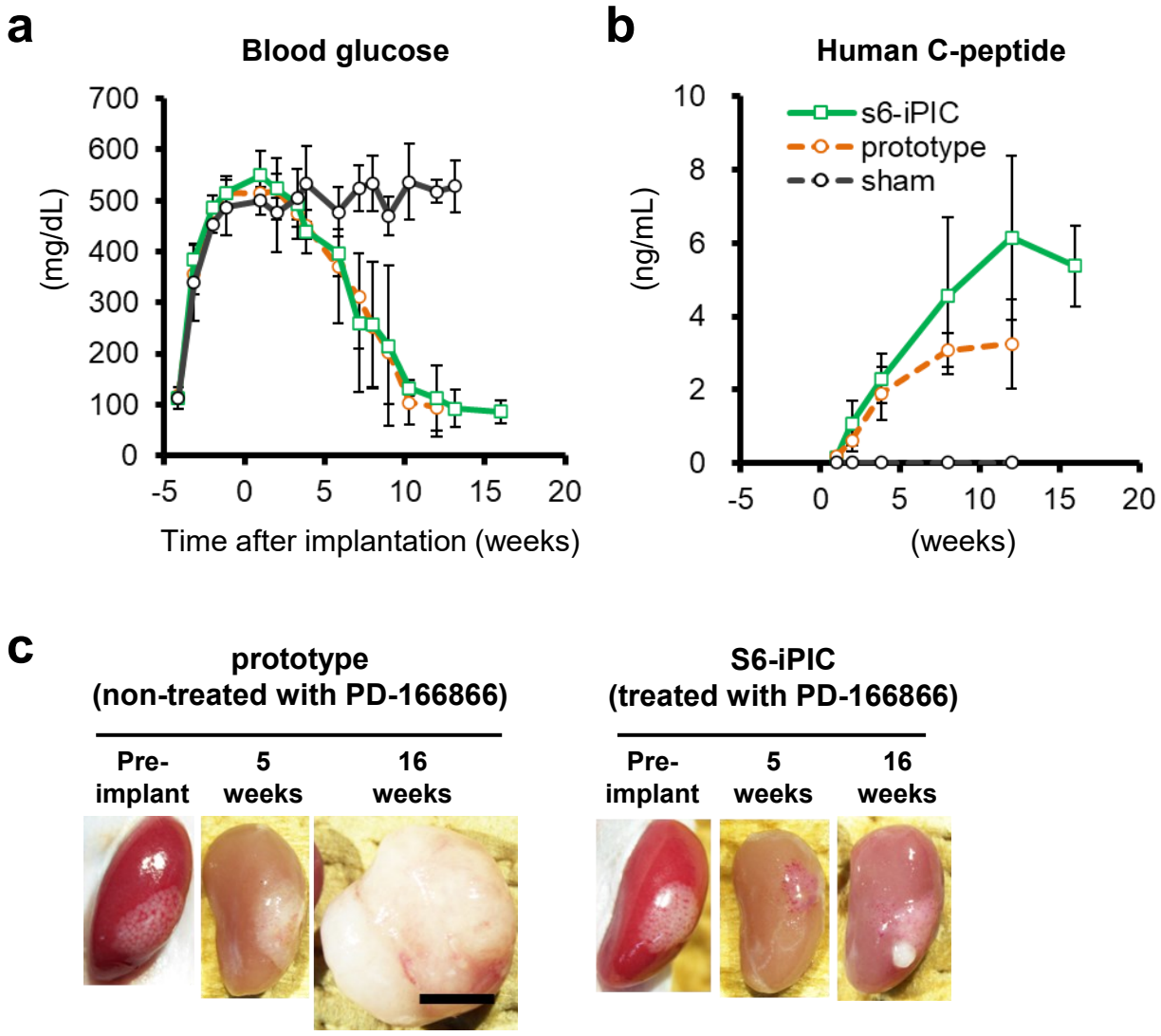
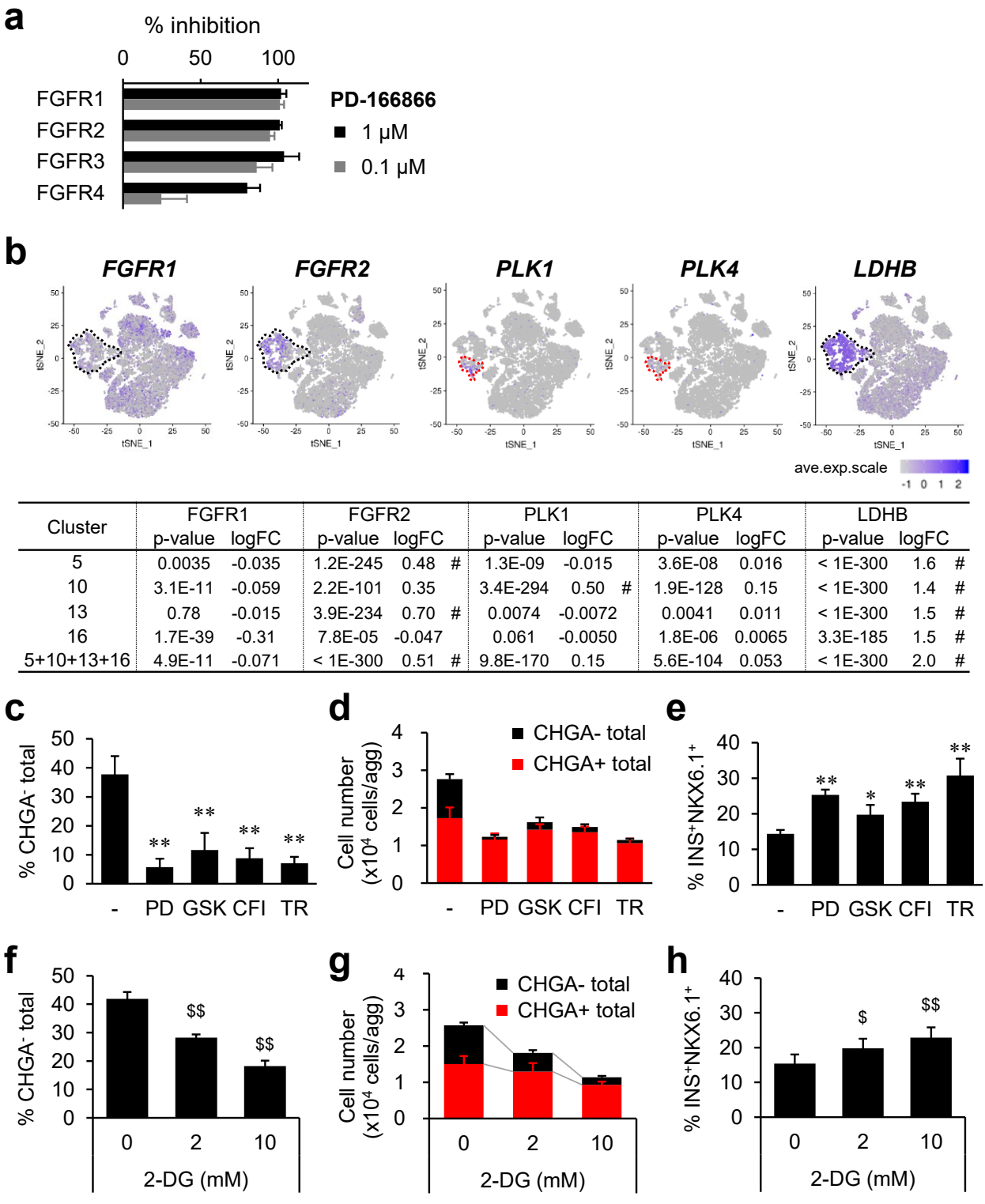


Figure 10 | Kidney capsular implantation of prototype and s6-iPIC.

a and **b**, Blood glucose and plasma human C-peptide levels after prototype (3.6×10^6 cells/mouse) or s6-iPIC (1.4×10^6 cells/mouse) implantation. Data are shown as the mean ± SD (prototype; n = 4⟶3, s6-iPIC; n = 4⟶3, sham; n = 5⟶4). The decrease in n number is due to unexpected death.

c, Macroscopic changes of grafts post implantation; n = 1 at each time point. The graft of s6-iPIC showed slight hypertrophy at 16 weeks post implantation. In contrast, the graft of prototype gradually became larger than the kidney at 16 weeks post implantation. Reproducibility was confirmed in several similar experiments. Scale bar, 5 mm.



Cluster	FGFR1		FGFR2			PLK1			PLK4		LDHB		
	p-value	logFC	p-value	logFC		p-value	logFC		p-value	logFC	p-value	logFC	
5	0.0035	-0.035	1.2E-245	0.48	#	1.3E-09	-0.015		3.6E-08	0.016	< 1E-300	1.6	#
10	3.1E-11	-0.059	2.2E-101	0.35		3.4E-294	0.50	#	1.9E-128	0.15	< 1E-300	1.4	#
13	0.78	-0.015	3.9E-234	0.70	#	0.0074	-0.0072		0.0041	0.011	< 1E-300	1.5	#
16	1.7E-39	-0.31	7.8E-05	-0.047		0.061	-0.0050		1.8E-06	0.0065	3.3E-185	1.5	#
5+10+13+16	4.9E-11	-0.071	< 1E-300	0.51	#	9.8E-170	0.15		5.6E-104	0.053	< 1E-300	2.0	#

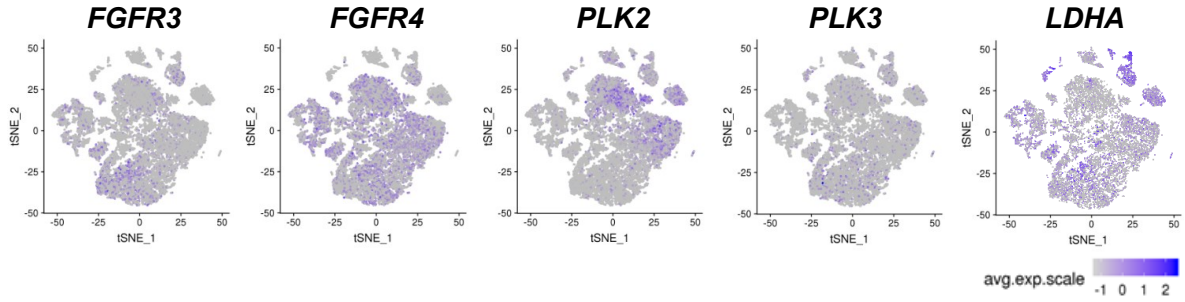
Figure 11 | Reduction of the number of non-endocrine cells by inhibition of non-endocrine specific factors.

a, Inhibitory activities of PD-166866, known as a selective FGFR1 inhibitor, for all four FGF receptor isoforms in the TR-FRET-based competitive binding assay. Data are shown as the mean ± SD (n = 4, collected from independent experiments).

b, Single-cell gene expression analysis showing that *FGFR1*, *FGFR2*, *PLK1*, *PLK4* and *LDHB* were specifically expressed in non-endocrine clusters. The dotted line in black represents the non-endocrine population from the iPICs. The dotted line in red represents proliferative cluster 10. The table contains the *p*-value and fold change for each indicated non-endocrine cluster compared to the others. # for *p*-values less than 0.05 and fold changes more than 1.5. See Figure 12a for single cell gene expression of other *FGFR* and *PLK* isoforms.

c-h, Evaluation of the non-endocrine cell reduction effects of kinases and glycolytic inhibitors. S6-iPIC cells were additively treated with the indicated inhibitors from day 4 at stage 6 for purification and then collected and analyzed for the ratios of putative non-endocrine cells ($CHGA^-$) (**c, f**), cell number (**d, g**) and the ratios of putative β-cells ($INS^+/NKX6.1^+$) (**e, h**). Data are shown as the mean ± SD (n= 4-5, collected from independent experiments). $^{**}P < 0.01$ versus s6-iPIC, Dunnett's test. $^{\$}P < 0.025$, $^{\$\$}P < 0.005$, versus s6-iPIC, one-tailed Williams' test. PD; PD-166866, GSK; GSK 461364, CFI; CFI-400945, TR; TR06141363, 2-DG; 2-deoxy-D-Glucose.

a



Cluster	FGFR3		FGFR4		PLK2		PLK3		LDHA	
	p-value	logFC	p-value	logFC	p-value	logFC	p-value	logFC	p-value	logFC
5	4.7E-17	0.019	5.9E-17	0.11	4.1E-70	-0.32	5.5E-18	-0.064	1.0E-61	-0.25
10	5.9E-24	-0.014	1.4E-10	0.056	2.5E-32	-0.30	3.9E-12	-0.062	1.4E-45	0.14
13	4.5E-05	-0.00073	2.9E-10	-0.081	3.4E-29	-0.28	1.7E-09	-0.058	6.6E-06	0.066
16	3.6E-10	-0.086	2.5E-24	0.0069	3.2E-21	-0.27	2.4E-11	-0.081	3.0E-05	0.037
5+10+13+16	6.3E-31	-0.0050	2.0E-10	0.052	1.1E-145	-0.33	3.7E-42	-0.071	2.5E-56	-0.046

b

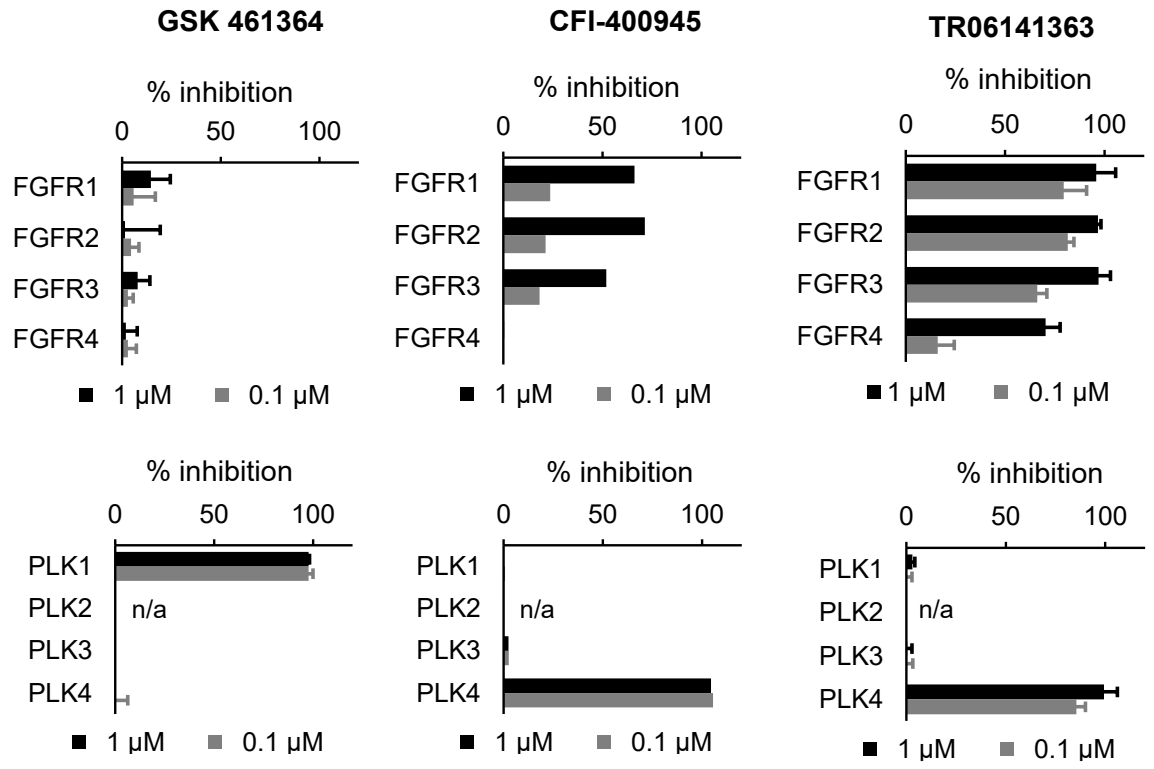


Figure 12 | Single-cell gene expression related to Figure 11b and inhibitory activities of GSK 461364, CFI-400945 and TR06141363 for FGFR and PLK isoforms.

a, Single-cell gene expression of *FGFR3*, *FGFR4*, *PLK2*, *PLK3* and *LDHA*. The table contains the *p*-value and fold change in each indicated non-endocrine cluster compared to the others.

b, Inhibitory activities of GSK 461364, CFI-400945 and TR06141363 for FGFR and PLK isoforms in the TR-FRET-based competitive binding assay. Data are shown as the mean ± SD (n = 3-4, collected from independent experiments), the mean without SD bar (n = 2, collected from independent experiments) or raw value (n = 1). n/a; not available.

5. Discussion

Off-target cells account for a small proportion of iPSC-derived islet cells, but their characteristics are significant for cell therapy against type 1 diabetes. In the present study, three types of iPICs with different degrees of induction efficiency were analyzed, revealing the potential remaining off-target cells in s7-iPIC. Single-cell dataset categorized off-target cells into 1) $PDX1^{+}$ pancreatic-lineage non-endocrine cells such as pancreatic duct-like cells and pancreatic progenitor cells (Cluster 5, 13, 16 and most of Cluster 10), and 2) heterogeneous proliferating cells (a small part of Cluster 10) (Fig. 4b, 4e, and 8a). The former cell population was fundamentally pancreatic, but partially retained traits of either liver or intestinal lineages; these are spatiotemporally close to the pancreatic region during development. It is reasonable to attribute the presence of these cells to artificial stepwise differentiation *in vitro*, as the *in vitro* environment probably lacks unknown factors that are needed to stabilize the pancreatic developmental process *in vivo*. Nonetheless, it is worth noting that maturation to liver or intestinal cell types was not apparent in implanted s7-iPIC grafts (Fig. 1e and f), indicating that there was little contamination with cells displaying multiple tissue traits or that the cells are prone to lose cell fates other than pancreas, at least on the scale of this study.

The presence of non-endocrine off-target cells has been observed in previous studies using scRNA-seq analysis (14, 17, 22). In some cases, the cells actively proliferated to enlarge the graft massively after implantation (11, 16). To reduce the number of non-endocrine cells, there are two approaches: concentrating only endocrine cells using cell sorting; and reducing non-endocrine cells through compound treatments or metabolic modification. The former approach, which assesses cells one by one, is useful for handling cells at a small scale (49). Regarding the latter approach, chemical treatment, such as a yes-associated protein (YAP) inhibitor, reduces the number of pancreatic progenitors via increased efficiency of endocrine cell induction (18). This approach targets multiple cells concomitantly, in principle, and may be applicable for handling cells on a large scale. In the present study, evidence was provided that non-endocrine off-target cell populations differentially express some genes, such as *FGFR2*, *PLK1/4* and *LDHB* (Fig. 11b), and that targeting the function of these genes with chemicals selectively reduced the number of non-endocrine cells (Fig. 11d and g). Because each of these methods is based on a different mechanism of action, combining previously reported methods with the findings obtained in this study may improve the cell quality in a large-scale culture.

PD-166866 selectively reduced the number of non-endocrine off-target cells mainly via the inhibition of FGFR2 in non-endocrine cells (Fig. 11a-e and 12a). PD-166866 has been reported to

increase *NGN3* expression via FGFR1 inhibition when added at the time of induction of pancreatic progenitor cells into endocrine cells (46). In this study, the timing of PD-166866 treatment was different from that in the previous report, and PD-166866 was added after 4 days of induction of endocrine cells with a γ-secretase inhibitor (Fig. 2). PD-166866 treatment during this period specifically decreased the number of non-endocrine cells without increasing the number of $CHGA^+$ endocrine cells (Fig. 11d). Therefore, the effects of PD-166866 seen in the present study are probably different from the previous one and mediated directly by FGFR2 inhibition in non-endocrine cells.

In the PLK family, PLK1 is a well-studied molecule that is known to control the progression of the M phase in the cell cycle process (50). In this study, *PLK1* was highly expressed in Cluster 10, in which the cell cycle was ongoing (Fig. 11b and 7b), and treatment with a selective PLK1 inhibitor (GSK 461364) reduced the number of cells in Cluster 10, as expected (Fig. 11d). Unexpectedly, the PLK1 inhibitor also reduced the number of non-endocrine cells in Clusters 5 and 13, in which *PLK1* was hardly expressed (Fig. 11b and d). The expectation is that the non-biased scRNA-seq analysis classified cells in which the proliferation process was ongoing into Cluster 10. However, since the cells in Clusters 5 and 13 were more potent proliferation than endocrine cells, the cells entered the proliferation process during the 7-day treatment with the PLK1 inhibitor, resulting in the cells becoming targets of the PLK1 inhibitor. Thus, the PLK isoform inhibition is effective for removing proliferating and highly proliferative cells, in other words, is effective for selecting low-proliferative cell types.

Metabolic selection is an approach that removes off-target cells based on the difference in energy sources among cell types. Recent reports have demonstrated that metabolic selection reduces off-target cells, mainly composed of incompletely differentiated cells, in directed differentiation from pluripotent stem cells, such as cardiomyocytes, neural progenitors and even in definitive endoderm (51-53). The findings of this study indicated that the differential expression of *LDHB* in non-endocrine cells led to sufficient function because the number of non-endocrine cells was preferentially reduced by treatment with 2-DG (Fig. 11b, f and g). As inhibition of glycolysis is one of the approaches to remove proliferative cancer cells (54), glycolysis inhibition has the potential to be an approach to remove unintended cells in directed differentiation towards less proliferative terminally differentiated cells.

