Diabetologia (2015) 58:1013–1023
DOI 10.1007/s00125-015-3507-x

ARTICLE

SOX4 cooperates with neurogenin 3 to regulate endocrine pancreas formation in mouse models

Eric E. Xu · Nicole A. J. Krentz · Sara Tan · Sam Z. Chow · Mei Tang · Cuilan Nian · Francis C. Lynn

Received: 23 August 2014 / Accepted: 9 January 2015 / Published online: 5 February 2015

Abstract
Aims/hypothesis The sex-determining region Y (SRY)-related high mobility group (HMG) box (SOX) family of transcription factors is essential for normal organismal development. Despite the longstanding knowledge that many SOX family members are expressed during pancreas development, a role for many of these factors in the establishment of insulin-producing beta cell fate remains to be determined. The aim of this study is to elucidate the role of SOX4 during beta cell development.
Methods We used pancreas and endocrine progenitor mouse knockouts of *Sox4* to uncover the roles of SOX4 during pancreas development. Lineage tracing and in vitro models were used to determine how SOX4 regulates beta cell formation and understand the fate of *Sox4*-null endocrine lineage cells.
Results This study demonstrates a progenitor cell-autonomous role for SOX4 in regulating the genesis of beta cells and shows that it is required at multiple stages of the process. SOX4 deletion in the multipotent pancreatic progenitors resulted in impaired endocrine progenitor cell differentiation. Deletion of SOX4 later in the *Neurog3*-expressing cells also caused reductions in beta cells. Lineage studies showed loss of *Sox4* in endocrine progenitors resulted in a block in terminal islet cell differentiation that was attributed to reduction in the production of key beta cell specification factors.
Conclusions/interpretation These results demonstrate that SOX4 is essential for normal endocrine pancreas development both concomitant with, and downstream of, the endocrine fate decision. In conclusion, these studies position *Sox4* temporally in the endocrine differentiation programme and provide a new target for improving in vitro differentiation of glucose-responsive pancreatic beta cells.

Keywords Basic science · Cell lines · Human · Islet development · Islet transplantation · Islets · Knockout mice · Transcription factors

Abbreviations

$^{+}$	Hormone-immunoreactive (cells)
CHGA	Chromogranin A
ChIP	Chromatin immunoprecipitation
Cre	Cre recombinase
E	Embryonic day
EMS4KO	*Neurog3*-Cre mT/mG *Sox4*$^{flox/flox}$
ES4KO	*Neurog3*-Cre *Sox4*$^{flox/flox}$
GCG	Glucagon
GF	Green fluorescence
(e)GFP	(enhanced) Green fluorescent protein
GHRL	Ghrelin
HA	Haemagglutinin
hESC	Human embryonic stem cell
HMG	High mobility group
INS	Insulin
Mafa	v-maf Avian musculoaponeurotic fibrosarcoma oncogene homologue A

Electronic supplementary material The online version of this article (doi:10.1007/s00125-015-3507-x) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

E. E. Xu · N. A. J. Krentz · S. Tan · S. Z. Chow · M. Tang · C. Nian · F. C. Lynn (✉)
Diabetes Research Program, Child and Family Research Institute, A4-184, 950 West 28 Ave, Vancouver, BC V5Z 4H4, Canada
e-mail: francis.lynn@ubc.ca

E. E. Xu · N. A. J. Krentz · M. Tang · C. Nian · F. C. Lynn
Department of Surgery, University of British Columbia, Vancouver, BC, Canada

F. C. Lynn
Department of Cellular and Physiological Sciences, University of British Columbia, Vancouver, BC, Canada

mPAC	Mouse pancreatic ductal adenocarcinoma
mT/mG	Membrane-targeted tdTomato/membrane-targeted enhanced GFP
NEUROD1	Neurogenic differentiation 1
NEUROG3	Neurogenin 3
NKX2.2	NK2 homeobox 2
NKX6.1	NK6 homeobox 1
ONECUT1	One cut homeobox 1
PAX4	Paired box 4
PDX1	Pancreatic and duodenal homeobox 1
PPY	Pancreatic polypeptide
PS4KO	*Pdx1*-Cre *Sox4*$^{flox/flox}$
PTF1A	Pancreas transcription factor 1 subunit α
q(PCR)	Quantitative (PCR)
RF	Red fluorescence
RFP	Red fluorescent protein
RFX6	Regulatory factor X, 6
SOX	SRY-related high mobility group box
SRY	Sex-determining region Y
SST	Somatostatin
WT	Wild type

Introduction

In mice, pancreas development begins when pancreatic and duodenal homeobox 1 (PDX1), pancreas transcription factor 1 subunit α (PTF1A) and sex-determining region Y (SRY)-related high mobility group (HMG) box (SOX)9-expressing epithelial cells bud from the posterior foregut endoderm at embryonic day (E) 9 [1–3]. Signals from adjacent tissues support budding and as development proceeds, two waves of differentiation occur [4–8]. The primary transition, from E8.5–10.5, is marked by low numbers of hormone-immunoreactive ($^{+}$) cells; the secondary transition, from E12.5–15.5, is denoted by a large wave of exocrine and endocrine cell differentiation [9]. The restriction towards a terminally differentiated endocrine cell occurs progressively [10].