Three different approaches selectively reduced non-endocrine cells, namely inhibition of FGFR2, PLK1/4, and glycolysis. Through inhibition of these factors, five compounds (PD-166866, GSK 461364, CFI-400945, TR06141363, and 2-DG) showed specific reduction of non-endocrine cells

(Fig. 11c-h). PD-166866 and TR06141363 were used to induce s7-iPIC; implantation results of iPICs suggested that these two compounds eliminated most potential non-endocrine cells while maintaining adequate *in vivo* efficacy (Fig. 1d-k and 10a-c). Therefore, PD-166866 and TR06141363 treatment have clinical application potentials. However, scRNA-seq data revealed that non-endocrine cells remained in s7-iPIC (Clusters 10 and 16), although very few (Fig. 5b). This slight difference could have a significant impact in clinical applications, which require $>10^8$ order cells. When removal of non-endocrine cells is insufficient in s7-iPIC in an over 10^8 cell scale, further removal of these off-target cells might be achieved by adapting PLK or glycolysis inhibition strategy. The PLK inhibition strategy may have little impact on *in vivo* efficacy since GSK 461364 and CFI-400945 had little effect on endocrine cell number (Fig. 11d). This is also supported by the fact that s7-iPIC differentiated following TR06141363 treatment, which has PLK4 inhibitory activity, showed sufficient *in vivo* efficacy (Fig. 1g-j and 12b). Therefore, the PLK inhibition strategy is the first choice as additional removal of non-endocrine cells. In particular, GSK 461364 is promising due to its PLK1 inhibitory activity, which was not covered by TR06141363 (Fig. 12b). For the glycolysis inhibition strategy, there is a limitation in using 2-DG, as observed in the current study: 2-DG treatment reduced not only non-endocrine cells but also endocrine cells, although the reduction rate of endocrine cells was smaller than that of non-endocrine cells (Fig. 11g). This is likely since 2-DG dose not target LDHB (lactate dehydrogenase B), which was specifically expressed in non-endocrine cells (Fig. 11b). LDHB subunit-specific inhibitors are expected to reduce the number of non-endocrine cells without affecting endocrine cells, although further studies are warranted.

6. Summary

This chapter highlighted $CHGA^-$ non-endocrine off-target cells, which are contaminating *in vitro*-generated pancreatic islet-like cells by directed differentiation from pluripotent stem cells. It was revealed that $CHGA^-$ non-endocrine cells mostly consist of $PDX1^+$ pancreatic-lineage cells such as pancreatic duct-like cells and pancreatic progenitor cells. In addition, novel approaches for non-endocrine cell specific reduction were investigated, such as inhibiting FGFR2, PLK1/4, and glycolysis, which were predominantly activated in non-endocrine cells. Among the non-endocrine cell reduction candidates, PD-166866 and TR06141363 were used to induce s7-iPIC, and these compounds reduced most of the potential non-endocrine off-target cells while maintaining adequate *in vivo* efficacy. Although the iPIC safety assessment was inconclusive due to an insufficient *in vivo* sample size of this chapter, findings in this chapter could contribute to mitigating the safety risks of iPICs for future clinical applications that require cell manufacturing on a large scale.

Chapter 2.

Identification and removal of unexpected proliferative off-target cells emerging after iPSC-derived pancreatic islet cell implantation

1. Introduction

As shown in chapter 1, the majority of off-target cells, which are contaminated in the induction process of pancreatic endocrine cells from iPSCs/ESCs, are non-endocrine cells of the pancreatic lineage. Non-endocrine off-target cells were significantly reduced in originally developed s7-iPICs that are generated with treatment of kinase inhibitors such as PD-166866 and TR06141363. When s7-iPICs were implanted into T1DM model mice in the order of 10^6, no graft abnormalities, such as cyst-like enlargement due to proliferation of pancreatic lineage off-target cells, were observed. However, for clinical application in patients with T1DM, induced cells are required in the order of 10^8–10^9 (20). Therefore, it was deemed necessary to investigate the long-term safety of s7-iPICs using large scale experiments.

Implantation studies in mice generally use 10^6 cells (9-11, 17), and implantation at higher cell densities may result in lower cell survival and an underestimation of the risk of off-target cells. Thus, in this chapter, multiple implantation experiments were performed with s7-iPICs in the order of 10^6 to evaluate cumulative cells in the order of 10^8 or higher. Consequently, novel proliferative off-target cells, which were cells outside the pancreatic lineage, were unexpectedly observed. As the unexpected off-target cells could pose a major risk for clinical applications, the purpose of this study was to identify the characteristic profiles of this unexpected off-target cell population to propose an effective detection system and reduction strategy.

2. Materials and methods

2.1 Cell culture and s7-iPIC differentiation

Maintenance culture of undifferentiated iPSCs and stepwise differentiation into s6-iPICs and s7-iPICs were basically conducted as described in the method of chapter 1. Only the following details are different.

- During the differentiation period from Stage 1 to Stage 6, the cells were cultured in a spinner-type 30-mL bioreactor (ABLE Corporation & Biott) or a vertical mixing 0.25-L bioreactor (SATAKE MultiMix Corporation).
- To generate s7-iPICs for implantation, cells were dissociated and re-sized ($1.0–1.4 \times 10^3$ cells/microwell) in an Elplasia microwell plate (Corning Incorporated) or a gas-permeable microwell culture bag (Toyo Seikan Group Holdings) at the start of Stage 7.
- When inducing docetaxel-treated s7-iPICs, 1 μM docetaxel (FUJIFILM Wako) was added from day 4 of Stage 6.

2.2 Type 1 diabetes mouse model

Type 1 diabetes model was prepared as described in the method of chapter 1. All animal studies were conducted at Shonan iPark, an AAALAC international accreditation facility, and approved by the iPark Institutional Animal Care and Use Committee. Animal experiment approval numbers: AU-00011795, AU-00020231, AU-00020664, and AU-00021121. All experiments were performed in accordance with the relevant guidelines and regulations, including the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines.

2.3 Implantation and in vivo assessment

Implantation and *in vivo* assessment were basically conducted as described in the method of chapter 1. Only the following details are different.

- Number of implanted cells: $2.0–4.5 \times 10^6$ cells/mouse
- For implantation using alginate gel (Fig. 15A and B), cell aggregates were suspended in 3% alginate (NovaMatrix) in 25 mM HEPES buffer, cross-linked with 75 mM strontium chloride hexahydrate (FUJIFILM Wako) in 25 mM HEPES buffer, and formed into discs.

2.4 Tissue processing and immunostaining

Grafts were collected 23–30 weeks after implantation. Graft retrieval, tissue fixation,

sectioning, and staining were conducted as described in the method of chapter 1. The primary antibodies are listed in Table 2. To determine the presence or absence of proliferative MSC- and SMC-like cells (PMSCs) and cysts, sections were prepared to ensure that the cross-sectional area of the graft was as large as possible and performed pathological and morphological determination by hematoxylin and eosin (HE) staining. Considering that PMSCs and cysts have characteristic structures, they can be determined by HE staining. However, in cases such as docetaxel-treated s7-iPICs, where neither PMSCs nor cysts were present or could not be determined by HE staining alone, a definitive diagnosis was performed using immunostaining (insulin, glucagon, PDX1, HuN, Ki67, α-smooth muscle actin (αSMA), and cytokeratin 19 (CK19)). The absence of PMSCs and cysts was determined if the following three conditions were met: i) αSMA- and CK19-positive areas were absent, ii) areas with a high percentage of Ki67-positive cells were absent, and iii) most HuN-positive areas comprised $PDX1^{+}$/$insulin^{+}$/$glucagon^{+}$ cells.

Table 2. List of primary antibodies used for immunofluorescence staining.

Antigen	Species	Manufacturer	Catalog #	Clonality	Dilution for immunohistochemistry	Dilution for flow cytometry
INS	Rabbit	CST	3014	Monoclonal (C27C9)	1 : 200	-
INS	Rat	DSHB	GN-ID4	Monoclonal (GN-ID4)	-	1 : 600
GCG	Mouse	Sigma	G2654	Monoclonal (K79bB10)	1 : 200	-
PDX1	Goat	R&D systems	AF2419	Polyclonal	1 : 200	1 : 200
NKX6.1	Rabbit	CST	54551S	Monoclonal (D8O4R)	-	1 : 600
CK19	Mouse	Dako	M0888	Monoclonal (RCK108)	1 : 200	
Ki67	Rabbit	CST	9129	Monoclonal (D3B5)	1 : 100	-
Ki67	Mouse	BD Biosciences	556003	Monoclonal (B56)	-	1 : 100
HuN	Mouse	Takara Bio	Y40400	Monoclonal	1 : 100	-
IGF1	Mouse	Merck Millipore	05-172	Monoclonal (Sm1.2)	1 : 100	-
α-SMA	Mouse	Abcam	ab7817	Monoclonal (1A4)	1 : 500	1 : 6000
Vimentin	Rabbit	CST	5741S	Monoclonal (D21H3)	1 : 100	-
Collagen Ⅲ	Rabbit	Abcam	ab7778	Polyclonal	1 : 50	-
SFRP4	Rabbit	Abcam	ab154167	Monoclonal (EPR9389)	1 : 50	-
CHGA	Rabbit	Abcam	ab68271	Monoclonal (EP1030Y)	-	1 : 500

2.5 Single-cell RNA sequencing library preparation, sequencing, and data processing

A total of seven samples (one sample of Vitro s6-iPICs, two samples of Vivo s6-iPICs, one

sample of reference human islets, one sample of Vitro s6-iPICs cultured without PD-166866, and two samples of s6-iPICs after 4 weeks of extended culture, including EGF treatment) underwent scRNA-seq. Vivo s6-iPIC samples were de-crosslinked using 100 mM sodium citrate solution. Human islets were purchased from PRODO (HP-18304-01, Donor age: 21 years, Donor sex: Male, Donor body mass index: 27.3 kg/m^2, Donor HbA1c: 5.4%, Estimated purity: 85%, Estimated viability: 95%). Library preparation for scRNA-seq, sequencing, and data processing were basically conducted as described in the method of chapter 1. Only the following details are different.

- The sequencing reads of the *in vivo* s6-iPIC samples were first aligned to the human GRCh38 and mouse mm10 genome references. Cells with at least one UMI count of the mouse genes were collected. Thereafter, the sequencing reads were aligned to the human GRCh38 genome reference and the collected cells containing mouse genes were removed for further analyses.
- Principal components were used for Seurat's shared nearest neighbor graph clustering, and t-distributed stochastic neighbor embedding (*t*-SNE) or uniform manifold approximation and projection (UMAP) dimensional reduction were used to visualize the data. The similarity between clusters was examined using hierarchical clustering analysis of average expression.
- For pseudotime analysis, the processed UMI count matrices were imported into a single-cell dataset for the monocle package (29-31). The genes for ordering the cells were selected using 'dpFeature' in monocle, to set Cluster 15 as the root (pseudotime was zero), and the single-cell pseudotime was conducted via the 'DDRTree' algorithm. Differential gene expression analyses along with pseudotime were performed using the likelihood-ratio test in monocle, and genes with q-values less than 0.05 were identified as differentially expressed genes along with pseudotime. Differentially expressed genes were classified into gene clusters using pseudotime expression patterns, and functional enrichment analysis of each gene cluster was performed using the clusterProfiler package (55).
- Previously reported scRNA-seq data using *in vivo* grafts of iPSC/ESC-derived islet-like cells were obtained from publicly available database (GSE151117). The UMI count matrices from three ESC (HUES8)-derived islet-like cell grafts (GSM4567001, GSM4567002, and GSM4567003) and two iPSC ($WS4^{corr}$)-derived islet-like cell grafts (GSM4567004 and GSM4567005) were imported into Seurat package and low-quality cells were removed. After normalization and scaling the data, the principal component analysis was performed, and clustering and dimensional reduction were conducted.

2.6 Extended culture

The induced s6-iPICs were cultured for 4 weeks in MCDB 131 medium with 1% P/S, 2% fat-free BSA, 14.44 mM glucose (added to generate a final concentration of 20 mM), 1.5 g/L $NaHCO_3$, and 1% GlutaMAX, 0.5% ITS-X. 50 ng/mL EGF was added at the start of the extended culture as needed. The same extended culture was performed on cells treated with 2 μM R428, 1 mM N-acetylcysteine, 10 μM Trolox, 3 μM lenvatinib mesylate (Carbosynth Limited), 3 μM cisplatin (FUJIFILM Wako), 1 μM docetaxel, 10 μM cyclophosphamide monohydrate (FUJIFILM Wako), 1 μM tamoxifen citrate (Merck Millipore), 10 μM anastrozole (AstaTech), or 0.3 μM lapatinib ditosylate (LKT Laboratories, Inc.). These compounds were added from day 4 of Stage 6 until the end of the stage. The concentrations of each compound that did not affect s6-iPIC induction or cell number in prior experiments were selected (data not shown).

2.7 Flow cytometry

Differentiation efficacy and quality analysis by flow cytometry was conducted as described in the method of chapter 1. The primary antibodies are listed in Table 2.

2.8 Plasma glucose and hormone measurements

Plasma glucose, human C-peptide, and mouse C-peptide levels were measured as described in the method of chapter 1.

2.9 Statistical analysis

Data are expressed as the mean and SD values. Dunnett's multiple-comparison test was performed based on the results of the homogeneity of variance test (Bartlett’s test) at a significance level of $P < 0.05$, as shown in Fig. 26C, 26E, and 27B. Additionally, the Aspin–Welch test was performed at a significance level of $P < 0.05$ to determine the statistical significance between the two groups (Fig. 13G and 26E). As shown in Fig. 29K, Fisher's exact test was performed for each of the two prioritized tiers (Tier 1: s7-iPICs vs. docetaxel-treated s7-iPICs; Tier 2: s7-iPICs vs. docetaxel-treated s6-iPICs), and adjustment for multiple testing was performed using the Bonferroni method. As shown in chapter 1, since evaluation at single cell resolution showed no prominent batch effect, the *in vivo* implantation data from multiple induction batches were combined for statistical analysis (Fig. 13G and 29K). All statistical analyses were performed using the Statistical Analysis System version 9.3 (SAS Institute, NC, USA).

2.10 Data and code availability

Data supporting the findings of this study are available from the author upon reasonable request. Single-cell RNA sequencing data were deposited in the Gene Expression Omnibus database (GSE213617). The computer code is available upon reasonable request.

3. Results

3.1 Unexpected abnormal outgrowth in s7-iPIC grafts is outside the pancreatic lineage and continues to proliferate after implantation

To evaluate the long-term safety profile of s7-iPICs on a large scale, experiments were repeated in which fibrin gel-embedded s7-iPICs in the order of 10^6 were subcutaneously implanted into immunodeficient mice (Fig. 13A). Similarly, s6-iPICs, which were cultured for the same period under Stage 6 conditions instead of receiving Stage 7 treatment, were also evaluated. At 23–26 weeks after implantation, most s7-iPIC grafts were composed of endocrine cell clusters, host-derived blood vessels, and host-derived fibrous tissue, and did not show abnormal cell proliferation (Fig. 13B). However, s7-iPIC grafts with unexpected abnormal outgrowth were infrequently observed (13%, 12/96 mice) (Fig. 13B and C). In addition, s6-iPIC grafts showed abnormal outgrowth more frequently than s7-iPIC grafts (84%, 38/45 mice) (Fig. 13C). Although the composition of the cell population was the same in s6-iPIC and s7-iPIC at single-cell resolution, the content ratio of each population was different as shown in chapter 1. In s6-iPICs, the content ratio of the non-endocrine progenitor cell population was reduced by PD-166866 treatment but was slightly higher than that in s7-iPICs. Therefore, the frequency of abnormal outgrowth appears to be associated with the degree of residual non-endocrine progenitor cells.

To characterize the abnormal outgrowth, immunostaining was performed; however, neither pancreatic endocrine cell markers (PDX1, insulin, and glucagon) nor markers of pancreatic duct-like cystic structures (PDX1 and CK19) were detected (Fig. 13D and E). In contrast, the abnormal outgrowth was stained with a human nuclear antibody (HuN), indicating that the structure was derived from the implanted s7-iPICs, and expressed the proliferation marker Ki67 at a high frequency (Fig. 13F). Furthermore, several grafts with abnormal outgrowth weighed more than 3-fold (> 60 mg) the average weight (15.8 ± 4.4 mg) of the grafts without abnormal outgrowth (Fig. 13G and 14A). These results suggest that unexpected cells constituting the outgrowth are outside the pancreatic lineage, and that they continue to proliferate and gradually become apparent after implantation. Although the frequency of abnormal outgrowth in s7-iPIC grafts appeared to be relatively low (13%, 12/96 mice) (Fig. 13C), this was the frequency when implanted in the order of 10^6; and as cells in the order of 10^8–10^9 are required for patients with T1DM (20), this abnormal proliferation is assumed to be a major risk for clinical application.

3.2 The unknown cell population that cannot be classified as known cell types in single-cell analysis could be responsible for abnormal outgrowth

To understand and eliminate abnormal outgrowth, scRNA-seq was conducted to extract and profile the cells constituting abnormal outgrowth using. Considering the infrequency of abnormal outgrowth in s7-iPIC grafts, s6-iPICs were used to increase the possibility of obtaining the cells responsible for abnormal outgrowth (Fig. 15A). In addition, to prevent reduction in sequencing depth due to contamination by host-derived cells, s6-iPICs were embedded in alginate gel, which is a non-biodegradable material (56, 57). Two and six months after subcutaneous implantation, the grafts could be separated from the host mouse while maintaining the shape of the alginate gel (Fig. 15B). In these *in vivo* samples, the number of cells excluded during quality control by the mitochondrial gene ratio was limited, and host cell contamination was minimal (Fig. 16A and B). To extract the characteristics of the *in vivo* samples, their data were combined with sequence data from the cells immediately before implantation (Vitro s6-iPICs) and from adult human islets as references, and performed non-biased cluster classification (Fig. 15A, C, and D). As there was little difference in the cluster classifications of the *in vivo* samples at 2- and 6-months post-implantation (Fig. 16C), the two *in vivo* samples were collectively treated as Vivo s6-iPICs.

Based on previous reports (14, 17, 22), each cluster was assigned to a known cell type using characteristic gene expression as an indicator (Fig. 15D and 16D). As a candidate cell population responsible for abnormal outgrowth, Cluster 18 initially caught attention because it had the highest expression of the proliferation marker *MKI67* (Fig. 16D). However, Cluster 18 was included cells from all samples, including human islets, and the majority of cells in Cluster 18 had high S-score and G2M-score in cell cycle phase assignment analysis (Fig. 16C and E). Therefore, Cluster 18 was concluded to be a heterogeneous population of mitotic cells (Fig. 15D), and it was assumed that extracting the cells responsible for abnormal outgrowth from an in-depth analysis of Cluster 18 would be difficult. Hence, the focus was placed on clusters specific to Vivo s6-iPICs. Among the clusters derived from Vivo s6-iPICs, five (Clusters 0, 4, 7, 16, and 17) were classified as endocrine cells, and three (Clusters 6, 13, and 15) were classified as non-endocrine cells. In the three non-endocrine cell clusters, Clusters 6 and 13 were pancreatic duct-like cells expressing *KRT19*, *HIF1B*, and *SOX9*, and pancreatic acinar-like cells expressing *CTRB1* and *CTRB2*, respectively (Fig. 16D). Such pancreatic exocrine cell clusters were also naturally present in adult human islets (Clusters 6, 12, and 14) (Fig. 15D and 16D). In contrast, Cluster 15 showed no expression of pancreatic lineage markers and expressed mesenchymal stem cells (MSC)-related genes (17), such as *VIM*, *COL3A1*, and *COL1A1* (Fig. 16D).

In addition, Cluster 15 contained several cells with high S-score and G2M-score (Fig. 16E), leading to the hypothesize that Cluster 15 is responsible for abnormal outgrowth after implantation.

3.3 Abnormal outgrowth consists of proliferative MSC- and SMC-like cells (PMSCs)

To scrutinize the characteristics of Cluster 15, reference component analysis (RCA), which indicates transcriptome similarities with known tissues or cell lines (28), was performed (Fig. 17A). Cluster 15 was classified as "green" and had high scores for smooth muscle, uterus, MSCs, and iPSCs (Fig. 17A, 17B, and 18A). As stromal cells in adult human islets (Cluster 9) also belonged to the "green" group and showed MSC-related marker expression (Fig. 16D), the analysis proceeded by comparing these two clusters. To extract the differences between Clusters 15 and 9, the pseudotime was estimated from Cluster 15 to Cluster 9 (Fig. 19A). In addition, gene clusters composed of genes with common pseudotime expression kinetics were calculated, and gene ontology (GO) term analysis was performed (Fig. 19B and C). Compared to Cluster 9, Cluster 15 had relatively active pathways related to developmental processes (Gene Cluster 1), gene transcription processes (Gene Cluster 2), and mitotic cell cycle transition/catabolic processes (Gene Cluster 3). These results suggest that Cluster 15 is an immature and proliferative population, unlike true stromal cells (Cluster 9).

The top 20 differentially expressed genes, which were highly expressed in Cluster 15 but minimally expressed in the other clusters, were extracted (Fig. 17C). These 20 genes were then inputted into an external database, TissueEnrich (https://tissueenrich.gdcb.iastate.edu/) (58), and the uterus-related genes *OGN*, *IGF1*, and *SFRP4* were observed to be included (Fig. 20A and B). The expression of *OGN*, *IGF1*, and *SFRP4* was validated on the *t*-SNE projections, revealing that these uterus-related genes were specifically expressed in Cluster 15, whereas MSC-related genes, such as *VIM* and *COL3A1*, were expressed not only in Cluster 15 but also in Cluster 9 (Fig. 17D). In addition, Cluster 15 specifically expressed the gene for α-SMA, *ACTA2*, possibly reflecting smooth muscle traits (Fig. 17C and D). Tissue immunostaining was performed to evaluate whether the Cluster 15-specific markers were expressed at the protein level. Insulin like growth factor 1 (IGF1), α-SMA, and secreted frizzled related protein 4 (SFRP4) proteins were confirmed to be expressed in the abnormal outgrowth, together with MSC-related proteins Vimentin and Collagen III (Fig. 17E, 17F, and 21A). Based on the results thus far, it was concluded that the cell population in Cluster 15 constitutes the abnormal outgrowth emerging after implantation. This population exhibited traits associated with MSCs and smooth muscle cells (SMCs), including uterine characteristics. Accordingly, this *in vivo* population is hereafter referred to as proliferative MSC- and SMC-like cells (PMSCs).

To demonstrate that PMSCs are not a cell type specific to iPICs, previously reported data were analyzed. A previous report was highlighted that provided scRNA-seq data using *in vivo* grafts of islet-like cells derived from ES and iPS cell lines, which are different from the cell lines of iPICs (22). Reanalysis of the scRNA-seq data in this report revealed the presence of a cell population expressing PMSC-specific markers such as *VIM*, *COL3A1*, *OGN*, *IGF1*, *SFRP4*, and *ACTA2* in the grafts (Fig. 22A and B). Remarkably, this *in vivo* graft, including a PMSC-equivalent cell population, was not only derived from cell lines different from iPICs but also differed in the detailed induction and implantation methods (22). Hence, PMSCs are not a phenomenon specific to cell line, induction protocol, or implantation method but are commonly present in the grafts of iPSC/ESC-derived islet-like cells.

Notably, the PMSC-specific markers were not expressed before implantation (Vitro s6-iPICs) (Fig. 17D). To evaluate the origin of PMSCs, a hierarchical clustering analysis was conducted. Cluster 11 was identified as the nearest neighbor of PMSCs (Cluster 15) (Fig. 17G). Cluster 11 is the non-endocrine progenitor population, which is included in Vitro s6-iPICs as a by-product (Fig. 15D), most of which are expected to mature into pancreatic duct-like cells (Cluster 6) that form cysts after implantation as shown in chapter 1. Hierarchical clustering analysis suggested that cells in the non-endocrine progenitor population could not only mature into cysts, but also into PMSCs. This hypothesis is consistent with the higher frequency of PMSCs in s6-iPICs (Fig. 13C), which contain more non-endocrine progenitor cells than s7-iPICs.

3.4 An EGF-supplemented extended culture system exposes a cell population (putative PMSCs) that closely resembles in vivo PMSCs

The analyses thus far suggest that PMSCs likely originate from non-endocrine progenitor cells; however, their fate is not determined prior to implantation, and PMSCs only become apparent after several months of *in vivo* maturation. Thus, the exploration of PMSC removal methods requires implantation trials and months of growth and maturation, which results in an extremely low throughput. Therefore, a detection system was developed to predict the appearance of PMSCs without relying on implantation. Because an extended culture of pre-implanted cells *in vitro* can mimic the *in vivo* maturation of pancreatic endocrine cells to some extent (17, 59), it was considered that extended culture might induce cells with a profile similar to that of PMSCs (Fig. 23A). However, extended culture of s6-iPICs for 4 weeks in a simple basal medium resulted in dominant CHGA-positive endocrine cells and no obvious increase in off-target cells (Fig. 23B). Thus, it was noted that the addition of EGF receptor ligands, such as betacellulin and EGF, to the s6-iPIC-induction process,

resulted in a dose-dependent increase in the CHGA-negative non-endocrine population, which is likely the origin of PMSCs (Fig. 24A and B). Accordingly, an extended culture was attempted with EGF supplementation and a marked increase in the PDX1^{-}/CHGA^{-} population was observed (Fig. 23B). Furthermore, the PDX1^{-}/CHGA^{-} population contained cells expressing α-SMA, a PMSC specific marker (Fig. 23C).

To examine the relationship between the PDX1^{-}/CHGA^{-} population and *in vivo* PMSCs, the scRNA-seq data were reanalyzed by adding sequence data from EGF-supplemented extended culture samples. Similar to the flow cytometry results, the cells after extended culture could be classified into three types: *PDX1*$^{-}$/*CHGA*$^{-}$, *PDX1*$^{+}$/*CHGA*$^{-}$, and *CHGA*$^{+}$ (Fig. 23D and 25A-C). RCA revealed that the *PDX1*$^{-}$/*CHGA*$^{-}$ population after extended culture scored highly for smooth muscle, uterus, and MSCs, as well as *in vivo* PMSCs (Fig. 23E, 23F, and 25D). In addition, the *PDX1*$^{-}$/*CHGA*$^{-}$ population characteristically expressed *in vivo* PMSC-related genes such as *OGN* and *COL3A1*, and a portion of the population expressed *ACTA2* (Fig. 23G). These results suggest that the PDX1^{-}/CHGA^{-} population that emerges after EGF-supplemented extended culture is almost equivalent to the PMSCs that appear after implantation.

3.5 Cisplatin and docetaxel effectively remove putative PMSCs through mechanisms other than kinase inhibition

As EGF-supplemented extended culture enabled the detection of putative PMSCs (PDX1^{-}/CHGA^{-} population) without implantation, methods to remove putative and *in vivo* PMSCs were explored. Given that s7-iPICs had a lower frequency of *in vivo* PMSCs than s6-iPICs (Fig. 13C), the cells generated by adding the s7-iPIC-specific inducers R428, N-acetylcysteine, and Trolox to s6-iPICs were evaluated (Fig. 26A). Although the addition of N-acetylcysteine and Trolox did not change the population after extended culture, the addition of R428, an anexelekto (AXL) inhibitor, significantly reduced the putative PMSCs (Fig. 26B and C). Since AXL inhibitors are kinase inhibitors and PMSCs are a proliferative cell population with uterine characteristics, a multi-kinase inhibitor, Lenvatinib, which is effective against endometrial carcinoma, was examined (60). Lenvatinib induced a similar reduction in putative PMSCs to R428 (Fig. 26B and C). These results suggest that kinase inhibitors can reduce *in vivo* PMSCs, and that factors that reduce *in vivo* PMSCs can be detected using extended culture. However, *in vivo* PMSCs appeared even in s7-iPICs that were treated with the three kinase inhibitors R428, PD-166866, and TR06141363. Thus, a mechanism of action other than kinase inhibition are necessary to ensure the removal of residual *in vivo* PMSCs from s7-iPIC grafts.

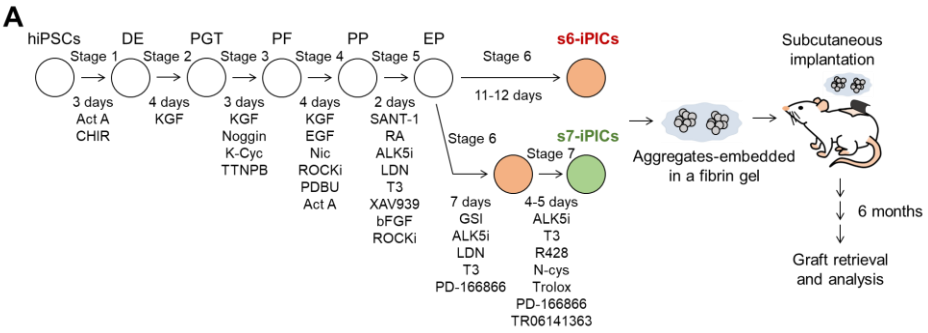
Considering that PMSCs have proliferative abilities and uterine characteristics, the focus was placed on the platinum complex, cisplatin, and taxanes, docetaxel, which are used as first-line chemotherapies for endometrial carcinoma (61, 62). In cells induced with cisplatin or docetaxel treatment, a significant reduction in the number of putative PMSCs was observed (Fig. 26D and E). Docetaxel exhibited a more pronounced putative PMSC removal than cisplatin (Fig. 26E). In addition, docetaxel clearly reduced the $PDX1^{+}/CHGA^{-}$ population (putative cysts), which is the other population that increases after extended culture and the pancreatic lineage non-endocrine population that can form cysts after implantation. Conversely, cisplatin and docetaxel treatments did not affect the number of $CHGA^{+}$ endocrine cells (Fig. 26E), suggesting that the percentage increase in the $PDX1^{-}/CHGA^{-}$ and $PDX1^{+}/CHGA^{-}$ populations after extended culture (Fig. 26D) is due to proliferation of non-endocrine cell population rather than a shift from the $CHGA^{+}$ endocrine cell population. Other chemotherapeutic agents were also evaluated, including the alkylating agent, cyclophosphamide; the anti-estrogenic agents, tamoxifen and anastrozole; and the human epidermal growth factor receptor (HER) 1/2 inhibitor, lapatinib. However, they had little effect on putative PMSC and cyst populations (Fig. 27A and B). These results indicate that platinum complexes and taxanes are effective compounds for removing *in vivo* PMSCs via a mechanism other than kinase inhibition. In addition, the compounds that appeared to be effective in removing PMSCs had little effect or tended to increase the $insulin^{+}/NKX6.1^{+}$ and $insulin^{+}/NKX6.1^{-}$ populations, which were likely to mature into β- and α- cells after implantation (Fig. 28A-C and 30A-B).

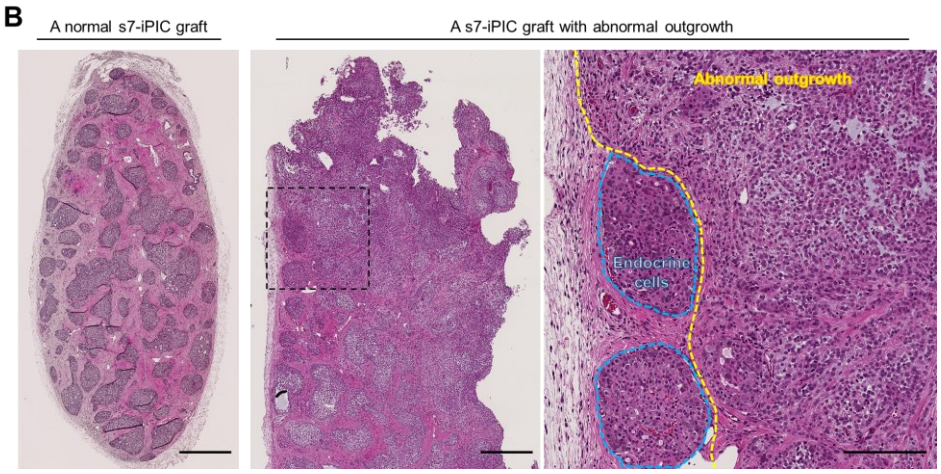
3.6 Docetaxel treatment of s7-iPICs abrogates the appearance of off-target cells while showing therapeutic efficacy

Finally, the effectiveness of docetaxel in removing putative PMSCs and cysts *in vivo*, when combined with s7-iPICs, was evaluated (Fig. 29A). Similar to the results obtained after implanting s7-iPICs without docetaxel treatment (Fig. 30C and D), streptozotocin-induced diabetic mice implanted with docetaxel-treated s7-iPICs showed human C-peptide levels of > 2,000 pmol/L (> 6 ng/mL) in the plasma within 8 weeks of implantation. Accordingly, the blood glucose levels in these mice were normalized (Fig. 29B and C). To evaluate insulin secretion in response to glucose levels, an oral glucose tolerance test was performed 23 weeks post-implantation. Plasma human C-peptide levels increased within 15 min of glucose loading (Fig. 29D and E). Immunohistochemical evaluation of grafts revealed that most HuN-positive implanted cells were insulin- and glucagon-positive pancreatic endocrine cells, with neither α-SMA-positive PMSCs nor CK19-positive cysts observed 24–30 weeks

after implantation (Fig. 29F-J). Same experiments were repeated, finally implanting docetaxel-treated s7-iPICs in 85 mice (cumulative implanted cell number > 3.0×10^8 cells), and neither PMSCs nor cysts were observed (Fig. 29K and 30E-H). To clarify the effect of docetaxel alone *in vivo*, docetaxel-treated s6-iPICs were evaluated. Docetaxel treatment resulted in a reduction in PMSC and cyst frequency (s6-iPICs vs. docetaxel-treated s6-iPICs), which was more evident than the effect of s7 factors (s6-iPICs vs. s7-iPICs) (Fig. 29K). In summary, docetaxel treatment is not only effective in *in vitro* extended culture but also in *in vivo* implantation. Furthermore, when combined with s7-iPICs, docetaxel can significantly reduce the risk of off-target cell appearance after implantation while maintaining therapeutic efficacy.

4. Figures





C

cell type	Number of induction batches	Implanted cell number ($x10^6$ cells/mouse)	Total number of implanted mice	Number of mice with outgrowth appearance	**Abnormal outgrowth appearance frequency (%)**
s6-iPICs	7	2.0-4.0	45	38	**84**
s7-iPICs	8	2.5-4.0	95	12	**13**

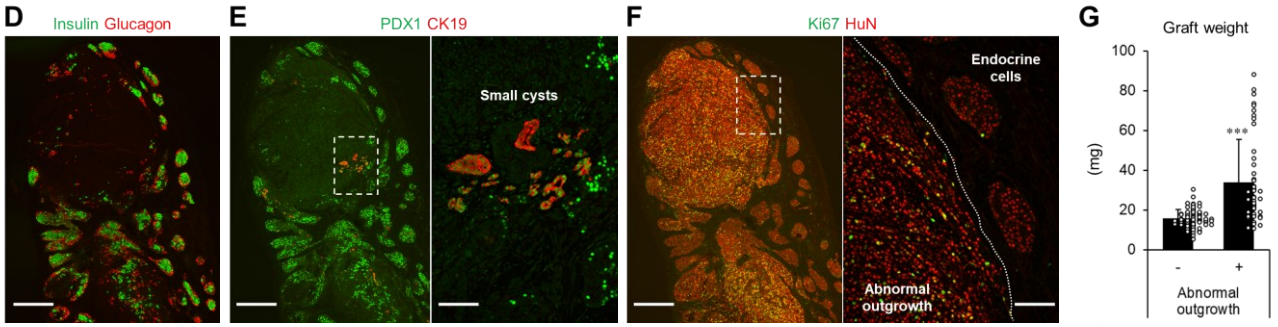


Figure 13 | Unexpected abnormal outgrowth in s7-iPIC grafts is outside the pancreatic lineage and continues to proliferate after implantation.

(A) Schematic representation of s6-iPIC and s7-iPIC differentiation and subcutaneous implantation using fibrin gel. DE, definitive endoderm; PGT, primitive gut tube; PF, posterior foregut; PP, pancreatic progenitors; EP, endocrine progenitors.

(B–G) Cell implantation experiments in streptozotocin-injected and normal NOD-scid mice. Mice were implanted with s6-iPICs or s7-iPICs (2.0–4.0 × 10^6 cells/mouse) embedded in a fibrin gel into the subcutaneous space.

(B) HE-stained sections 24 weeks after implantation. The left image shows a normal s7-iPIC graft, and the two right images show a s7-iPIC graft with abnormal outgrowth. The high magnification image is an enlarged image of the area enclosed by the dotted line in the low magnification image. Black scale bars indicate 500 μm at low magnification and 200 μm at high magnification. The images are representative of dozens of samples showing similar results.

(C) Details of abnormal outgrowth appearance frequency after s6-iPIC and s7-iPIC implantation.

(D–F) Immunohistochemical images of an s7-iPIC graft with abnormal outgrowth at 24 weeks post-implantation. White scale bars indicate 500 μm at low magnification and 100 μm at high magnification. HuN; human nucleus. Images were taken from serial sections of the same sample and are representative of dozens of samples showing similar results.

(G) Graft weight of samples without abnormal outgrowth (n = 82, combined total number of s6-iPIC and s7-iPIC grafts); and samples with abnormal outgrowth (n = 47, combined total number of s6-iPIC and s7-iPIC grafts). See Fig. 14A for graft weight distribution according to the type and number of implanted cells. Data are shown as the mean ± SD. ***$P < 0.001$, Aspin–Welch test.

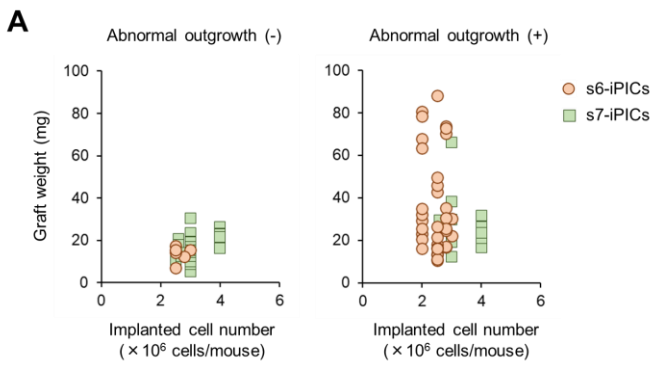


Figure 14 | Additional information for graft weight. Related to Fig. 13G.

(A) Individual graft weight distribution by type and number of implanted cells for grafts without abnormal outgrowth (s6-iPICs: n = 6, s7-iPICs: n = 76) and with abnormal outgrowth (s6-iPICs: n = 35, s7-iPICs: n = 12), respectively.

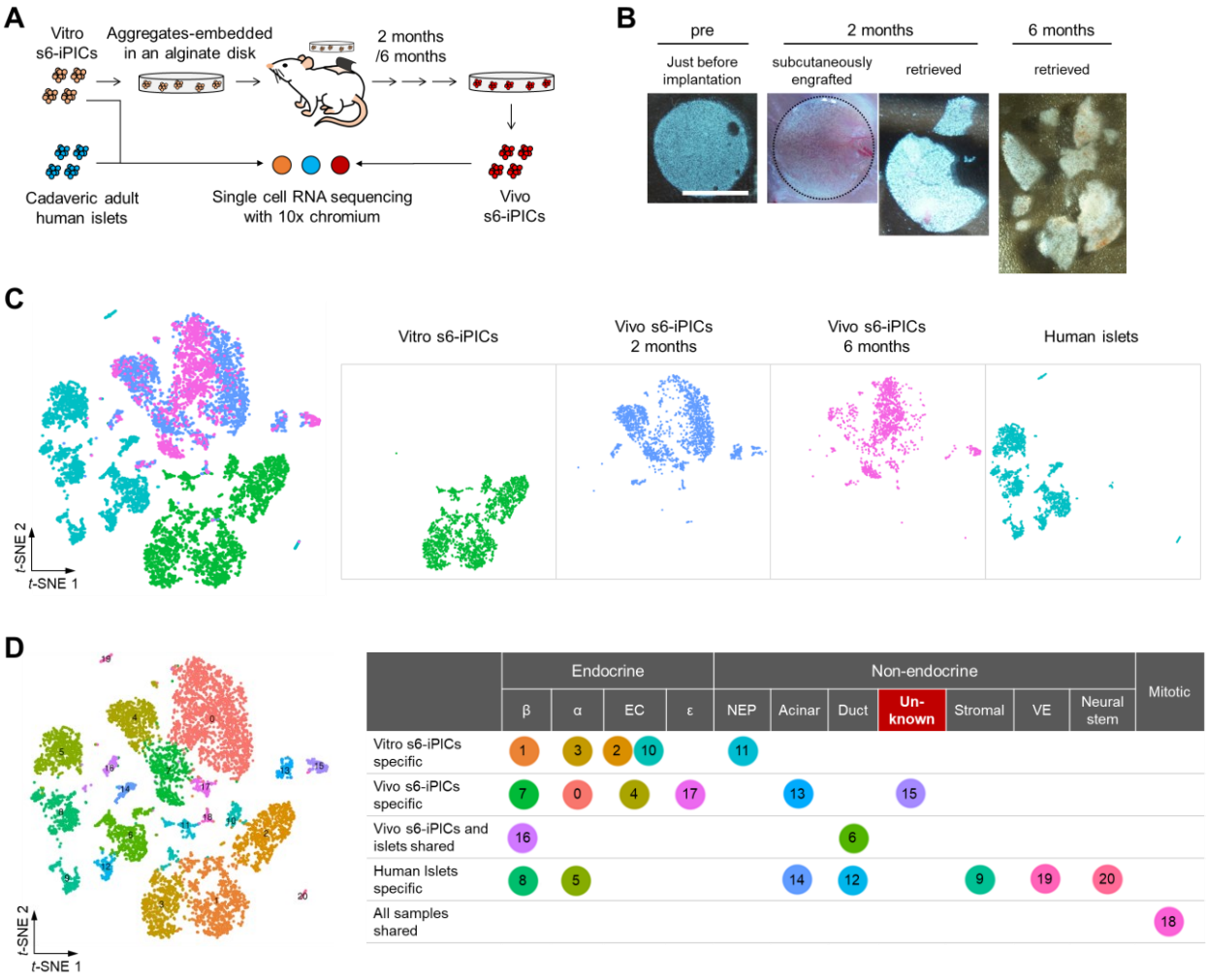


Figure 15 | The unknown cell population that cannot be classified as known cell types in single-cell analysis could be responsible for abnormal outgrowth.

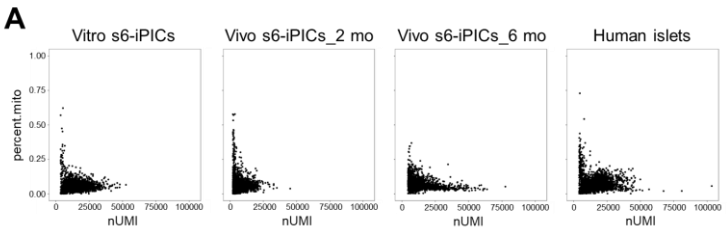
(A) Schematic representation of Vitro and Vivo s6-iPIC samples subjected to scRNA-seq.

(B) Macroscopic photographs of Vivo s6-iPIC samples. From left to right, alginate gel-embedded s6-iPICs before implantation, the graft was subcutaneously engrafted or retrieved 2 months after implantation, and the graft was retrieved 6 months after implantation. The area surrounded by a dotted line indicates the area where implanted cells have engrafted subcutaneously. The white bar indicates 5 mm.

(C) Cell distribution on *t*-SNE projections for four combined samples and each sample. The four samples were broken down as follows: one sample of Vitro s6-iPICs, two samples of Vivo s6-iPICs (2 and 6 months after implantation), and one sample of reference human islets.

(D) Shared nearest neighbor clustering identified five clusters specific to Vitro s6-iPICs (1, 2, 3, 10,

and 11), six clusters specific to Vivo s6-iPICs (0, 4, 7, 13, 15, and 17), and seven clusters specific to human islets (5, 8, 9, 12, 14, 19, and 20). Clusters 6 and 16 were shared by Vivo s6-iPICs and human islets. Cluster 18, a population of proliferative cells, was shared by all the samples. See Fig. 16C for further details. Each cluster was assigned to a known cell type using the characteristic gene expression shown in Fig. 16D. β, β-fate cells or actual β-cells; α, α-fate cells or actual α-cells; EC, enterochromaffin-fate cells; ε, ε-fate cells; NEP, non-endocrine progenitor cells; Acinar, acinar-like cells or actual acinar cells; Duct, duct-like cells or actual duct cells; Unknown, unknown cells that could not be assigned to known cells; Stromal, stromal cells; VE, vascular endothelial cells; Neural stem, neural stem cells; Mitotic, Mitotic cells.