Endocrine differentiation requires inhibition of NOTCH signalling and transient activation of neurogenin 3 (NEUROG3) in epithelial trunk progenitor cells, and is accompanied by exit from the cell cycle and migration into the mesenchyme [11–18]. Downstream, regulators of beta cell specification include: paired box 4 (PAX4), neurogenic differentiation 1 (NEUROD1), regulatory factor X, 6 (RFX6), NK2 homeobox 2 (NKX2.2), NK6 homeobox 1 (NKX6.1) and v-maf avian musculoaponeurotic fibrosarcoma oncogene homologue A (MAFA), among others [19–25]. While a number of these factors are essential for normal beta cell development, it is unclear if any transcription factor couples the endocrine progenitor fate choice with downstream terminal differentiation. Here, we demonstrate that SOX4 may play this bridging role during endocrine cell genesis.

SOX proteins are an evolutionarily ancient family of complex-forming transcriptional activators and repressors that are identified by a conserved HMG DNA-binding domain [26, 27]. A plethora of SOX genes has been demonstrated to be expressed within the developing pancreas, including most of the C, D, E, F, G and H subfamily members [28]. Early during pancreas formation, expression of an E group family member, SOX9, is critical for the maintenance of pancreatic progenitor identity [1].

Previous work has demonstrated SOX4 expression in the embryonic human, mouse and zebrafish pancreas [28–30]. Null mutation of *Sox4* results in cardiac abnormalities, defects in blood flow and embryonic lethality by E14.5 prior to completion of the bulk of endocrine cell differentiation. However, significant reductions of insulin (INS) and/or glucagon (GCG) expression were observed by us and others [29, 31].

This study expands on these observations to define the spatiotemporal importance of SOX4 during endocrine formation in vivo. SOX4 is demonstrated to be required in a cell-autonomous capacity for efficient endocrine cell differentiation both prior to and downstream of NEUROG3. As such, the function of SOX4 during pancreatic endocrine cell formation mirrors roles of SOX family genes in other systems: to bridge key developmental transitions and couple fate decisions.

Methods

Chemicals and reagents Chemicals were purchased from Sigma Aldrich (Oakville, ON, Canada) or Fisher Scientific (Ottawa, ON, Canada). The antibodies used are listed in electronic supplementary material (ESM) Table 1. Oligonucleotides were purchased from IDT (Coralville, IA, USA; ESM Table 2). Tissue culture reagents were from Hyclone (Logan, UT, USA) and cultureware from BD-Falcon (Fisher Scientific).

Tissues The University of British Columbia (UBC) Animal Care Committee approved all mouse breeding and experiments, with E0.5 set as noon after discovery of a vaginal plug. Mouse strains included membrane-targeted tdTomato/membrane-targeted enhanced (e) green fluorescent protein (GFP; mT/mG) [32], *Neurog3*-Cre recombinase (Cre) [33] (Jackson Laboratory, Bar Harbor, ME, USA), *Pdx1*-Cre [17] (Mutant Mouse Regional Resource Centers [MMRRC], Davis, CA, USA) and *Sox4*^{flox} [34]. *Pdx1*-Cre *Sox4*$^{flox/flox}$ (PS4KO) and *Neurog3*-Cre *Sox4*$^{flox/flox}$ (ES4KO) controls were Cre negative and *Sox4*$^{flox/WT}$ or *Sox4*$^{flox/flox}$ littermates (where WT is wild type). *Neurog3*-Cre mT/mG *Sox4* $^{flox/flox}$ (EMS4KO) controls were: *Neurog3*-Cre mT/mG$^{flox/WT}$ or mT/mG$^{flox/flox}$ and *Sox*$^{flox/WT}$ double heterozygotes. The

University of Alberta Research and Transplant Islet Isolation Cores obtained human islets under informed consent.

Tissue harvest and preparation Tissue was harvested and fixed in 4% (wt/vol.) paraformaldehyde in PBS (4°C). Tissue samples were then dehydrated and paraffin embedded or frozen as described [35]. Sections, 5 μm paraffin or 10 μm frozen, were rehydrated as necessary prior to antigen retrieval.

Immunostaining and imaging Please see ESM for antibodies, in situ hybridisation, TUNEL and 5-ethynyl-2′-deoxyuridine (EdU) staining methods. Antigen retrieval was performed in citrate buffer (pH 6 at 95°C for 20 min). Slides were blocked for 1 h each in 5% (vol./vol.) horse serum followed by donkey anti-mouse IgG. Sections were incubated overnight at 4°C with primary antibody and for 1 h at 20°C with secondary antibodies plus nuclear stains. Tyramide signal amplification (TSA) was carried out for detection of SOX4 (Perkin Elmer, Woodbridge, ON, Canada). Imaging was carried out with sequential scanning on a Leica SP8 confocal imaging system (Concord, ON, Canada).