B

Sample ID	processing status	# of cells	median UMI counts per cell	median genes detected per cell
Vitro s6-iPICs	before QC	3125	13265	4014
	after QC	2936	13354.5	4036.5
Vivo s6-iPICs_2 months	before QC	2802	6130	2069
	exclude mouse cells	2755	6215	2084
	after QC	2422	6410	2122
Vivo s6-iPICs_6 months	before QC	1924	10455.5	2736.5
	exclude mouse cells	1912	10484	2741
	after QC	1584	11267.5	2832
Human islets	before QC	2959	14935	3400
	after QC	2419	14720	3410

C

Cluster No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Vitro s6-iPICs	0	1015	813	721	0	0	1	0	0	0	202	172	0	0	0	0	1	0	11	0	0
Vivo s6-iPICs 2 months	1112	7	0	0	542	0	59	337	3	0	1	0	0	129	0	44	85	77	26	0	0
Vivo s6-iPICs 6 months	948	1	0	0	127	1	21	303	3	0	0	0	0	12	0	88	21	45	13	0	1
Human islets	0	0	0	0	0	645	564	0	505	239	0	0	161	0	135	0	18	0	33	69	50

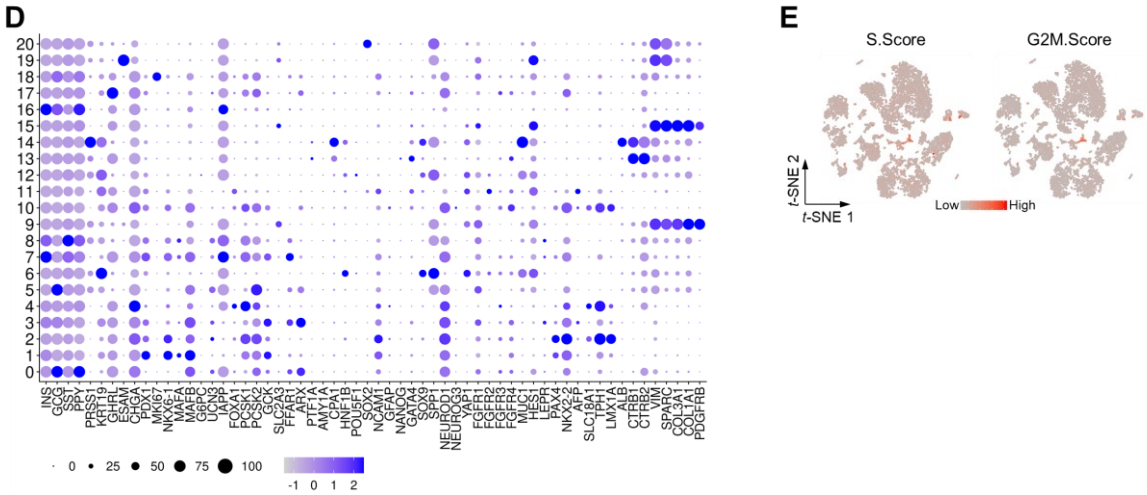


Figure 16 | Additional information for single-cell RNA sequencing analysis. Related to Fig. 15.

(A) Mitochondrial gene ratio and number of UMI counts for each individual cell in each sample.

(B) Number of cells, median UMI counts per cell, and median genes detected per cell before and after quality control by the mitochondrial gene ratio. For Vivo s6-iPIC samples, cells containing mouse genes were excluded prior to quality control.

(C) Number of cells in the clusters classified in Fig. 15D. As Vivo s6-iPIC samples at 2 and 6 months post-implantation did not show any differences in cluster classification, the two *in vivo* samples were

treated collectively.

(D) Bubble plot of endocrine and non-endocrine signature genes in the clusters classified in Fig. 15D. Color intensity indicates average relative expression levels. The bubble size indicates the percentage of expressing cells.

(E) Cell cycle phase assignments based on the S-, G2-, and M-phase gene signatures indicate a highly proliferative population (in red) within the *t*-SNE projection.

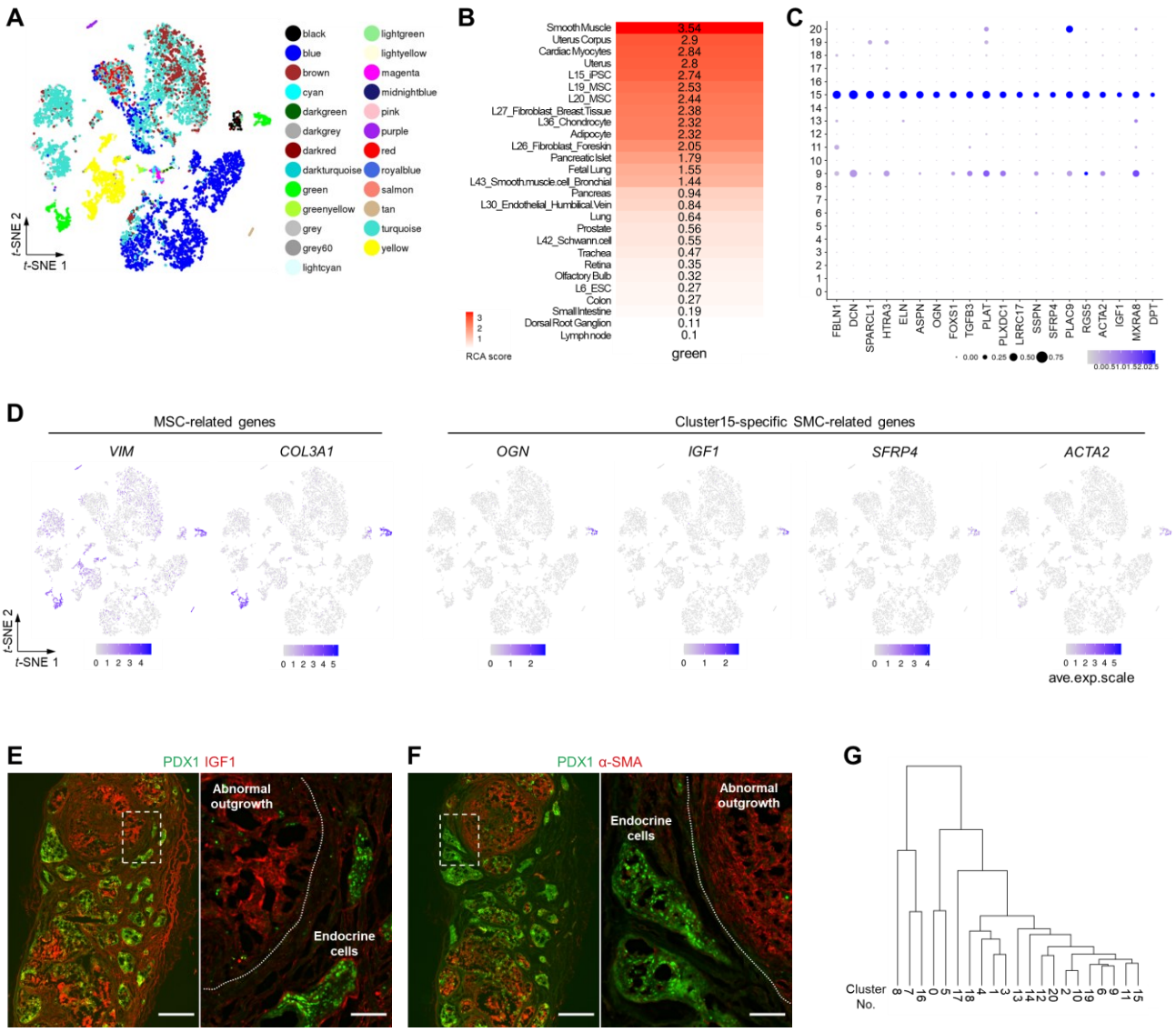


Figure 17 | Abnormal outgrowth consists of proliferative MSC- and SMC-like cells (PMSCs).

(A) Newly classified cell populations identified by reference component analysis (RCA) on the *t*-SNE projection.

(B) Heatmap of tissues and cell lines with similarity to cells annotated as "green" (Clusters 15 and 9 in Fig. 15D) in the RCA. See Fig. 18A for a full RCA heatmap.

(C) Bubble plot of the top 20 differentially expressed genes in Cluster 15 ($P < 0.05$, fold-change > 1.2, and pct.2 < 0.05). Color intensity indicates average relative expression levels. The bubble size indicates the percentage of expressing cells.

(D) Single-cell gene expression of MSC- and Cluster 15-specific SMC-related markers on the *t*-SNE projection.

(E and F) Immunohistochemical images of an s7-iPIC graft with abnormal outgrowth at 24 weeks post-implantation. White scale bars indicate 500 μm at low magnification and 100 μm at high magnification. Images were taken from serial sections of the same sample as in Fig. 13D-F and are representative of dozens of samples showing similar results. See Fig. 21A for immunohistochemical data on other PMSC-related markers.

(G) Hierarchical clustering analysis. Cluster 15 was most closely related to the non-endocrine progenitor cells in Vitro s6-iPICs (Cluster 11) and most distantly related to the β-cell population (Clusters 8, 7, and 16).

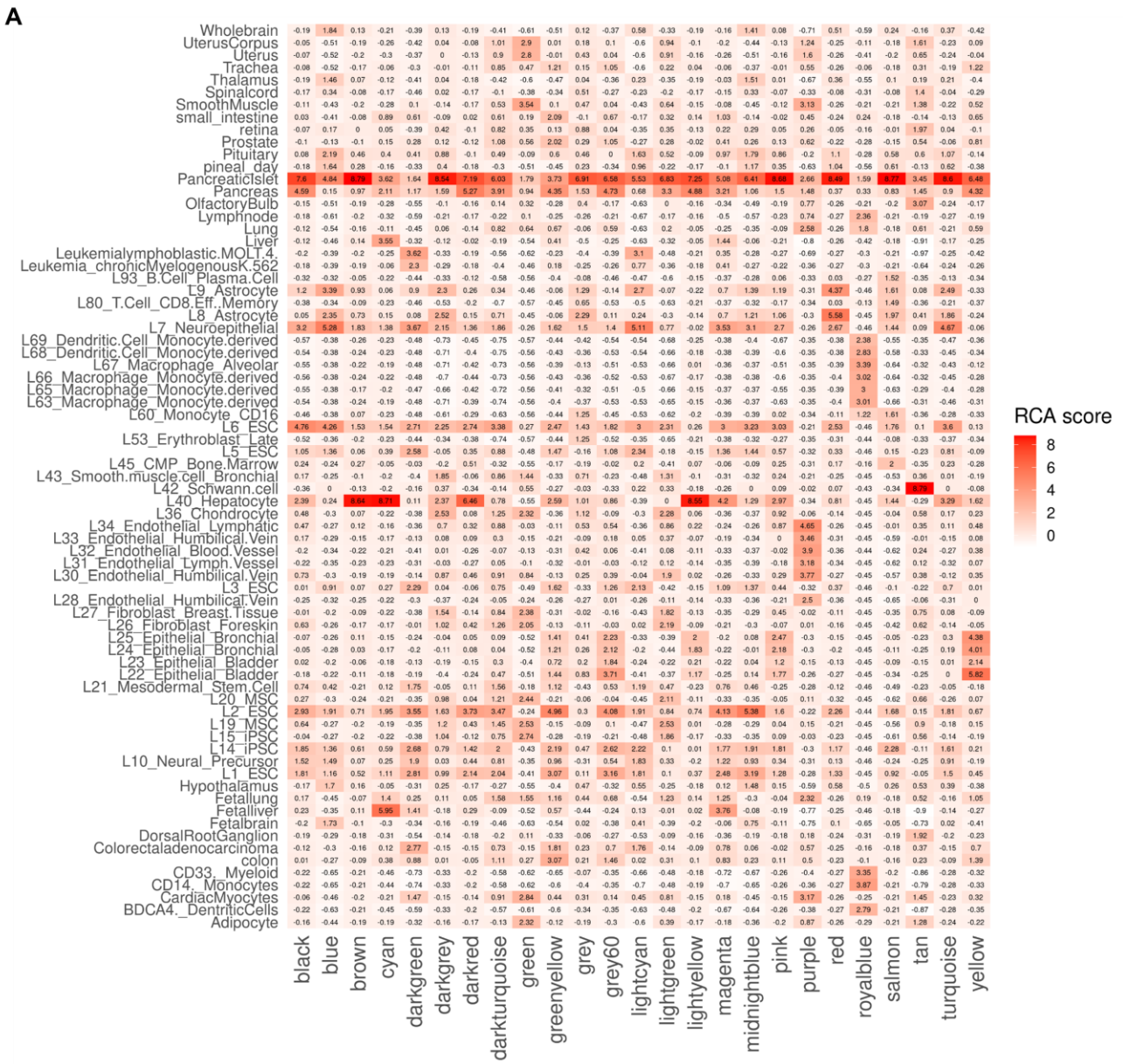
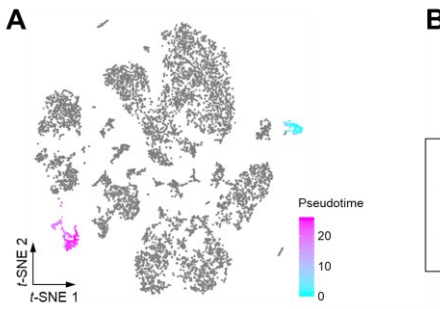
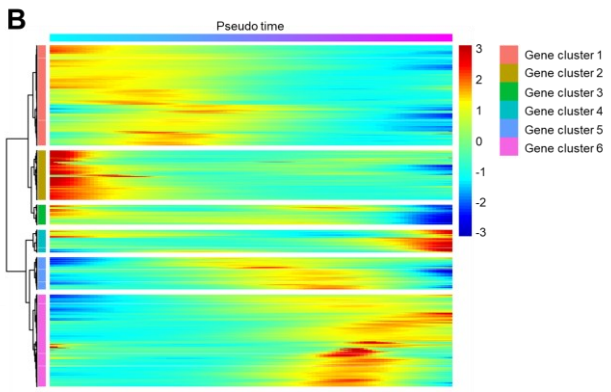


Figure 18 | Detailed information of RCA. Related to Fig. 17.

(A) Heatmap of similar tissues or cell lines for each color classified by the RCA in Fig. 17A.





C

cluster	1	2	3	4	5	6
# of genes	1301	641	259	292	417	1198
top 10 significant genes (q < 0.05)	MGP, APOE, DLK1, IGF2, PPY, C7, SPARCL1, CHGA, FBLN1, MFAP4	ASPN, OGN, JUNB, IGF1, DPT, HOPX, NGFR, CAPN6, JUN, KLF2	RGS5, PLAT, APOC1, ACTB, CALM2, ANXA5, CYR61, ITM2B, UBC, BTG1	TIMP3, POSTN, CTHRC1, COL6A3, SCD, GREM1, COL1A2, FN1, CTSK, ARL4C	MMP1, IL24, TFPI2, GPX3, ADIRF, CHI3L1, CDH6, TPM1, MYL6, GDF15	MT2A, CXCL8, FABP4, IL11, CCL26, SPP1, SFRP2, CTSD, STC1, SOD2
# of GO terms	709	121	422	169	761	1276
top 10 significant GO terms (q < 0.05)	developmental process	regulation of transcription from RNA polymerase II promoter	positive regulation of ubiquitin-protein ligase activity involved in regulation of mitotic cell cycle transition	SRP-dependent cotranslational protein targeting to membrane	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	immune system process
	anatomical structure development	transcription from RNA polymerase II promoter	positive regulation of proteolysis involved in cellular protein catabolic process	protein targeting to ER	antigen processing and presentation of exogenous peptide antigen via MHC class I	cellular response to chemical stimulus
	single-multicellular organism process	smoothened signaling pathway	positive regulation of cellular catabolic process	cotranslational protein targeting to membrane	antigen processing and presentation of peptide antigen via MHC class I	defense response
	single-organism developmental process	positive regulation of RNA metabolic process	positive regulation of ubiquitin protein ligase activity	establishment of protein localization to endoplasmic reticulum	response to organic substance	immune response
	multicellular organism development	positive regulation of gene expression	positive regulation of catabolic process	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	antigen processing and presentation	regulation of immune system process
	negative regulation of cellular process	positive regulation of RNA biosynthetic process	regulation of ubiquitin protein ligase activity	protein localization to endoplasmic reticulum	antigen processing and presentation of peptide antigen	response to stress
	system development	positive regulation of developmental process	positive regulation of cellular protein catabolic process	translational initiation	cellular response to organic substance	response to organic substance
	negative regulation of biological process	positive regulation of macromolecule biosynthetic process	negative regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle	viral gene expression	cellular response to chemical stimulus	localization
	anatomical structure morphogenesis	positive regulation of transcription, DNA-templated	stimulatory C-type lectin receptor signaling pathway	viral transcription	antigen processing and presentation of exogenous peptide antigen	regulation of response to stimulus
	regulation of developmental process	positive regulation of nucleic acid-templated transcription	regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle	nuclear-transcribed mRNA catabolic process	regulation of cellular amino acid metabolic process	cell migration

Figure 19 | Comparison of Clusters 15 and 9 using pseudotime analysis.

(A) Estimated pseudotime from Cluster 15 to Cluster 9 on the *t*-SNE projection.

(B) Expression heatmap of differentially expressed genes ordered by their common kinetics through pseudotime.

(C) Top 10 significant genes (q-value < 0.05) and enriched gene ontology (GO) terms (q-value < 0.05) of each gene cluster.

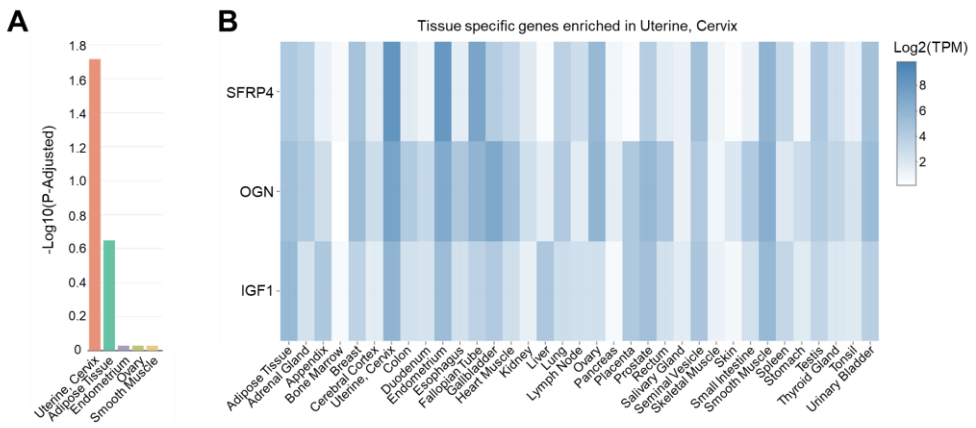


Figure 20 | Tissue-specific gene enrichment analysis of the Cluster 15 specific genes.

(A and B) Annotation of the top 20 differentially expressed genes in Cluster 15 (Fig. 17C) using TissueEnrich (https://tissueenrich.gdcb.iastate.edu/).

(A) Human tissues extracted as related tissues from the Human Protein Atlas dataset.

(B) Expression intensities of *SFRP4*, *OGN*, and *IGF1* (causative genes of high relevance to the uterus in Fig. 20A) in each human tissue.

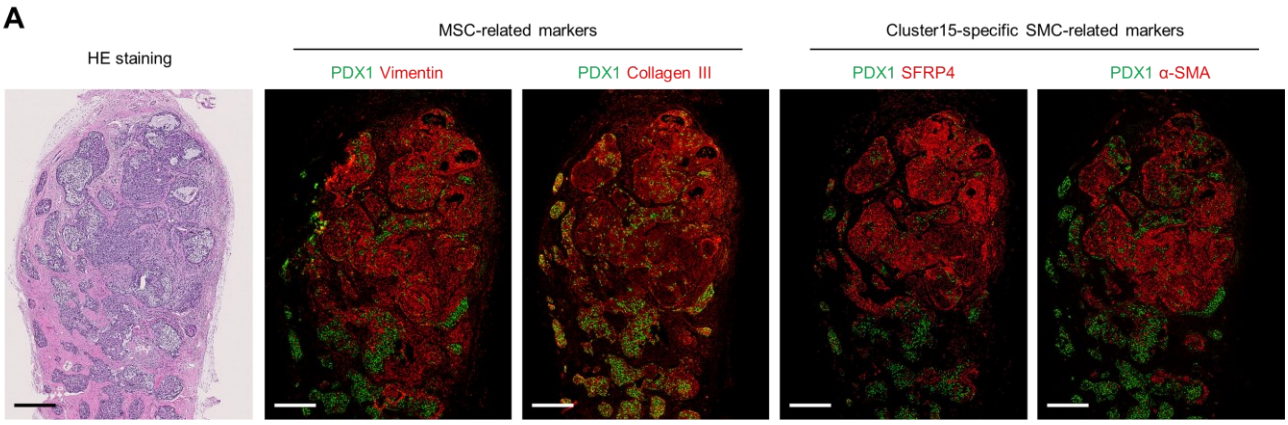


Figure 21 | Additional information for tissue immunostaining. **Related to Fig. 17E and F.**

(A) HE and immunohistochemical images of an s7-iPIC graft with abnormal outgrowth at 24 weeks post-implantation. Black and white scale bars indicate 500 μm. Images were taken from serial sections of the same sample as in Fig. 17E and F and are representative of dozens of samples showing similar results.

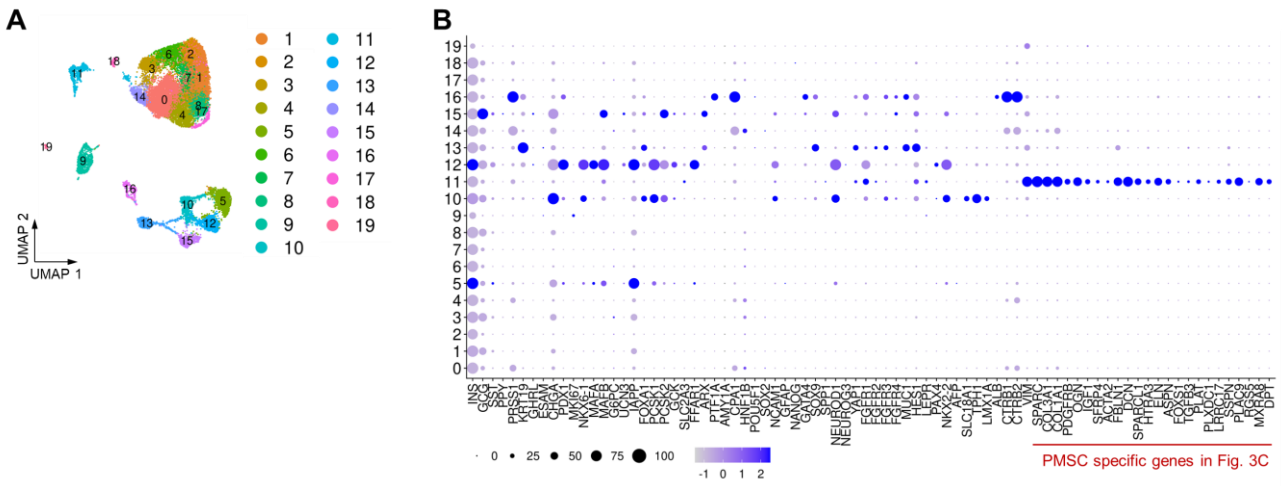


Figure 22 | Detection of PMSCs equivalent cells using previously reported scRNA-seq data of ESC/iPSC-derived islet-like cell grafts.

(A) Shared nearest neighbors clustering on the UMAP projection using scRNA-seq data from *in vivo* samples included in GSE151117. Data from three ESC (HUES8)-derived islet-like cell grafts (GSM4567001, GSM4567002, and GSM4567003) and two iPSC (WS4^{corr})-derived islet-like cell grafts (GSM4567004 and GSM4567005) were reanalyzed.

(B) Bubble plot of endocrine and non-endocrine signature genes in the clusters classified in Fig. 22A. Cluster 11 expressed PMSC-specific markers such as *VIM*, *COL3A1*, *OGN*, *IGF1*, *SFRP4*, and *ACTA2*.

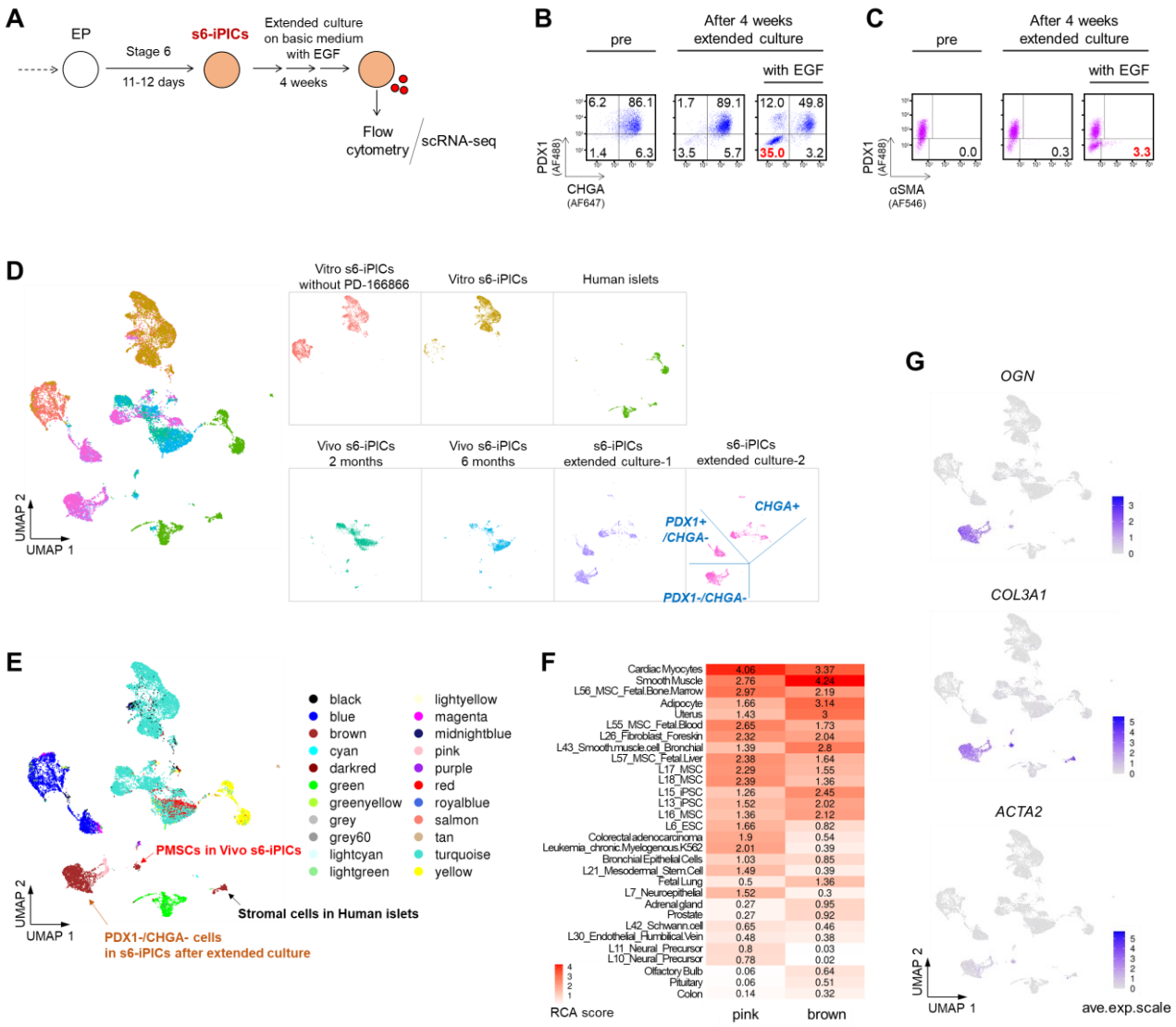


Figure 23 | An EGF-supplemented extended culture system exposes a cell population (putative PMSCs) that closely resembles *in vivo* PMSCs.

(A) Schematic representation of extended culture, subsequent flow cytometry analysis, and scRNA-seq using s6-iPICs.

(B and C) Representative flow cytometry plots illustrating the protein expression of s6-iPICs before and after extended culture. The numbers in each plot diagram show the percentage of each population. The reproducibility was confirmed in three independent experiments.

(D) Cell distribution on UMAP projections for seven combined samples and for each sample. The four samples shown in Fig. 15C were reanalyzed with the addition of the following three samples: one sample of Vitro s6-iPICs cultured without PD-166866 and two samples (technical duplicates) of s6-iPICs after 4 weeks of extended culture including EGF treatment. Cells in the samples after extended

culture were classified into three types based on gene expression intensity: *PDX1*$^{-}$/*CHGA*$^{-}$, *PDX1*$^{+}$/*CHGA*$^{-}$, and *CHGA*$^{+}$ populations. See Fig. 25A-C for further details.

(E) Newly classified cell populations identified by RCA on the UMAP projection. The *PDX1*$^{-}$/*CHGA*$^{-}$ population in the extended culture samples, PMSCs in Vivo s6-iPICs, and stromal cells in human islets were classified as "pink" and "brown". Fig. 25A-B show the position of PMSCs in Vivo s6-iPICs and stromal cells in human islets on UMAP projections.

(F) Heatmap of tissues and cell lines similar to cells annotated as "pink" and "brown" in the RCA. See Fig. 25D for a full RCA heatmap.

(G) Single-cell gene expression of PMSC markers on the UMAP projection.

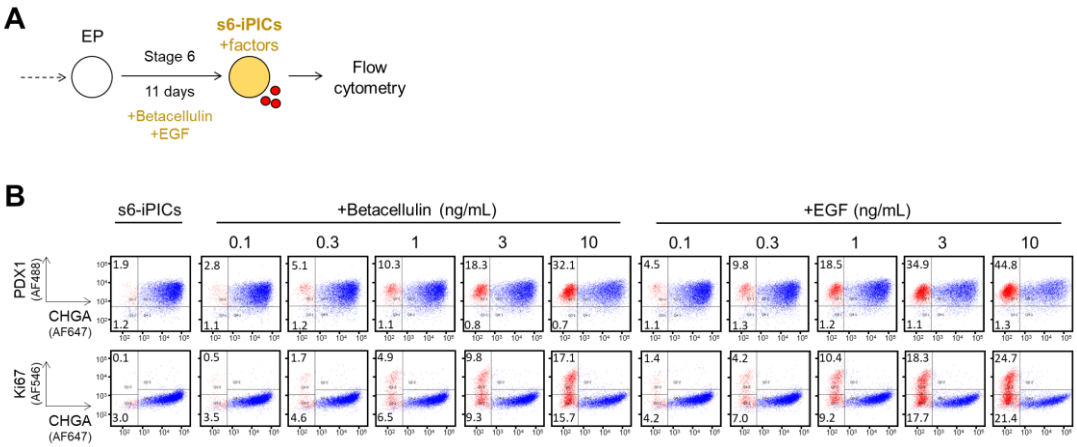
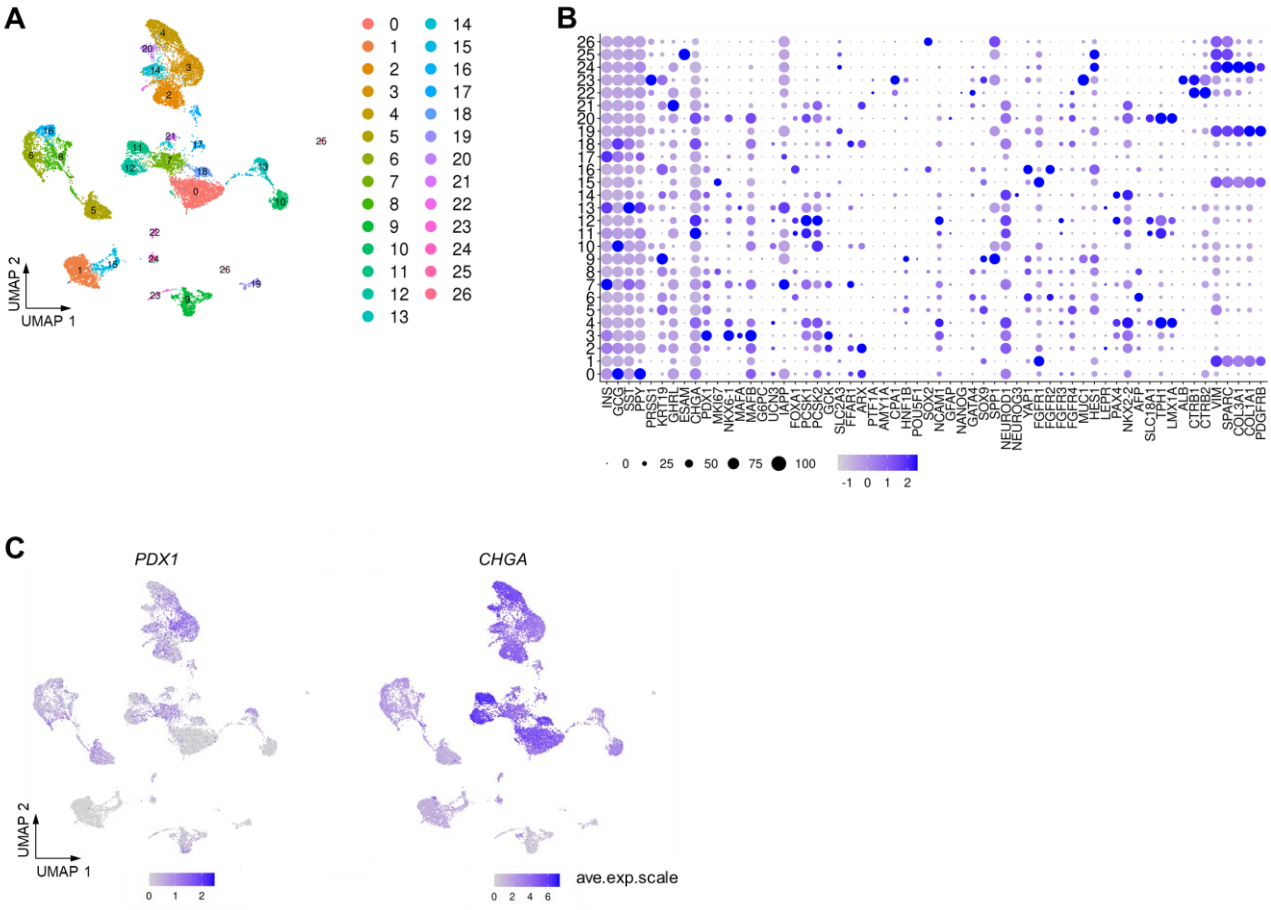


Figure 24 | Proliferation-promoting activity of betacellurin and EGF on the non-endocrine population in s6-iPICs.

(A) Schematic representation of s6-iPIC induction including betacellulin or EGF treatment, and subsequent flow cytometry analysis.

(B) Flow cytometry plots illustrating the protein expression of s6-iPICs differentiated with betacellulin or EGF treatment. The numbers in each plot diagram show the percentage of each population. The CHGA-negative non-endocrine population, shown in red, increased in a dose-dependent manner by betacellulin or EGF treatment.


A
UMAP 2
UMAP 1
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
B
INS GCG SST PPY PRSS1 KRT19 GHRL ESAM CPA1 PDX1 MKI67 NKX6.1 MAFA MAFB G6PC UCN3 IAPP FOXA1 FOXA2 PCSK2 GCK SLC2A3 FFAR1 ARX PTF1A AMY1A CPA1 HNF1B POU5F1 SOX2 NCAM1 GFAP NANOG GATA4 SOX9 SPP1 NEUROD1 NEUROG3 YAP1 FGFR1 FGFR2 FGFR3 FGFR4 MUC1 HES1 EPR1 KRT14 NKX2.2 AFP SLC18A1 TPH1 LMX1A ALB CTRB1 CTRB2 VIM SPARC COL3A1 COL1A1 PDGFRB
0 25 50 75 100
-1 0 1 2
C
PDX1
CHGA
UMAP 2
UMAP 1
0 1 2
ave.exp.scale
0 2 4 6

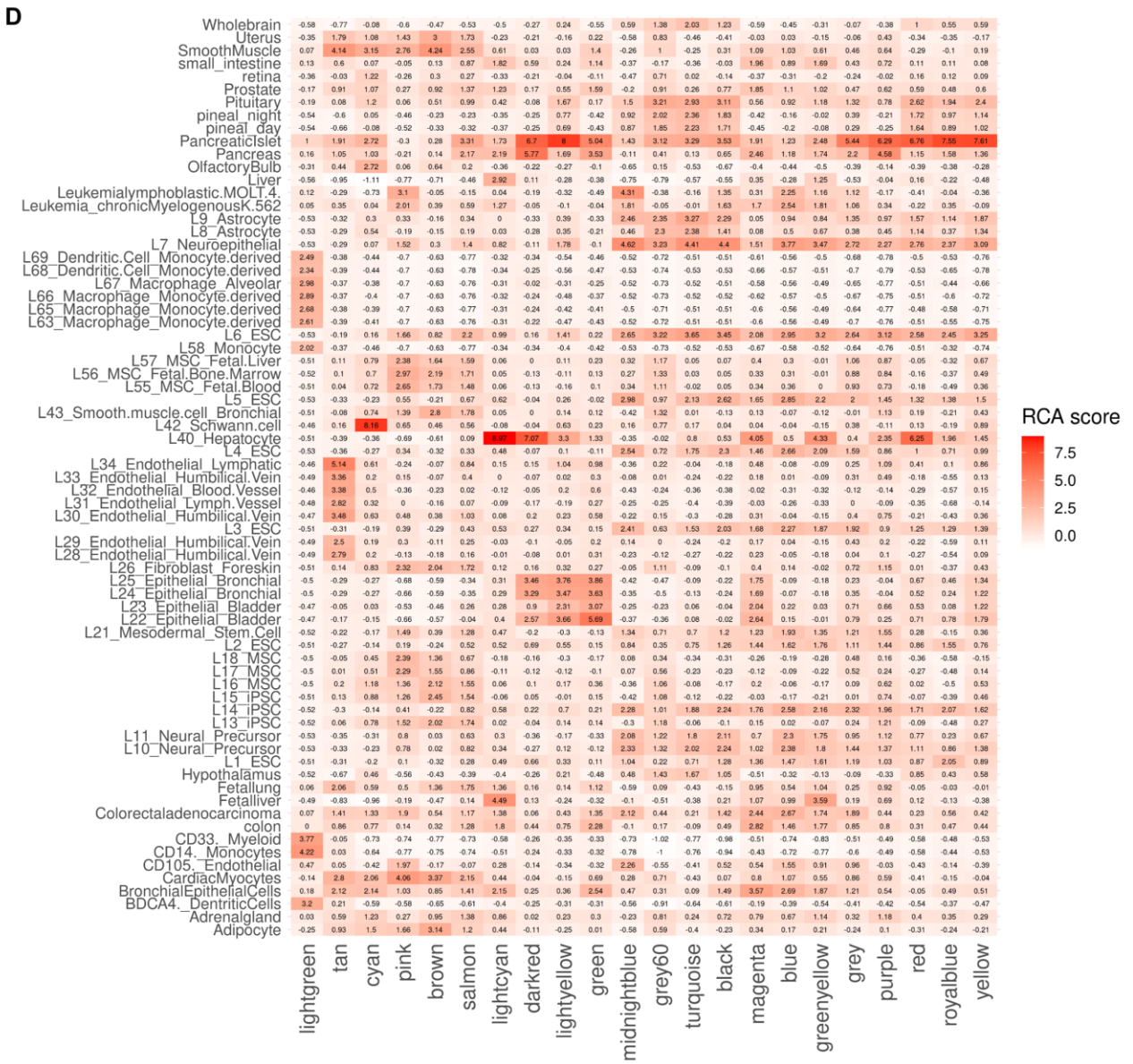


Figure 25 | Additional information for scRNA-seq analysis containing extended culture samples. Related to Fig. 23.

(A) Shared nearest neighbors clustering of the reanalyzed scRNA-seq data in Fig. 23D on the UMAP projection. Reanalysis with the additional samples re-assigned Clusters 15 and 9 in Fig. 15D to Clusters 24 and 19 in this figure, respectively.

(B) Bubble plot of endocrine and non-endocrine signature genes in the clusters classified in Fig. 25A. Color intensity indicates average relative expression levels. The bubble size indicates the percentage of expressing cells.

(C) Single-cell gene expression of *PDX1* and *CHGA* on the UMAP projection in Fig. 23D.

(D) Heatmap of similar tissues or cell lines for each color classified by the RCA in Fig. 23E.

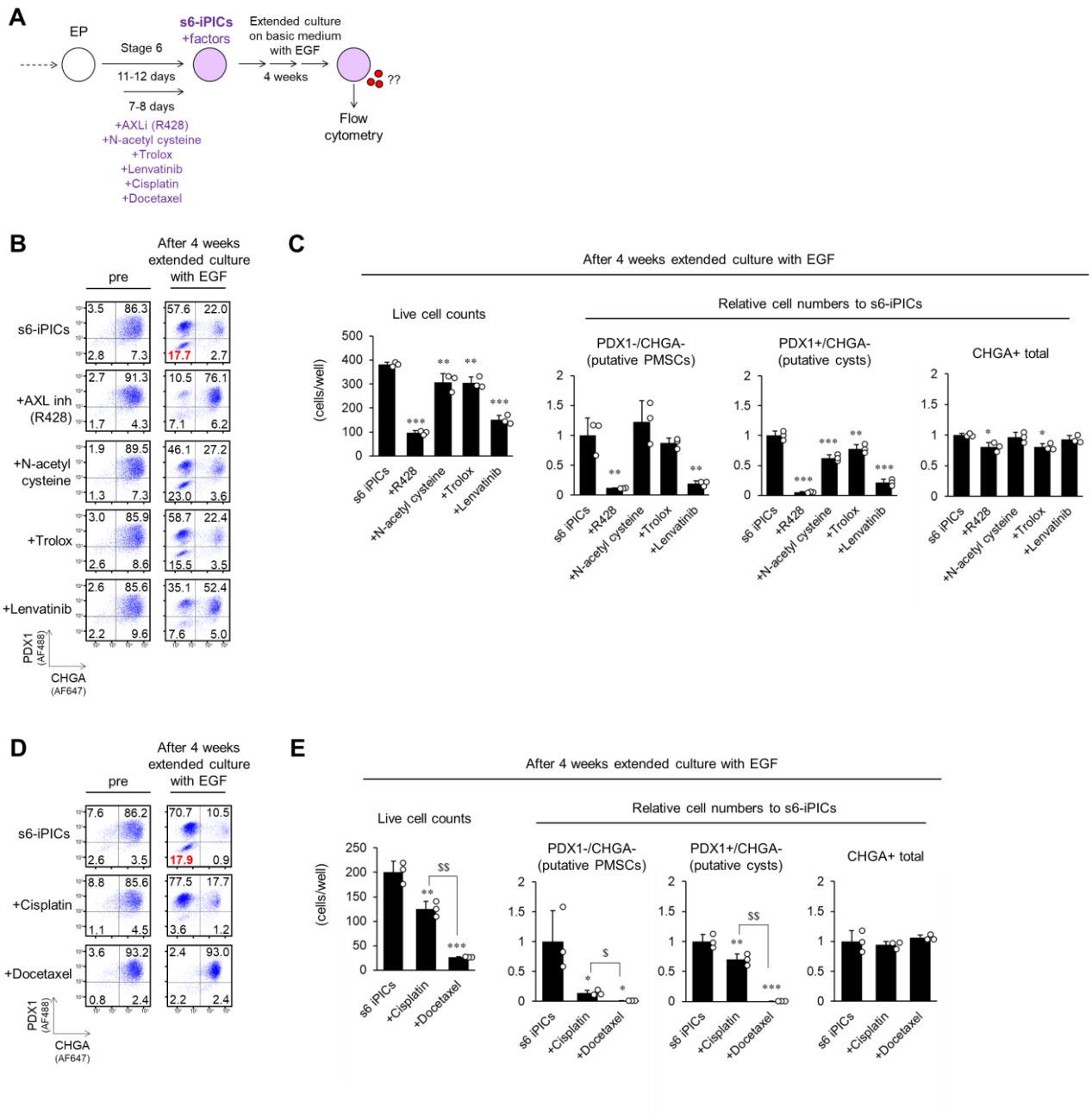


Figure 26 | Cisplatin and docetaxel effectively remove putative PMSCs through mechanisms other than kinase inhibition.