Morphometric analyses and quantification Serial sections were chosen at 100 or 50 μm intervals for E18.5 or E15.5 embryos, respectively. The whole section was imaged and tiled and immunopositive cells were counted using CellProfiler (Cambridge, MA, USA) [36]. Counts were normalised to total pancreatic nuclei using CellProfiler or the sum of pancreatic area calculated using Image-Pro Analyzer (Media Cybernetics, Rockville, MD, USA).

Adenoviral generation, infection and cell culture *Sox4* adenovirus was generated using the pAdTrack system (Addgene, Cambridge, MA, USA) [37]. Viruses were purified with the Adeno-X Maxi adenovirus purification system (Clontech, Mountain View, CA, USA). Please see ESM Methods for gene targeting, human embryonic stem cells (hESCs) and FACS methods. Mouse pancreatic ductal adenocarcinoma (mPACs) cells were infected at a multiplicity of infection (MOI) 100:1 for 2 h. Cells were washed briefly in PBS, media was replaced and they were cultured for 48 h.

Real-time PCR RNA was isolated with TRIzol and DNAse treated and cDNA was synthesised using Superscript II reverse transcriptase (Life Technologies, Burlington, ON, Canada). A 40 ng sample of cDNA was used as template for quantitative (q)PCR carried out in a ViiA7 qPCR machine (Life Technologies). Expression was determined using the $\Delta\Delta C_t$ method with glucouronidase B (*Gusb*) as a reference gene [35].

Luciferase assay mPACL20 cells were seeded (2×10^5 cells) in 24-well plates and transfected using Lipofectamine 2000 (Life Technologies) with 1 μg reporter, 50 ng cytomegalovirus promoter (pCMV)-renilla luciferase and 100 ng expression vectors, balanced with empty expression vector. Cells were incubated for 48 h prior to determination of luciferase activity using the Dual-Luciferase Reporter (DLR) Assay (Promega, Madison, WI, USA) on a SpectraMaxL luminometer (Molecular Devices, Sunnyvale, CA, USA).

Chromatin immunoprecipitation Chromatin immunoprecipitation (ChIP) was performed as previously described [35]. The only change being that ChIP was carried out in the presence of 160 mmol/l NaCl, 0.25% (wt/vol.) deoxycholate and 0.025% (vol./vol.) Triton X-100.

Statistical analysis All statistical analyses were performed with Prism5 (GraphPad Software, La Jolla, CA, USA). Two-tailed unpaired Student's *t* tests or one-way ANOVA with Tukey post-hoc tests were performed, with $p\leq0.05$ considered significant. Error bars represent SEM.

Results

SOX4 is dynamically expressed during mouse pancreas formation As seen in Fig. 1a, *Sox4* is expressed throughout pancreas development, with the highest levels at E11.5. A gradual decline in *Sox4* message was observed; however, significant expression was observed during the secondary transition when the majority of NEUROG3$^+$ endocrine progenitors are derived (E12.5–14; Fig. 1a). During human embryonic stem cell differentiation *SOX4* is expressed at high levels in stage 3 and 4 pancreatic progenitors (Fig. 1b).

Because suitable antibodies were not initially available, in situ hybridisation was carried out to localise *Sox4* expression (ESM Fig. 1). The epithelium at this stage was cord-like with GCG$^+$ cells found in proximity to the trunk (ESM Fig. 1a). Low expression of *Sox4* mRNA was observed throughout the epithelium (E14.5; ESM Fig. 1b). Occasional cells expressing higher levels of *Sox4* were noted scattered through the epithelium (arrowheads ESM Fig. 1b): an expression pattern reminiscent of the bipotent trunk cells that activate NEUROG3 [13, 38]. Next, immunostaining was used to determine the spatiotemporal distribution of SOX4. Early during pancreatic organogenesis (E11.5 and E12.5), SOX4 co-localised with PDX1 and NKX6.1 in the multipotent progenitors and trunk cells respectively (Fig. 1c, d; ESM Fig. 2b) [39]. Following the secondary transition, SOX4 was progressively excluded from pancreatic epithelial cells; however, it remained highly expressed in all NEUROG3-expressing endocrine progenitors (Fig. 1e, f) and in nascent endocrine cells (ESM Fig. 2b). By E17.5, SOX4 expression was confined largely to endocrine cells (ESM Fig. 2d).