(A) Schematic representation of s6-iPIC derivatives induced with or without additional compound treatment and subsequent extended culture and flow cytometry analysis.

(B–E) Representative flow cytometry plots illustrating protein expression before and after extended culture of s6-iPIC derivatives (**B and D**). Live cell counts and relative cell numbers of $PDX1^{-}/CHGA^{-}$, $PDX1^{+}/CHGA^{-}$, and $CHGA^{+}$ populations in each s6-iPIC derivative post-extended culture. The

number of cells in control s6-iPICs was set to "1" (**C and E**). The number of cells in each population was calculated from flow cytometry results (percentage of each population) and live cell counts. Data are shown as the mean ± SD (n = 3, technical replicates). The reproducibility of the effect of docetaxel was confirmed in three independent experiments. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ versus s6-iPICs, Dunnett's test. $^{\$}P < 0.05$, $^{\$\$}P < 0.01$, Aspin–Welch test.

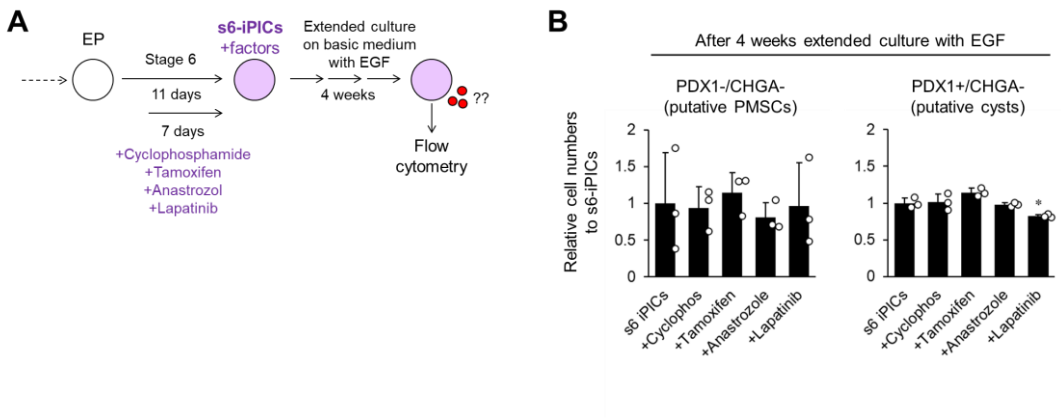


Figure 27 | Evaluation of the effects of cyclophosphamide, tamoxifen, anastrozole, and lapatinib treatment on putative PMSC and cyst populations.

(A) Schematic representation of s6-iPIC derivatives induced with or without additional compound treatment and subsequent extended culture and flow cytometry analysis.

(B) Relative cell numbers of $PDX1^{-}/CHGA^{-}$, $PDX1^{+}/CHGA^{-}$, and $CHGA^{+}$ populations in each s6-iPIC derivative post-extended culture. The number of cells in control s6-iPICs was set to "1". The number of cells in each population was calculated from flow cytometry results (percentage of each population) and live cell counts. Data are shown as the mean ± SD (n = 3, technical replicates). $*P < 0.05$, versus s6-iPICs, Dunnett's test.

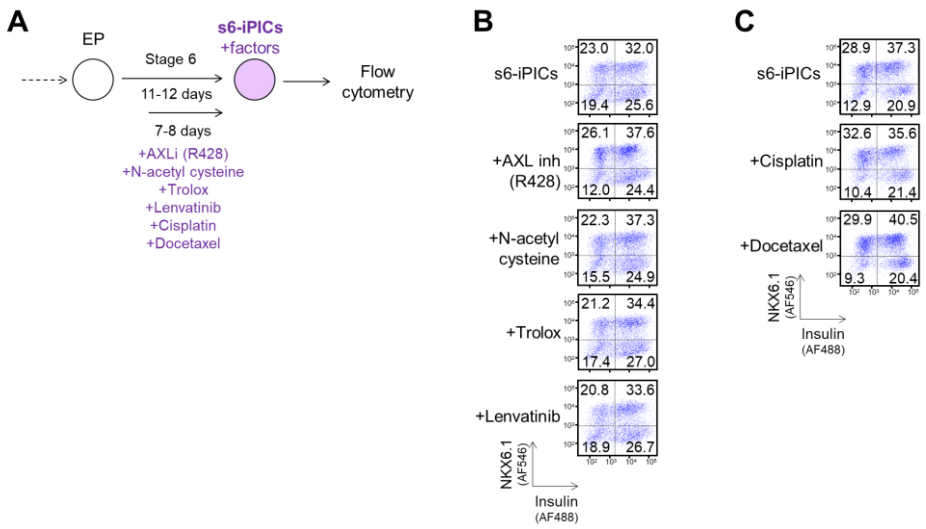
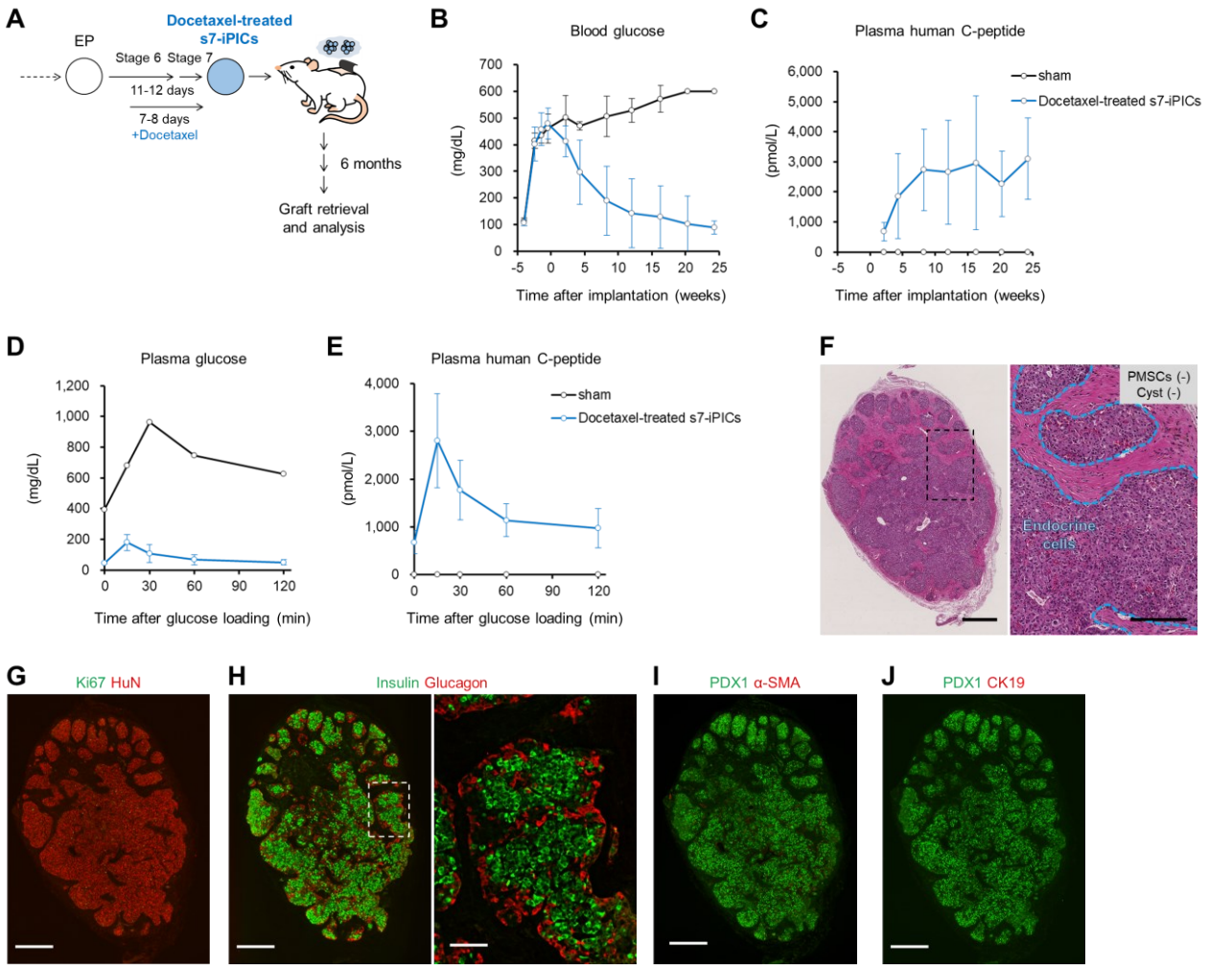


Figure 28 | Evaluation of the effect of kinase inhibitors, cisplatin, and docetaxel treatment on insulin-positive cell rates.

(A) Schematic representation of s6-iPIC derivatives induced with or without additional compound treatment and flow cytometry analysis.

(B and C) Representative flow cytometry plots illustrating insulin and NKX6.1 protein expression before an extended culture of s6-iPIC derivatives.



K

cell type	Number of induction batches	Implanted cell number (x10^6 cells/mouse)	Total number of implanted mice	Number of mice with appearance			Appearance frequency (%)		
				PMSCs	Large cysts (>500 μm)	Small cysts (<500 μm)	**PMSCs**	**Large cysts (>500 μm)**	**Small cysts (<500 μm)**
s6-iPICs	7	2.0-4.0	45	38	30	45	**84**	**67**	**100**
s7-iPICs	8	2.5-4.0	95	12	13	86	**13**	**14**	**91**
Docetaxel-treated s6-iPICs	4	4.5	21	0	0	2	**0** (p=0.120)	**0** (p=0.121)	**10** (***)
Docetaxel-treated s7-iPICs	6	3.0-4.5	85	0	0	0	**0** (***)	**0** (***)	**0** (***)

Figure 29 | Docetaxel treatment of s7-iPICs abrogates the appearance of off-target cells while showing therapeutic efficacy.

(A) Schematic representation of docetaxel-treated s7-iPIC differentiation and subcutaneous implantation using fibrin gel.

(B–J) Cell implantation experiments in streptozotocin-injected NOD-scid mice. Mice were implanted with docetaxel-treated s7-iPICs (3.0–4.5 × 10^6 cells/mouse) embedded in a fibrin gel into the subcutaneous space.

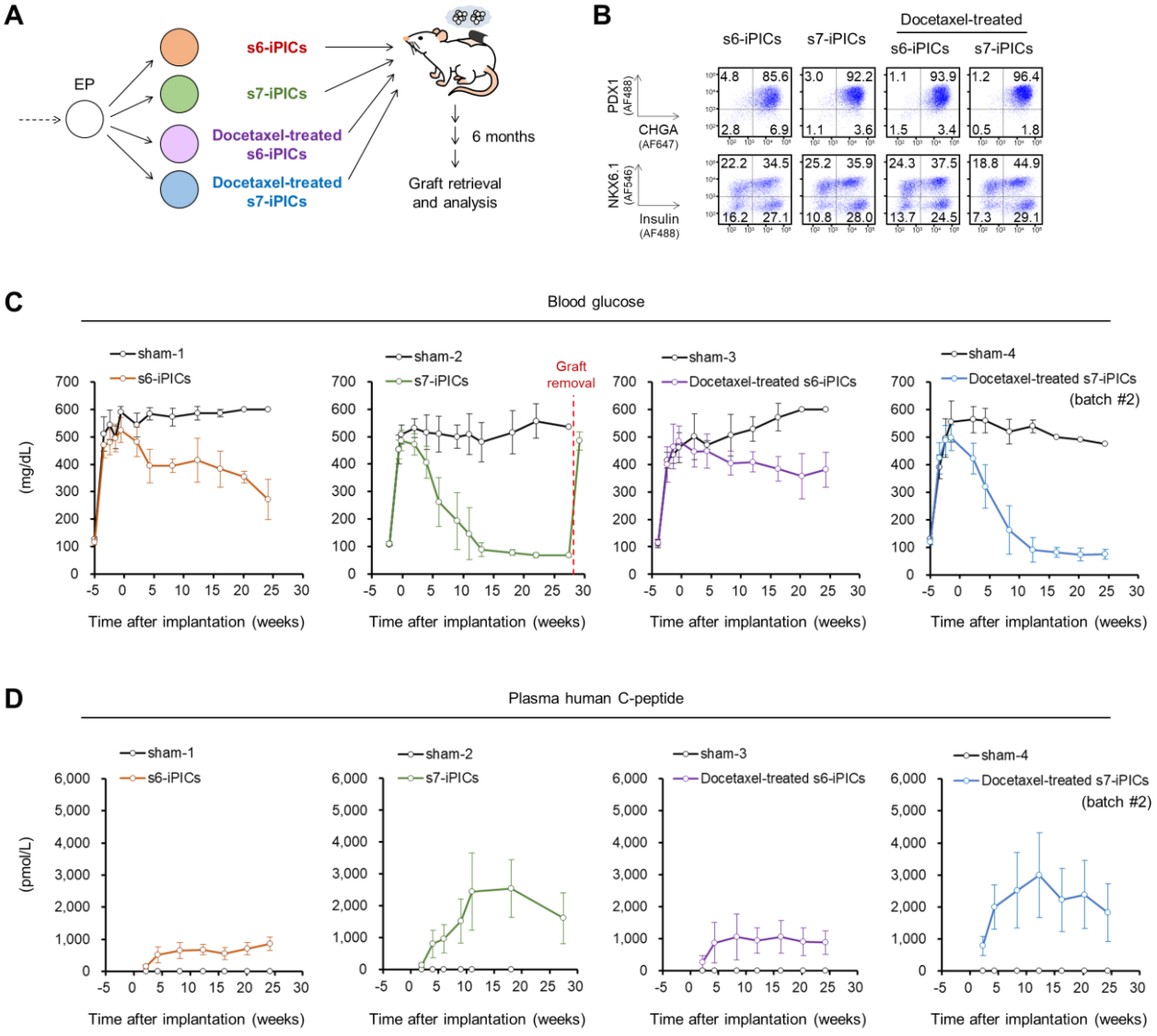
(B and C) Blood glucose and plasma human C-peptide levels after docetaxel-treated s7-iPIC (4.0×10^6 cells/mouse) implantation. Data are shown as mean ± SD (sham: n = 5→1, docetaxel-treated s7-iPICs: n = 18→13). The decrease in the n number was caused by unexpected death. See Fig. 30A-D for efficacy data of the other iPICs.

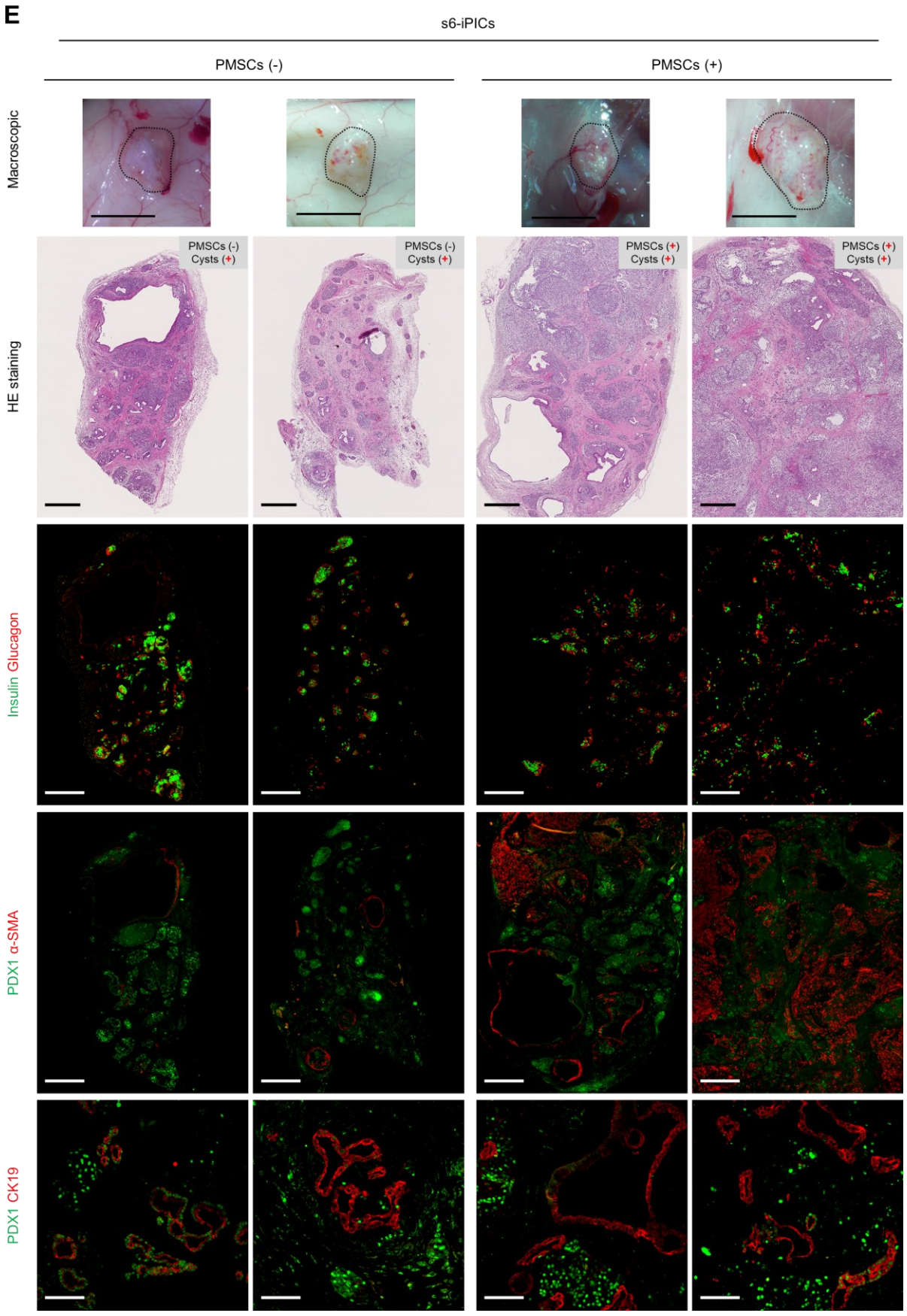
(D and E) Plasma glucose, and human C-peptide levels during the oral glucose tolerance test 23 weeks after docetaxel-treated s7-iPIC (4.0×10^6 cells/mouse) implantation. Data are shown as mean ± SD (sham: n = 1, docetaxel-treated s7-iPICs: n = 13).

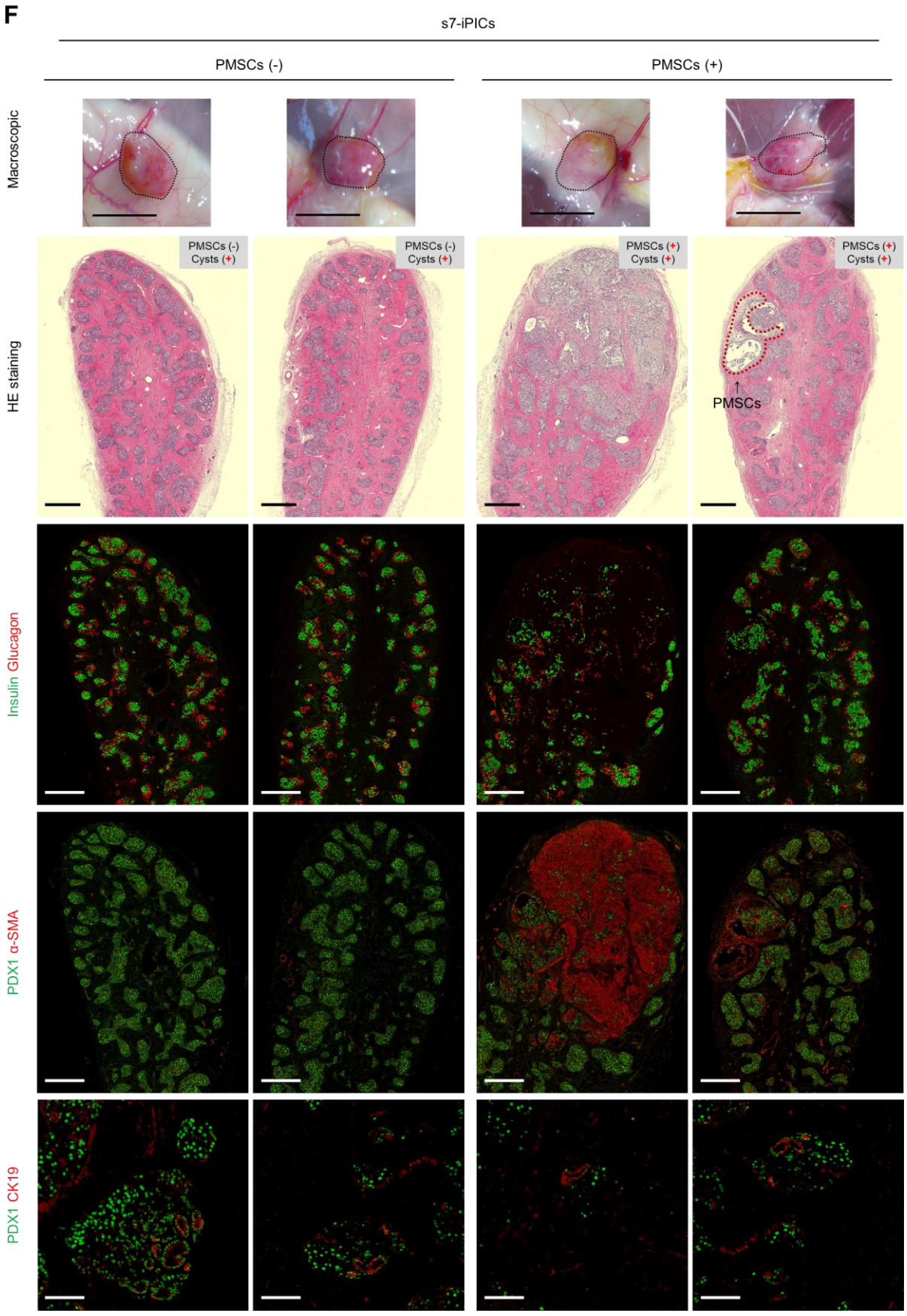
(F) HE-stained sections 24 weeks after implantation. Black scale bars indicate 500 μm at low magnification and 200 μm at high magnification. The images are representative of dozens of samples showing similar results.