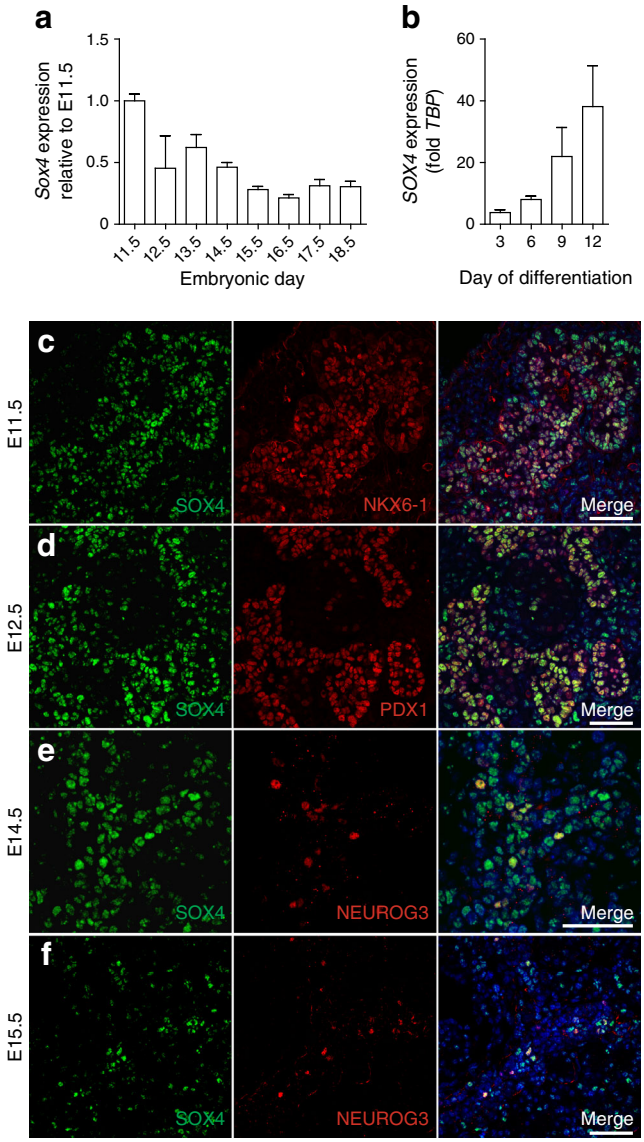

Fig. 1 SOX4 is expressed during pancreas development and becomes restricted to endocrine cells. qPCR analyses of *Sox4* expression in E11.5 to E18.5 mouse pancreas (**a**) and following directed differentiation of hESCs (**b**). Coimmunostaining of SOX4 (green) (**c**–**f**) with: (**c**) NKX6-1 (red), E11.5; (**d**) PDX1 (red), E12.5; and (**e**, **f**) NEUROG3 (red) (**e**) E14.5 and (**f**) E15.5. $n \geq 3$; scale bars, 50 μm. TBP, TATA-binding protein

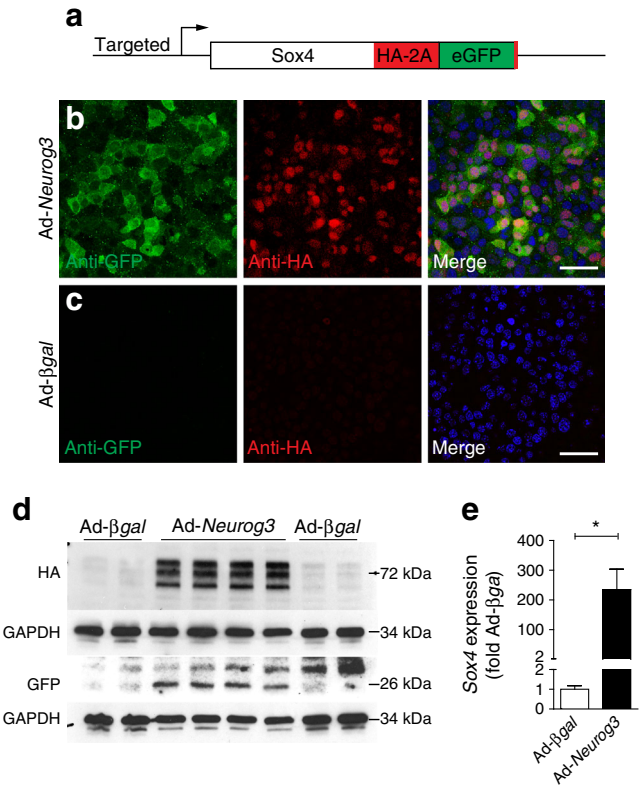

Fig. 2 Endogenous SOX4 is induced by NEUROG3 in the mPAC model of endocrine development. Schematic of *Sox4*-HA-2A-eGFP knock-in add-on allele; 3′ red box is stop codon (**a**; see ESM Methods for details). Immunofluorescence analyses show induction of HA-tagged SOX4 and GFP following treatment with Ad-*Neurog3* in mPAC-*Sox4*-HA cells (**b**) but not in control cells (**c**). Western blotting of mPAC-*Sox4*-HA cell lysates showed an increase in HA and GFP immunoreactivity at predicted electrophoretic mobilities (~47 kDa and ~70 kDa^{+}) following Ad-*Neurog3* transduction (**d**). *Sox4* expression is increased by Ad-*Neurog3* in unmodified mPACL20 cells (**e**). $n \geq 3$; $^{*}p \leq 0.05$. *βgal* is also known as *bgaL*. GAPDH, glyceraldehyde-3-phosphate dehydrogenase