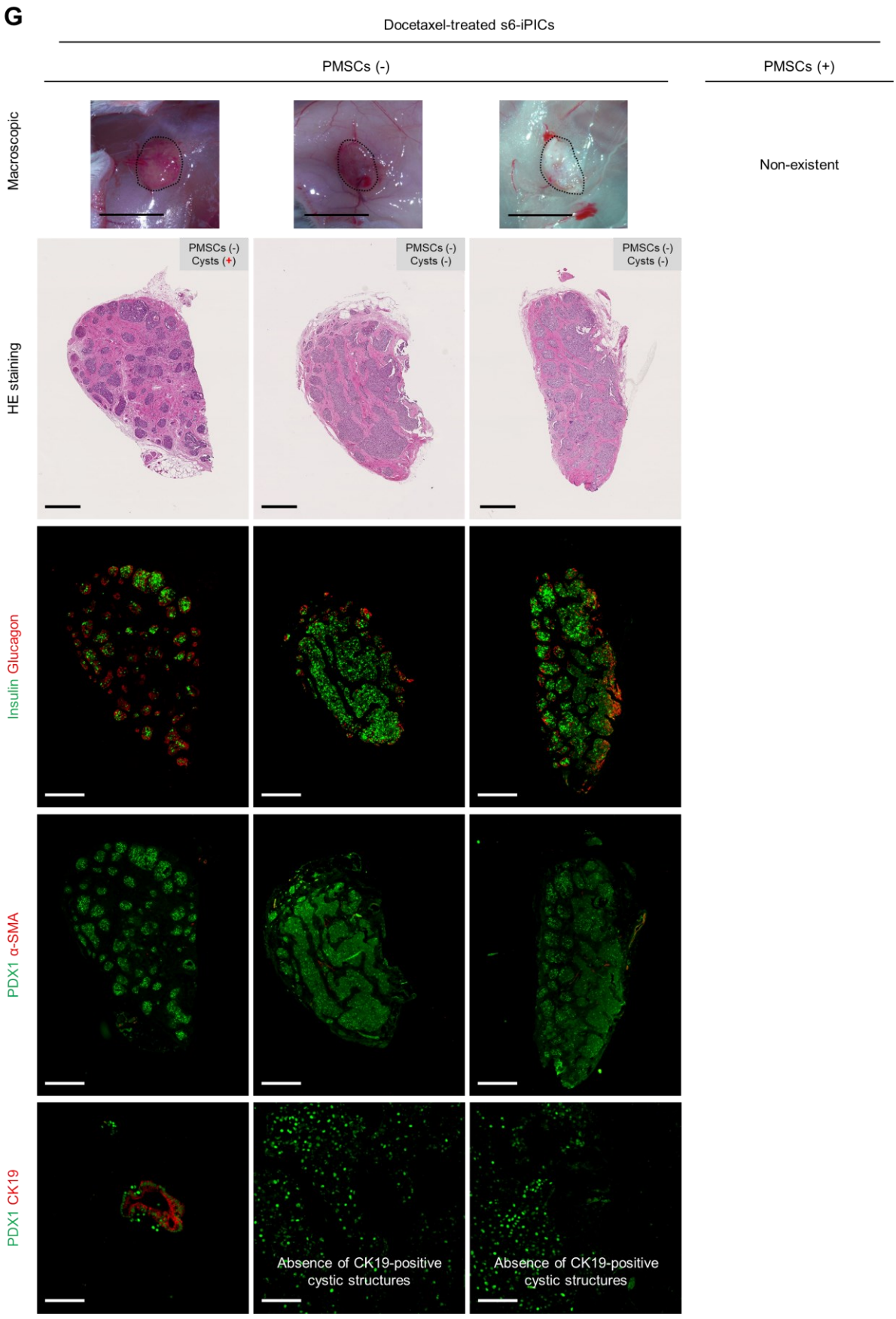
(G–J) Immunohistochemical images of a docetaxel-treated s7-iPIC graft at 24 weeks post-implantation. White scale bars indicate 500 μm at low magnification and 100 μm at high magnification. HuN; human nucleus. Images were taken from serial sections of the same sample and are representative of dozens of samples showing similar results. See Fig. 30E-H for HE and immunohistochemical images on other iPICs.

(K) Details of PMSC, large cyst (> 500 μm in diameter), and small cyst (< 500 μm in diameter) appearance frequency after fibrin gel-embedded subcutaneous implantation of s6-iPIC, s7-iPIC, docetaxel-treated s6-iPIC, and docetaxel-treated s7-iPIC. The parts of s6-iPICs and s7-iPIC data are previously shown in Fig. 13C and are repeated here for easy comparison between all groups. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ versus s7-iPICs, Fisher's exact test.


A
EP
s6-iPICs
s7-iPICs
Docetaxel-treated s6-iPICs
Docetaxel-treated s7-iPICs
6 months
Graft retrieval and analysis
B
s6-iPICs
s7-iPICs
Docetaxel-treated
s6-iPICs
s7-iPICs
PDX1 (AF488)
CHGA (AF647)
NKX6.1 (AF546)
Insulin (AF488)
4.8 85.6 2.8 6.9
3.0 92.2 1.1 3.6
1.1 93.9 1.5 3.4
1.2 96.4 0.5 1.8
22.2 34.5 16.2 27.1
25.2 35.9 10.8 28.0
24.3 37.5 13.7 24.5
18.8 44.9 7.3 29.1
C
Blood glucose
sham-1
s6-iPICs
sham-2
s7-iPICs
Graft removal
sham-3
Docetaxel-treated s6-iPICs
sham-4
Docetaxel-treated s7-iPICs (batch #2)
(mg/dL)
Time after implantation (weeks)
D
Plasma human C-peptide
sham-1
s6-iPICs
sham-2
s7-iPICs
sham-3
Docetaxel-treated s6-iPICs
sham-4
Docetaxel-treated s7-iPICs (batch #2)
(pmol/L)
Time after implantation (weeks)


E
s6-iPICs
PMSCs (-)
PMSCs (+)
Macroscopic
HE staining
PMSCs (-) Cysts (+)
PMSCs (-) Cysts (+)
PMSCs (+) Cysts (+)
PMSCs (+) Cysts (+)
Insulin Glucagon
PDX1 α-SMA
PDX1 CK19


F
s7-iPICs
PMSCs (-)
PMSCs (+)
Macroscopic
HE staining
PMSCs (-) Cysts (+)
PMSCs (-) Cysts (+)
PMSCs (+) Cysts (+)
PMSCs (+) Cysts (+)
PMSCs
Insulin Glucagon
PDX1 α-SMA
PDX1 CK19


G
Docetaxel-treated s6-iPICs
PMSCs (-)
PMSCs (+)
Macroscopic
Non-existent
HE staining
PMSCs (-)
Cysts (+)
PMSCs (-)
Cysts (-)
PMSCs (-)
Cysts (-)
Insulin Glucagon
PDX1 α-SMA
PDX1 CK19
Absence of CK19-positive cystic structures
Absence of CK19-positive cystic structures

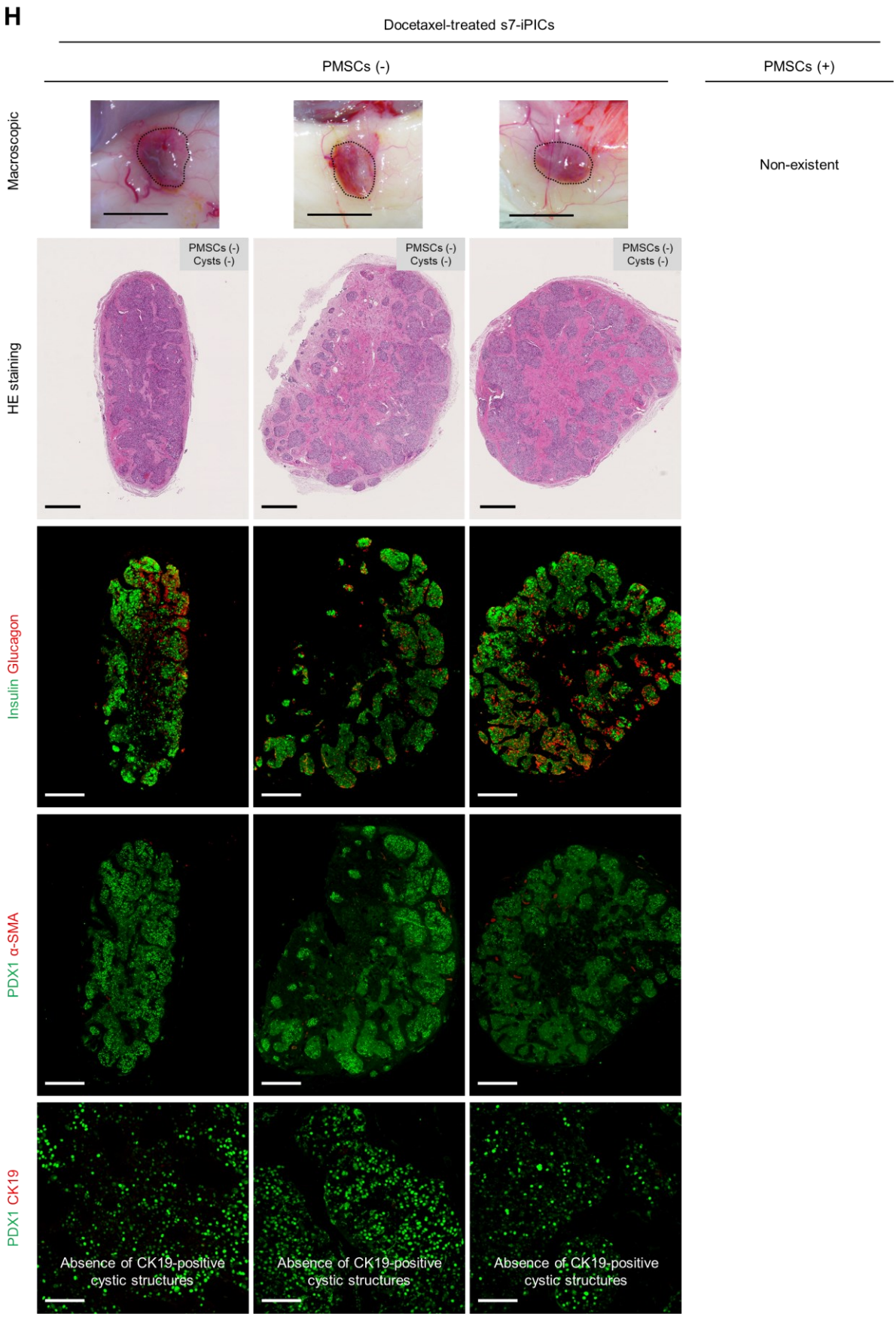

H
Docetaxel-treated s7-iPICs
PMSCs (-)
PMSCs (+)
Macroscopic
Non-existent
HE staining
PMSCs (-)
Cysts (-)
PMSCs (-)
Cysts (-)
PMSCs (-)
Cysts (-)
Insulin Glucagon
PDX1 α-SMA
PDX1 CK19
Absence of CK19-positive cystic structures
Absence of CK19-positive cystic structures
Absence of CK19-positive cystic structures

Figure 30 | Additional information for implantation studies. **Related to Fig. 29.**

(A) Schematic representation of four types of iPICs differentiation and subcutaneous implantation using fibrin gel.

(B) Representative flow cytometry plots illustrating protein expression before implantation of four types of iPICs. The numbers in each plot diagram show the percentage of each population.

(C and D) Blood glucose and plasma human C-peptide levels after implantation of s6-iPICs (3.0×10^6 cells/mouse), s7-iPICs (3.0×10^6 cells/mouse), docetaxel-treated s6-iPICs (4.0×10^6 cells/mouse), and docetaxel-treated s7-iPICs (batch #2) (4.0×10^6 cells/mouse). Data are shown as mean ± SD (sham-1: n = 5→1, s6-iPICs: n = 5→3, sham-2: n = 5→2, s7-iPICs: n = 16→6, sham-3: n = 5→1, docetaxel-treated s6-iPICs: n = 11→8, sham-4: n = 5→1, docetaxel-treated s7-iPICs (batch #2): n = 22→17). The decrease in the n number was caused by unexpected death.

(E–H) Macroscopic, HE, and immunohistochemical images of s6-iPIC grafts **(E)**, s7-iPIC grafts **(F)**, docetaxel-treated s6-iPIC grafts **(G)**, and docetaxel-treated s7-iPICs grafts **(H)**. In the macroscopic images, the areas surrounded by dotted lines indicate areas where the implanted cells engrafted subcutaneously. Black and white scale bars indicate 5 mm in macroscopic images, 500 μm in HE, insulin/glucagon, PDX1/α-SMA images, and 100 μm in PDX1/CK19 images.

5. Discussion

In the present study, PMSCs, which have MSC and SMC traits, were identified as unexpected proliferative off-target cells that emerged after the implantation of iPSC-derived islet-like cells (Fig. 13, 15, and 17). As PMSCs are expected to appear at a high frequency when s7-iPICs are implanted in the order of 10^8, it was assumed that their clinical application would be challenging without addressing these PMSCs. However, the detection and removal of *in vivo* PMSCs was difficult because of the following reasons: i) PMSC appearance is infrequent in implantation studies using cells on the order of 10^6, especially with s7-iPICs (Fig. 13C); and ii) Detection of PMSCs takes several months because PMSCs do not express their specific markers before implantation and only become apparent after several months of *in vivo* maturation (Fig. 17D). Therefore, a four-week-extended culture system was developed to expose putative PMSCs ($PDX1^-/CHGA^-$ population) without implantation (Fig. 23). Moreover, the reduction of putative and *in vivo* PMSCs was demonstrated through compound treatment, particularly with the taxane docetaxel (Fig. 23 and 26). In fact, in mice implanted with docetaxel-treated s7-iPICs, the frequency of *in vivo* PMSC appearance was suppressed, whereas the anti-diabetic activity remained clear (Fig. 29).

Cells equivalent to *in vivo* PMSCs were present in grafts of iPSC/ESC-derived islet-like cells induced in a previous study (22) (Fig. 22). Another study showed proliferative cells with MSC-related gene expression observed during scRNA-seq using samples obtained from extended cultures of ESC-derived islet-like cells (17). Therefore, PMSCs are not regarded as a cell type specific to iPICs. The lack of attention paid to PMSCs could be due to the dominance of cyst formation in situations involving abundant residual non-endocrine progenitor cells (11, 16, 19); thus, PMSCs in the grafts could have been overlooked. In fact, the presence of PMSCs was not detected in early implantation experiments using cells with a large number of non-endocrine progenitors, as rapid graft enlargement caused by pancreatic duct-like cysts was observed (Fig. 10c). In this study, *in vivo* PMSCs were detected because of the following reasons: i) the use of s6-iPICs and s7-iPICs, in which non-endocrine progenitors were markedly reduced by original modification, suppressed cyst formation; and ii) repeated implantation experiments facilitated the evaluation of a large number of cells. As several clinical trials have been initiated using cells containing off-target cells that lead to cyst formation (16, 63), clinical application would be possible without the complete removal of off-target cells. For example, the risk to off-target cells can be significantly reduced by limiting their distribution using robust devices (63). However, because the risk of device breakage due to physical impact or other factors cannot be completely eliminated, the number of off-target cells should be reduced as much as possible. In this study, off-

target cells were successfully removed, paving the way for the safe and long-term implantation of iPSC/ESC-derived islet-like cells.

Possible approaches for removing off-target cells contaminating iPSC/ESC-derived cells include cell sorting (11, 17), metabolic selection (51-53), and compound treatment (18, 19). Among these approaches, compound treatment was selected due to its ability to easily process more than 10^8 *in vitro*-generated cells. The kinase inhibitors, cisplatin, and docetaxel were identified as compounds that are effective at reducing putative PMSCs through different mechanisms of action (Fig. 26). However, these compounds are not novel from the perspective of elimination approaches for proliferating cells. In contrast, in the process of identifying these compounds, the development of an off-target cell detection system based on extended culture (Fig. 23) is worth highlighting. This detection system, in which putative PMSCs were successfully exposed *in vitro* by adding growth factors that were effective for the population while promoting the maturation of the PMSC origins via extended culture, was the core of the present study (Fig. 23A-C). Furthermore, comprehensive gene expression data verified that the putative PMSCs were almost equivalent to *in vivo* PMSCs (Fig. 23D-G), leading to the discovery of effective compounds for *in vivo* PMSC removal. These methods are expected to be effective not only for PMSC removal but also for the removal of off-target cells contaminating other types of iPSC/ESC-derived products.

Docetaxel was not only the most potent agent for eliminating PMSCs but also reduced cystic structures composed of pancreatic duct-like cells to undetectable levels *in vivo* and *in vitro* (Fig. 26E and 29K). According to a previous report, nocodazole, which inhibits microtubule polymerization, the opposite mechanism of action to docetaxel, promotes the induction of pancreatic duct-like cells (64). Therefore, docetaxel may exhibit cyst removal effects by inhibiting differentiation into the pancreatic duct, independent of its PMSC removal activity. Although both *in vivo* PMSCs and cysts are presumed to be derived from non-endocrine progenitor cells (Fig. 17G), the origin of PMSCs and cysts may not be completely identical based on the following observations: i) the ratio of putative PMSCs ($PDX1^-/CHGA^-$) and putative cysts ($PDX1^+/CHGA^-$) after extended culture differed between experiments (Fig. 23B, 26B, and 26D); and ii) the putative PMSCs and putative cysts showed different reduction ratios after cisplatin treatment (Fig. 26E). In chapter 1, the non-endocrine progenitor population was found to include multiple cell types. However, further studies are warranted to determine the differences between *in vivo* PMSC and cyst origins.

6. Summary

This chapter demonstrates that the implantation of seven-stage iPSC-derived pancreatic islet cells (s7-iPICs) leads to the emergence of unexpected off-target cells with proliferative capacity via *in vivo* maturation. These cells displayed characteristics of both mesenchymal stem cells (MSCs) and smooth muscle cells (SMCs), termed proliferative MSC- and SMC-like cells (PMSCs). The frequency of PMSC emergence was found to be high when 10^8 s7-iPICs were used. Given that clinical applications involve the use of a greater number of induced cells than 10^8, it is challenging to ensure the safety of clinical applications unless PMSCs are adequately addressed. Accordingly, a detection system and removal methods for PMSCs were developed. To detect PMSCs without implantation, a four-week-extended culture system was implemented, demonstrating that putative PMSCs could be reduced by compound treatment, particularly with the taxane docetaxel. When docetaxel-treated s7-iPICs were implanted, the PMSCs were no longer observed. This study provides useful insights into the identification and resolution of safety issues, which are particularly important in the field of cell-based medicine using iPSCs and ESCs.

Conclusion

The unrestricted capacity of human induced pluripotent stem cells (iPSCs) to proliferate and differentiate is a major advantage for the development of cell-based therapies. However, this advantage is accompanied by a risk of off-target cell proliferation and tumorigenesis. Therefore, minimizing the risk of off-target cell contamination is imperative for the safe use of iPSC-derived products. In particular, cell therapy using iPSC/ESC-derived pancreatic endocrine cells for T1DM has already been initiated in clinical trials, and the identification of potential off-target cells and their removal is urgently needed. Thus, the purpose of this thesis is to expose and profile all types of off-target cells that contaminate in the induction process of iPSC-derived pancreatic endocrine cells. Furthermore, based on the characteristics of off-target cells, appropriate removal methods were proposed and their effectiveness was verified through *in vivo* implantation.

In Chapter1, $CHGA^-$ non-endocrine cells were highlighted as the major off-target cells contaminating *in vitro*-generated pancreatic endocrine cells derived from iPSCs and ESCs. It was showed that non-endocrine off-target populations mostly consist of $PDX1^+$ pancreatic-lineage cells such as pancreatic duct-like cells and pancreatic progenitor cells. Additionally, the inhibition of FGFR2, PLK1/4, and glycolysis, which were predominantly activated in non-endocrine cells, was demonstrated to effectively and specifically reduce non-endocrine off-target cells. Among the non-endocrine cell reduction candidate, kinase inhibitors, such as PD-166866 and TR06141363, were used to induce s7-iPICs, and these compounds reduced most of the potential off-target cells while maintaining adequate *in vivo* efficacy.

Thus, in s7-iPICs, $PDX1^+$ pancreatic-lineage non-endocrine cells were largely reduced, and there appeared to be few remaining safety issues. Indeed, when s7-iPICs were implanted into T1DM model mice in the order of 10^6, no graft abnormalities, such as cyst-like enlargement due to proliferation of pancreatic lineage off-target cells, were observed. However, in repeated implantation trials, unexpected off-target cells outside the pancreatic lineage were identified to proliferate within the graft. Therefore, Chapter 2 focused on detail profiling of this non-pancreatic lineage unexpected off-target cells.