To verify that SOX4 was expressed in NEUROG3^{+} endocrine progenitor cells, the mouse mPACL20 pancreatic ductal cell line, which has previously been used to model endocrine cell differentiation, was used [40, 41]. Using transcription activator-like effector nuclease (TALEN)-driven homologous recombination in the mPACL20 cells, a haemagglutinin (HA) tag followed by 2A-GFP was knocked onto *Sox4* (mPAC-*Sox4*-HA; Fig. 2a) and 11 clonal lines were derived with 23% targeting efficiency.

As predicted from the immunostaining results (Fig. 1c–f), when this clonal mPAC-*Sox4*-HA cell line was transduced with a *Neurog3* adenovirus (Ad-*Neurog3*) [41], marked increases in endogenous promoter activity led to appreciable HA and GFP that were detectable by both immunostaining (Fig. 2b, c) and western blotting (Fig. 2d). In support, unmodified mPACL20 cells transduced with Ad-*Neurog3* showed a 240-fold increase in *Sox4* message expression (Fig. 2e).

These observations show that SOX4 is widely produced in the multipotent pancreatic progenitors prior to endocrine differentiation and that its expression is maintained within the endocrine progenitors. In order to assess whether there is an important cell-autonomous role for SOX4 during pancreatic or endocrine cell formation, loss-of-function approaches were used.

SOX4 is necessary for endocrine progenitor cell formation To determine if SOX4 is important prior to endocrine progenitor formation, mice harbouring a floxed *Sox4* allele (*Sox4*^{flox}) were crossed with *Pdx1*-Cre transgenic mice to generate

pancreatic *Sox4*-knockout embryos (PS4KO). Lineage analyses indicate that recombination is extremely efficient early during pancreas organogenesis (>95% of epithelial cells at E12.5; ESM Fig. 2a). In agreement, SOX4 immunostaining at E15.5 verified that SOX4 protein level was significantly reduced compared with controls (ESM Fig. 2b, c).

Quantification of endocrine cell types at E18.5 revealed that pancreatic deletion of *Sox4* significantly reduced INS$^+$ GCG$^+$ (Fig. 3a, b, i, j), somatostatin$^+$ ([SST$^+$] Fig. 3c, d, l), and pancreatic polypeptide$^+$ ([PPY$^+$] Fig. 3e, f, m) cells, without changes in ghrelin$^+$ (GHRL$^+$) cells (Fig. 3c, d, k). To address the possibility of a defect in early endocrine cell differentiation, bipotent SOX9$^+$ cells and NEUROG3$^+$ endocrine progenitor cells were quantified. A significant reduction in NEUROG3$^+$ cells, but not SOX9$^+$ cells, was observed (Fig. 3g, h, n, o); suggesting that SOX4 is dispensable until endocrine differentiation is initiated within the SOX9$^+$ progenitor [38, 42]. As expected, neither apoptosis (ESM Fig. 3e–h) nor proliferation (ESM Fig. 3c, d) accounted for the differences in NEUROG3$^+$ cells at E15.5.

Despite the marked phenotype and loss of SOX4 protein (ESM Fig. 2c), *Sox4* message was significantly reduced but detectable in knockout pancreas by qPCR (Fig. 3p). This may result from a long half-life of *Sox4* message or expression from neurons, blood vessels and mesenchyme that are closely associated with the epithelium. Expression of related SOXC class family members, *Sox11* and *Sox12* (Fig. 3q, r), was not significantly different, suggesting their upregulation did not compensate for loss of *Sox4*. In sum, these findings demonstrate SOX4 is essential for efficient endocrine progenitor formation.

SOX4 activates the Neurog3 promoter and induces NEUROG3 expression To investigate whether the cooperative actions of NEUROG3 and SOX4 are important for endocrine cell genesis, unmodified mPACL20 pancreatic ductal cells were transduced with Ad-*Sox4* and Ad-*Neurog3* (Fig. 4a) [41, 42]. Co-transduction of mPACL20 cells resulted in a significant increase in endogenous *Neurog3* expression (Fig. 4a).

In order to determine if SOX4 binds and directly regulates NEUROG3 in the context of development, the modified mPAC-*Sox4*-HA cells were transduced with Ad-*Neurog3* to induce endogenous *Sox4* (Fig. 2e) and ChIP was carried out using anti-HA antibodies (Fig. 4b). Significant enrichment of SOX4-HA at a conserved region of the *Neurog3* promoter containing two SOX4 consensus binding sites (ESM Fig. 4) was observed. Luciferase assays next demonstrated that SOX4 can transactivate a 2 kb fragment of the *Neurog3* promoter (Fig. 4c) [43]. Collectively, these results suggest that SOX4 expression within a subset of bipotent trunk progenitors enhances NEUROG3 expression and drives subsequent endocrine progenitor cell formation.

SOX4 regulates endocrine cell differentiation downstream of NEUROG3 SOX4 expression is maintained at high levels in both mouse and human islets (ESM Figs 2 and 5); however, NEUROG3 expression is not [44]. In order to test if the

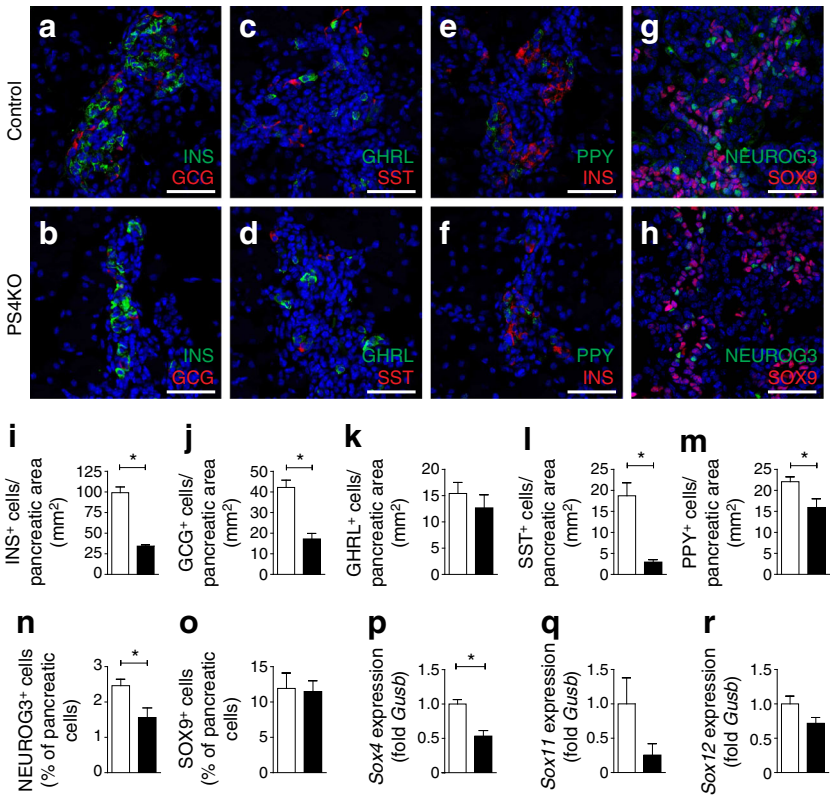

Fig. 3 PS4KO mice have defects in endocrine cell genesis. Immunohistochemical analyses of hormone immunoreactive (+) cells in control (**a**, **c**, **e**, **g**) and PS4KO (**b**, **d**, **f**, **h**) pancreas: (**a**, **b**) INS$^+$ (green) and GCG$^+$ (red), E18.5; (**c**, **d**) GHRL$^+$ (green) and SST$^+$ (red) E18.5; (**e**, **f**) PPY$^+$ (green) and INS$^+$ (red), E18.5; (**g**, **h**) NEUROG3$^+$ (green) and SOX9$^+$ (red), E15.5. PS4KO (black bars) had significantly reduced numbers of INS$^+$ (**i**), GCG$^+$ (**j**), SST$^+$ (**l**), PPY$^+$ (**m**) and NEUROG3$^+$ (**n**) cells but not in GHRL$^+$ (**k**) or SOX9$^+$ (**o**) cells compared with controls (white bars). *Sox4* expression (**p**) was significantly reduced at E15.5 in PS4KO pancreas without changes in *Sox11* (**q**) or *Sox12* (**r**). $n \geq 3$; *$p \leq 0.05$; scale bars, 50 μm; nuclei are stained blue