The unexpected off-target cells displayed characteristics of both mesenchymal stem cells (MSCs) and smooth muscle cells (SMCs), termed proliferative MSC- and SMC-like cells (PMSCs). The frequency of PMSC emergence was found to be high when 10^8 s7-iPICs were used. Given that clinical applications involve the use of a greater number of induced cells than 10^8, it is challenging to ensure the safety of clinical applications unless PMSCs are adequately addressed. Accordingly, a

detection system and removal methods for PMSCs were developed. To detect PMSCs without implantation, a four-week-extended culture system was implemented, demonstrating that putative PMSCs could be reduced by compound treatment, particularly with the taxane docetaxel. When docetaxel-treated s7-iPICs were implanted over 10^8 cells in total, the PMSCs were no longer observed. Hence, PMSCs, which are unexpected off-target cells outside the pancreatic lineage, were shown to be removable by docetaxel treatment.

Thus, docetaxel-treated s7-iPICs would be an appropriate candidate for clinical application to achieve remission of T1DM while ensuring safety. However, the cumulative number of docetaxel-treated s7-iPICs evaluated in this thesis amounts to approximately 2×10^8 cells, which is sufficient for only a few patients. Furthermore, given the potential for genetic mutations during the culture process, it is not feasible to ensure the complete absence of abnormal cells when 10^8 to 10^9 cells are implanted into multiple patients. To address this challenge, pre-implant efforts to manage abnormal cell proliferation must be combined with post-implant strategies. For example, selecting the subcutaneous space as an implantation site would facilitate post-implant monitoring and allow for emergency retrieval if necessary. Additionally, utilizing allogeneic cells instead of autologous cells would necessitate the use of immunosuppressants; however, in cases where abnormal cells arise, discontinuation of immunosuppressant dosing could enable their elimination through the host immune response. By integrating such post-implantation strategies with the findings presented in this thesis, clinical trials could be advanced while effectively addressing safety concerns.

In conclusion, studies in this thesis provide useful insights into the identification and resolution of safety issues which are encompassed in pluripotent stem cell-derived products. In particular, the methodology to expose and remove off-target cells via single cell analysis and extended culture can be widely applied in the field of cell therapy using pluripotent stem cell-derived products. Building on these findings, the clinical application of pluripotent stem cell-derived pancreatic endocrine cells can be expected to accelerate, including their potential future use in veterinary medicine.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Professor Yuko Okamatsu-Ogura (Laboratory of Biochemistry, Graduate School of Veterinary Medicine, Hokkaido University), for the inspection and critical review of this manuscript.

I would also like to express my deep appreciation to my deputy examiners, Professor Takashi Kimura (Laboratory of Comparative Pathology, Graduate School of Veterinary Medicine, Hokkaido University), Professor Ken-ichi Otsuguro (Laboratory of Pharmacology, Graduate School of Veterinary Medicine, Hokkaido University), and Associate Professor Hironobu Yasui (Laboratory of Radiation Biology, Graduate School of Veterinary Medicine, Hokkaido University) for their invaluable advice on improving this manuscript.

I am deeply grateful to Dr. Yoshiaki Kassai (Head of Global Advanced Platform, Takeda Pharmaceutical Company Limited) and Dr. Kazuhiro Matsumoto (Associate Director, Global Advanced Platform, Takeda Pharmaceutical Company Limited) for their constant encouragement and considerate guidance during my research and manuscript writing.

I greatly appreciate Dr. Taro Toyoda (Junior Associate Professor, Department of Life Science Frontiers, Center for iPS Cell Research and Application, Kyoto University) and Dr. Ryo Ito (Head of iPS cell-derived pancreatic islet cells Business Unit, Orizuru Therapeutics) for the guidance and supervision based on their expertise.

I also thank Kensuke Sakuma, Shinya Asano, Noriko Tsubooka-Yamazoe, Taisuke Mochida, Stephanie C. Napier, Shuhei Konagaya, Takeshi Watanabe, and Junji Yamaura for facilitating the research and advising me according to their expertise.

My great appreciation is also extended to Miho Ohra, Shika Inoue, Atsuno Kaneto, Ayumi Osawa, Aika Takahashi, Ayako Makita, Aki Kuwano, Sayaka Seki, Aya Takino, Shintarou Hokaiwado, and Keiko Nishijima for their technical support.

Finally, I am grateful to Eiichi Hiyoshi, Shigeko Hiyoshi, Chizuko Hiyoshi, Hinata Hiyoshi, Kaisei Hiyoshi, Asahi Hiyoshi, and other family members for understanding and heart-warming support.

References

1. J. A. Bluestone, K. Herold, G. Eisenbarth, Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature* **464**, 1293-1300 (2010).
2. A. M. Shapiro *et al.*, Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. *The New England journal of medicine* **343**, 230-238 (2000).
3. A. S. Klein *et al.*, Organ donation and utilization in the United States, 1999-2008. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons* **10**, 973-986 (2010).
4. M. C. Vantyghem, E. J. P. de Koning, F. Pattou, M. R. Rickels, Advances in β-cell replacement therapy for the treatment of type 1 diabetes. *Lancet (London, England)* **394**, 1274-1285 (2019).
5. E. Latres, D. A. Finan, J. L. Greenstein, A. Kowalski, T. J. Kieffer, Navigating Two Roads to Glucose Normalization in Diabetes: Automated Insulin Delivery Devices and Cell Therapy. *Cell metabolism* **29**, 545-563 (2019).
6. B. Catchpole, J. M. Ristic, L. M. Fleeman, L. J. Davison, Canine diabetes mellitus: can old dogs teach us new tricks? *Diabetologia* **48**, 1948-1956 (2005).
7. R. A. Delicano *et al.*, The shared risk of diabetes between dog and cat owners and their pets: register based cohort study. *Bmj* **371**, m4337 (2020).
8. H. Klingemann, Immunotherapy for Dogs: Still Running Behind Humans. *Front Immunol* **12**, 665784 (2021).
9. A. Rezania *et al.*, Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nature biotechnology* **32**, 1121-1133 (2014).
10. F. W. Pagliuca *et al.*, Generation of functional human pancreatic β cells in vitro. *Cell* **159**, 428-439 (2014).
11. G. G. Nair *et al.*, Recapitulating endocrine cell clustering in culture promotes maturation of human stem-cell-derived β cells. *Nature cell biology* **21**, 263-274 (2019).
12. T. Toyoda *et al.*, Cell aggregation optimizes the differentiation of human ESCs and iPSCs into pancreatic bud-like progenitor cells. *Stem cell research* **14**, 185-197 (2015).
13. T. Mochida *et al.*, Insulin-Deficient Diabetic Condition Upregulates the Insulin-Secreting Capacity of Human Induced Pluripotent Stem Cell-Derived Pancreatic Endocrine Progenitor Cells After Implantation in Mice. *Diabetes* **69**, 634-646 (2020).
14. N. A. J. Krentz *et al.*, Single-Cell Transcriptome Profiling of Mouse and hESC-Derived Pancreatic Progenitors. *Stem cell reports* **11**, 1551-1564 (2018).
15. T. Robert *et al.*, Functional Beta Cell Mass from Device-Encapsulated hESC-Derived Pancreatic Endoderm Achieving Metabolic Control. *Stem cell reports* **10**, 739-750 (2018).
16. A. R. Pepper *et al.*, Posttransplant Characterization of Long-term Functional hESC-Derived Pancreatic Endoderm Grafts. *Diabetes* **68**, 953-962 (2019).

17. A. Veres *et al.*, Charting cellular identity during human in vitro β-cell differentiation. *Nature* **569**, 368-373 (2019).
18. E. A. Rosado-Olivieri, K. Anderson, J. H. Kenty, D. A. Melton, YAP inhibition enhances the differentiation of functional stem cell-derived insulin-producing β cells. *Nature communications* **10**, 1464 (2019).
19. H. Liu *et al.*, Chemical combinations potentiate human pluripotent stem cell-derived 3D pancreatic progenitor clusters toward functional β cells. *Nature communications* **12**, 3330 (2021).
20. K. Verhoeff, S. J. Henschke, B. A. Marfil-Garza, N. Dadheech, A. M. J. Shapiro, Inducible Pluripotent Stem Cells as a Potential Cure for Diabetes. *Cells* **10** (2021).
21. A. Lewis *et al.*, Organ donation in the US and Europe: The supply vs demand imbalance. *Transplantation reviews (Orlando, Fla.)* **35**, 100585 (2021).
22. P. Augsornworawat, K. G. Maxwell, L. Velazco-Cruz, J. R. Millman, Single-Cell Transcriptome Profiling Reveals β Cell Maturation in Stem Cell-Derived Islets after Transplantation. *Cell reports* **32**, 108067 (2020).
23. Y. Du *et al.*, Human pluripotent stem-cell-derived islets ameliorate diabetes in non-human primates. *Nature medicine* **28**, 272-282 (2022).
24. H. Hiyoshi *et al.*, Characterization and reduction of non-endocrine cells accompanying islet-like endocrine cells differentiated from human iPSC. *Scientific reports* **12**, 4740 (2022).
25. S. Ogaki, M. Morooka, K. Otera, S. Kume, A cost-effective system for differentiation of intestinal epithelium from human induced pluripotent stem cells. *Scientific reports* **5**, 17297 (2015).
26. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature biotechnology* **36**, 411-420 (2018).
27. T. Stuart *et al.*, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821 (2019).
28. H. Li *et al.*, Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nature genetics* **49**, 708-718 (2017).
29. C. Trapnell *et al.*, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature biotechnology* **32**, 381-386 (2014).
30. X. Qiu *et al.*, Single-cell mRNA quantification and differential analysis with Census. *Nature methods* **14**, 309-315 (2017).
31. X. Qiu *et al.*, Reversed graph embedding resolves complex single-cell trajectories. *Nature methods* **14**, 979-982 (2017).
32. Y. Hirozane *et al.*, Structure-based rational design of staurosporine-based fluorescent probe with broad-ranging kinase affinity for kinase panel application. *Bioorganic & medicinal chemistry letters* **29**, 126641 (2019).
33. H. A. Russ *et al.*, Controlled induction of human pancreatic progenitors produces functional beta-like cells in vitro. *The EMBO journal* **34**, 1759-1772 (2015).
34. M. B. K. Petersen *et al.*, Single-Cell Gene Expression Analysis of a Human ESC Model of Pancreatic Endocrine Development Reveals Different Paths to β-Cell Differentiation. *Stem cell reports* **9**, 1246-1261 (2017).

35. J. S. Kim *et al.*, In situ application of hydrogel-type fibrin-islet composite optimized for rapid glycemic control by subcutaneous xenogeneic porcine islet transplantation. *Journal of controlled release : official journal of the Controlled Release Society* **162**, 382-390 (2012).
36. B. F. Salama, K. L. Seeberger, G. S. Korbutt, Fibrin supports subcutaneous neonatal porcine islet transplantation without the need for pre-vascularization. *Xenotransplantation* **27**, e12575 (2020).
37. J. Jeon, M. Correa-Medina, C. Ricordi, H. Edlund, J. A. Diez, Endocrine cell clustering during human pancreas development. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **57**, 811-824 (2009).
38. F. C. Pan, C. Wright, Pancreas organogenesis: from bud to plexus to gland. *Developmental dynamics : an official publication of the American Association of Anatomists* **240**, 530-565 (2011).
39. B. A. Sullivan, J. Hollister-Lock, S. Bonner-Weir, G. C. Weir, Reduced Ki67 Staining in the Postmortem State Calls Into Question Past Conclusions About the Lack of Turnover of Adult Human β -Cells. *Diabetes* **64**, 1698-1702 (2015).
40. S. T. Chen, S. H. Fu, S. Hsu, Y. Y. Huang, B. R. Hsu, Synergistic Effect of Hyperglycemia and p27(kip1) Suppression on Adult Mouse Islet Beta Cell Replication. *International journal of endocrinology* **2012**, 417390 (2012).
41. R. Rodriguez-Diaz *et al.*, Paracrine Interactions within the Pancreatic Islet Determine the Glycemic Set Point. *Cell metabolism* **27**, 549-558.e544 (2018).
42. Y. Suissa *et al.*, Gastrin: a distinct fate of neurogenin3 positive progenitor cells in the embryonic pancreas. *PloS one* **8**, e70397 (2013).
43. N. M. George, C. E. Day, B. P. Boerner, R. L. Johnson, N. E. Sarvetnick, Hippo signaling regulates pancreas development through inactivation of Yap. *Molecular and cellular biology* **32**, 5116-5128 (2012).
44. B. Dahal Lamichane *et al.*, AGR2 is a target of canonical Wnt/β -catenin signaling and is important for stemness maintenance in colorectal cancer stem cells. *Biochemical and biophysical research communications* **515**, 600-606 (2019).
45. R. L. Panek *et al.*, In vitro biological characterization and antiangiogenic effects of PD 166866, a selective inhibitor of the FGF-1 receptor tyrosine kinase. *The Journal of pharmacology and experimental therapeutics* **286**, 569-577 (1998).
46. Y. Yamashita-Sugahara *et al.*, An inhibitor of fibroblast growth factor receptor-1 (FGFR1) promotes late-stage terminal differentiation from NGN3+ pancreatic endocrine progenitors. *Scientific reports* **6**, 35908 (2016).
47. J. Van den Bossche *et al.*, Spotlight on Volasertib: Preclinical and Clinical Evaluation of a Promising Plk1 Inhibitor. *Medicinal research reviews* **36**, 749-786 (2016).
48. B. Pajak *et al.*, 2-Deoxy-d-Glucose and Its Analogs: From Diagnostic to Therapeutic Agents. *International journal of molecular sciences* **21** (2019).
49. D. Doi *et al.*, Pre-clinical study of induced pluripotent stem cell-derived dopaminergic progenitor cells for Parkinson's disease. *Nature communications* **11**, 3369 (2020).
50. F. A. Barr, H. H. Silljé, E. A. Nigg, Polo-like kinases and the orchestration of cell division. *Nature reviews. Molecular cell biology* **5**, 429-440 (2004).

51. N. Shiraki *et al.*, Methionine metabolism regulates maintenance and differentiation of human pluripotent stem cells. *Cell metabolism* **19**, 780-794 (2014).
52. S. Tohyama, K. Fukuda, "Future Treatment of Heart Failure Using Human iPSC-Derived Cardiomyocytes" in Etiology and Morphogenesis of Congenital Heart Disease: From Gene Function and Cellular Interaction to Morphology, T. Nakanishi *et al.*, Eds. (Springer Copyright 2016, The Author(s). Tokyo, 2016), 10.1007/978-4-431-54628-3_4, pp. 25-31.
53. H. Okano, S. Yamanaka, iPS cell technologies: significance and applications to CNS regeneration and disease. *Molecular brain* **7**, 22 (2014).
54. H. Pelicano, D. S. Martin, R. H. Xu, P. Huang, Glycolysis inhibition for anticancer treatment. *Oncogene* **25**, 4633-4646 (2006).
55. G. Yu, L. G. Wang, Y. Han, Q. Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology* **16**, 284-287 (2012).
56. A. J. Vegas *et al.*, Long-term glycemic control using polymer-encapsulated human stem cell-derived beta cells in immune-competent mice. *Nature medicine* **22**, 306-311 (2016).
57. O. Veiseh *et al.*, Size- and shape-dependent foreign body immune response to materials implanted in rodents and non-human primates. *Nature materials* **14**, 643-651 (2015).
58. A. Jain, G. Tuteja, TissueEnrich: Tissue-specific gene enrichment analysis. *Bioinformatics (Oxford, England)* **35**, 1966-1967 (2019).
59. D. Balboa *et al.*, Functional, metabolic and transcriptional maturation of human pancreatic islets derived from stem cells. *Nature biotechnology* 10.1038/s41587-022-01219-z (2022).
60. V. Makker *et al.*, Lenvatinib plus Pembrolizumab for Advanced Endometrial Cancer. *The New England journal of medicine* **386**, 437-448 (2022).
61. M. Markman, Taxanes in the management of gynecologic malignancies. *Expert review of anticancer therapy* **8**, 219-226 (2008).
62. H. Nomura *et al.*, Effect of Taxane Plus Platinum Regimens vs Doxorubicin Plus Cisplatin as Adjuvant Chemotherapy for Endometrial Cancer at a High Risk of Progression: A Randomized Clinical Trial. *JAMA oncology* **5**, 833-840 (2019).
63. A. Ramzy *et al.*, Implanted pluripotent stem-cell-derived pancreatic endoderm cells secrete glucose-responsive C-peptide in patients with type 1 diabetes. *Cell stem cell* **28**, 2047-2061.e2045 (2021).
64. N. J. Hogrebe, P. Augsornworawat, K. G. Maxwell, L. Velazco-Cruz, J. R. Millman, Targeting the cytoskeleton to direct pancreatic differentiation of human pluripotent stem cells. *Nature biotechnology* **38**, 460-470 (2020).

Summary in Japanese

ヒト人工多能性幹細胞（iPSCs, induced pluripotent stem cells）は、無制限に増殖可能であり、なおかつ種々の細胞への分化能を持つ。この性質により、iPSCs を用いた細胞医療の開発が、ドナー不足という課題を抱える臓器移植の代替候補として進められている。しかし、iPSCs の無限の増殖能や分化能は、目的外の細胞への分化や増殖、さらには腫瘍形成のリスクを内包する。そのため、iPSCs 由来製品の開発においては、目的外細胞の汚染リスクを最小限に抑えることが不可欠である。特に、多能性幹細胞由来の膵内分泌細胞を用いた細胞医療に関しては、1 型糖尿病に対する臨床試験がすでに開始されており、潜在的な目的外細胞の同定と除去が急務である。そこで本論文では、iPSCs 由来の膵内分泌細胞の誘導過程で混入する目的外細胞を幅広く同定し、その特性解析を行った。さらに、目的外細胞の特性に基づく除去方法を提案し、これらの除去方法が *in vivo* 移植試験においても有効かどうかを検証した。

第 1 章では、iPSCs 由来膵内分泌細胞の誘導過程で混入する主要な目的外細胞である、CHGA 陰性の非内分泌細胞に焦点を当てた。単一細胞レベルの解析から、非内分泌細胞の大部分は、膵管様細胞や膵前駆細胞といった PDX1 陽性の膵臓系譜の細胞であることが明らかとなった。また、これらの膵臓系譜の非内分泌細胞は、*FGFR2*、*PLK1/4*、ならびに解糖系の遺伝子を特異的に高発現していた。そこで、それぞれの遺伝子に対応する化合物処理を行った結果、非内分泌細胞を特異的に減少させることに成功した。このように、目的外細胞の除去には、その特性に応じた蛋白質や代謝経路を標的とする化合物処理が有効であることが示された。

膵臓系譜の非内分泌細胞削減に有効であった化合物の中には、キナーゼ阻害薬である PD-166866 および TR06141363 が含まれる。これらのキナーゼ阻害薬の処置を介して作成した s7-iPICs (seven-stage iPSC-derived pancreatic islet cells)では、CHGA 陰性の非内分泌細胞が *in vitro* においてほとんど検出されないレベルにまで減少した。さらに s7-iPICs を 10^6 オーダーで移植した場合、膵臓系譜の目的外細胞の増殖による移植片の明確な異常は認めなかった。そのため、s7-iPICs では移植後の安全性に関する懸念は解消されたと思われた。

しかし、臨床応用時には 10^8 オーダーの細胞移植が想定される。そこで、s7-iPICs を繰り返し移植することで、大量の細胞を移植した際の安全性評価を試みた。その結果、s7-iPICs の移植片において時折、膵臓系譜ではない予期せぬ目的外細胞が増殖することが判明した。したがって、第 2 章では、この膵臓系譜外の新規目的外細胞の詳細な解析および除去方法の検討を行った。

新たに同定された目的外細胞は、間葉系幹細胞（MSCs, mesenchymal stem cells）と平滑筋細胞（SMC, smooth muscle cells）の形質を有すことが明らかとなったため、PMSCs (proliferative MSC- and SMC-like cells)と呼称した。PMSCs は、臨床応用を想定して 10^8 オーダーの s7-iPICs を移植した場合には、ほぼ 100%出現することが示唆された。したがって、PMSCs を除去しない限り、臨床応用の安全性を担保するのは困難であると考えられた。また、PMSCs は移植前にはその特徴である形質を発現しておらず、移植前の検出が極めて困難であることが判明した。そこで、移植に依らない PMSCs 検出方法として、4 週間の延長培養系を確立した。延長培養系を介して PMSCs 除去に有効な化合物を探索した結果、特にタキサン系抗がん剤であるドセタキセルが PMSCs 除去に有効であることを見出した。最後に、ドセタキセル処理を施した s7-iPICs の移植試験を繰り返し実施し、累計で 10^8 オーダーの細胞を移植した場合であっても、PMSCs が出現しないことを確認した。

以上より、本研究では、多能性幹細胞由来の膵内分泌細胞の臨床応用に向け、分化誘導過程や移植後に生じうる目的外細胞の特性を詳細に分析した。得られた特性に応じた目的外細胞の除去方法の中から、最終的に辿り着いたドセタキセル処理と s7-iPICs の組み合わせは、想定されるすべての目的外細胞に対処可能であった。したがって、ドセタキセル処理を施した s7-iPICs を用いることで、高い安全性を担保しつつ、1 型糖尿病の寛解を目指した臨床応用を進めることが可能と思われる。また、ドセタキセル処理を施した s7-iPICs を見出す過程において確立した、目的外細胞の検出と除去に関する方法論は、多能性幹細胞由来製品を用いた細胞医療の分野で広く応用可能である。本研究で得られた知見により、多能性幹細胞由来の膵内分泌細胞の臨床応用が加速し、将来的には獣医療を含め、糖尿病に苦しむ患者様に対する根治療法の提供が一日でも早く実現することを期待する。