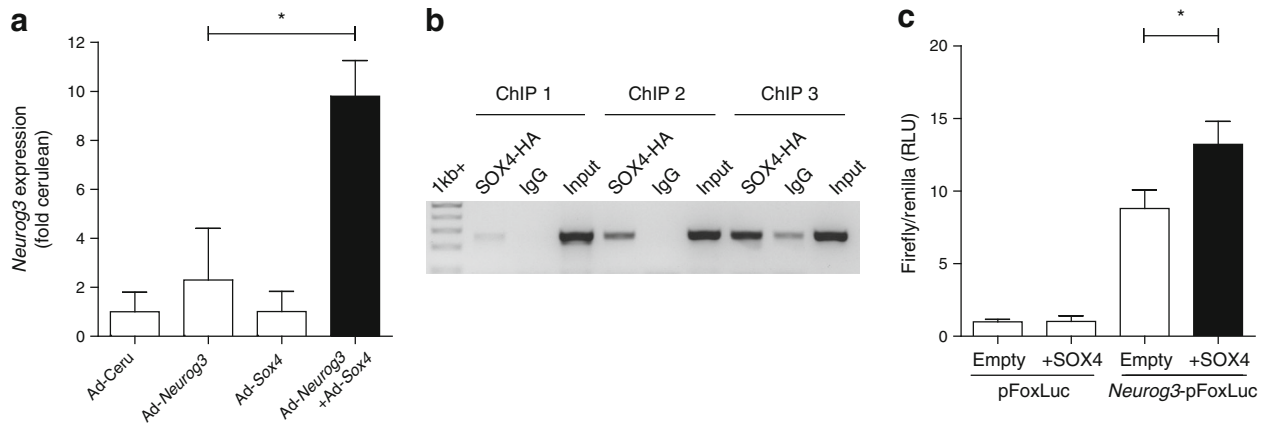

Fig. 4 SOX4 directly regulates *Neurog3* expression in the mPAC model of endocrine development. Ad-*Neurog3* and Ad-*Sox4* cooperatively increase expression from the *Neurog3* locus in mPACL20 cells (**a**). HA-ChIP (**b**) following Ad-*Neurog3* showed HA-tagged SOX4 protein enriched at the *Neurog3* promoter. Luciferase assays demonstrate that SOX4 activates a truncated 2 kb fragment of *Neurog3* in mPACL20 cells (**c**). $n \geq 3$; $^{*}p \leq 0.05$. RLU, relative light units

maintenance of *Sox4* expression downstream of *Neurog3* remains important for islet cell formation, mice harbouring *Sox4*^{flox} were crossed with the *Neurog3*-Cre transgenic mice (ES4KO). Because reduction of SOX4 expression would likely occur with some delay following *Neurog3* induction in the ES4KO mice, we hypothesised that no difference in the number of NEUROG3^{+} cells would be observed. To test this, we analysed the number of NEUROG3^{+} cells in the ES4KO pancreas at E15.5 and found no difference between genotypes (Fig. 5a–c). Despite this, quantification of endocrine cells within E18.5 ES4KO pancreas demonstrated significant reductions of INS^{+}, GCG^{+} (Fig. 5d–g), SST^{+} (Fig. 5h, i, k) and SOX4^{+} (ESM Fig. 2e) cells but not GHRL^{+} (Fig. 5h–j) or PPY^{+} (Fig. 5l–n) cells.

Loss of SOX4 does not change the number of Neurog3 lineage cells Because a significant reduction in endocrine mass was observed in the ES4KO mouse and SOX4 has previously been shown to be important for cell survival and cell proliferation [45], the effects of loss of SOX4 on proliferation (ESM Fig. 6a–c) and apoptosis (ESM Fig. 6d–g) were tested: neither was affected.

Previous studies have suggested 'dedifferentiated' endocrine lineage cells maintain expression of the neuroendocrine secretory granule protein chromogranin A (CHGA) [46]. Therefore, in order to address the possibility that islet endocrine cells do not fully differentiate in ES4KO mice, CHGA immunostaining was carried out. Remarkably, despite the reduction in INS^{+} (Figs 5d, e and 6b) and other endocrine cells (Fig. 5) no difference in the number of CHGA^{+} cells (Fig. 6a) was observed. Furthermore, co-staining for INS and GCG with CHGA uncovered a significantly increased number of cluster-forming (dotted area) CHGA-expressing cells that did not co-stain for either of these hormones (Fig. 6c, d). These data suggest that loss of SOX4 in the ES4KO impacts normal islet endocrine cell differentiation or maturation downstream of NEUROG3.

To show that these CHGA-expressing non-INS^{+}/GCG^{+} cells were derived from the *Sox4*-null *Neurog3* lineage, mice harbouring a *ROSA26*$^{mT/mG}$ reporter allele [32] were crossed with ES4KO mice (EMS4KO). In this model, Cre expression in endocrine progenitors drives a switch from red fluorescence (RF) to green fluorescence (GF). *Neurog3*-lineage GFP^{+} cells were quantified and no reduction in the total number of GFP^{+} cells in the EMS4KO mice was found (Fig. 6e), supporting the observation that neither cell proliferation nor death is impacted. Despite a similar number of pancreatic GFP^{+} cells, a significant increase in non-INS^{+}/GCG^{+}-staining GFP^{+} cells was observed in the EMS4KO mice (Fig. 6f–h). Notably, no differences in the numbers of GFP^{+} cells that expressed the acinar cell marker amylase (ESM Fig. 7a, b, e) or the ductal cell marker *Dolichos biflorus* agglutinin (DBA)-lectin (ESM Fig. 7c, d, f) were observed. These experiments reveal that *Sox4*-knockout cells remain present within the pancreas, incapable of differentiating into islet hormone-expressing endocrine cells and that they persist in a state not normally present in the pancreas, with CHGA^{+} and synaptophysin^{+} (not shown) but no islet hormone expression.

SOX4 regulates factors downstream of Neurog3 As lineage tracing revealed that *Sox4*-null *Neurog3* lineage cells remained fated to a neuroendocrine lineage, the possibility that SOX4-regulated genes important for beta cell differentiation downstream of *Neurog3* was explored. To that end, qPCR expression analyses were carried out at E17.5 (Fig. 7a–j). The expression profiles of the endocrine hormones (Fig. 7a–e) mirrored tissue morphometry results (Fig. 5). Importantly, gene expression of at least two transcription factors known to be downstream of endocrine cell specification, *Neurod1* and *Pax4*, was significantly reduced (Fig. 7f, g).

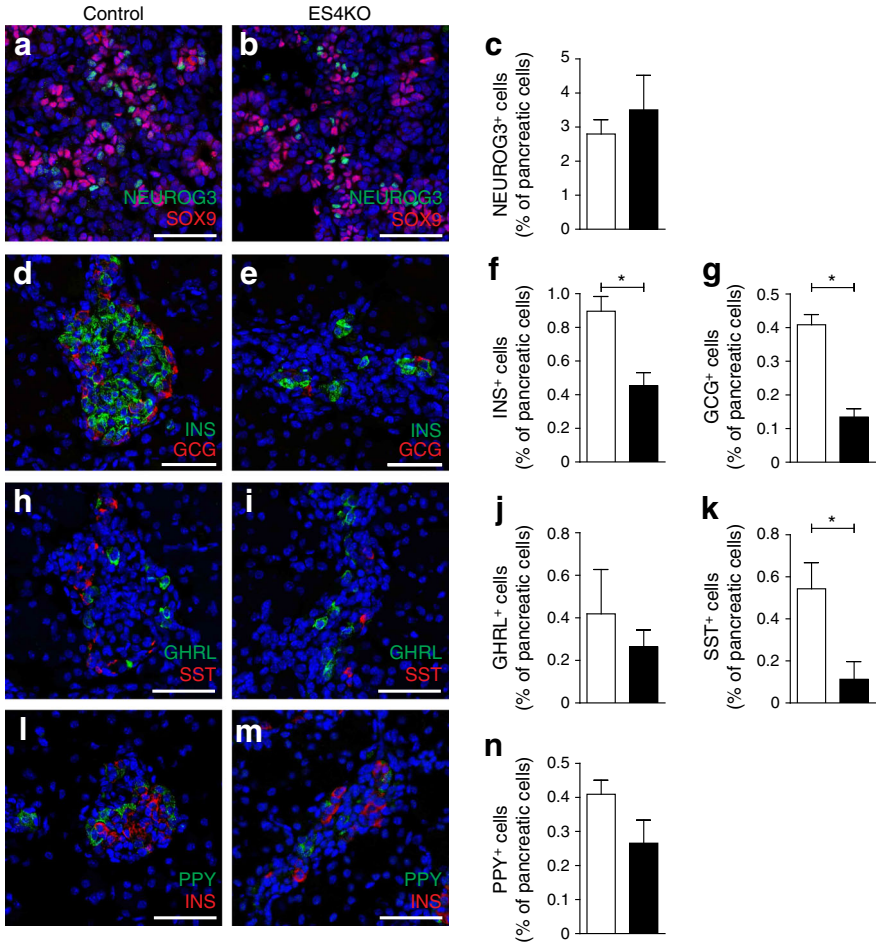

Fig. 5 ES4KO embryos have defective endocrine cell formation. Pancreatic immunostaining for (**a**, **b**) NEUROG3 (green) and SOX9 (red) and subsequent morphometric analyses showed no difference in NEUROG3$^+$ (**c**) or SOX9$^+$ cells (not shown) in the ES4KO (compared **b** vs **a**). Immunostaining in control (**d**, **h**, **l**) and ES4KO (**e**, **i**, **m**) E18.5 pancreases for: (**d**, **e**) INS$^+$ (green) and GCG$^+$ (red), (**h**, **i**) GHRL$^+$ (green) and SST$^+$ (red) and (**l**, **m**) PPY$^+$ (green) and INS$^+$ (red). Morphometric analyses showed that ES4KO (black bars) have significantly reduced INS$^+$ (**f**), GCG$^+$ (**g**) and SST$^+$ (**k**) cell populations but normal GHRL$^+$ (**j**) and PPY$^+$ (**n**) cell populations compared with controls (white bars). $n \geq 3$; *$p \leq 0.05$; scale bars, 50 μm; nuclei are stained blue

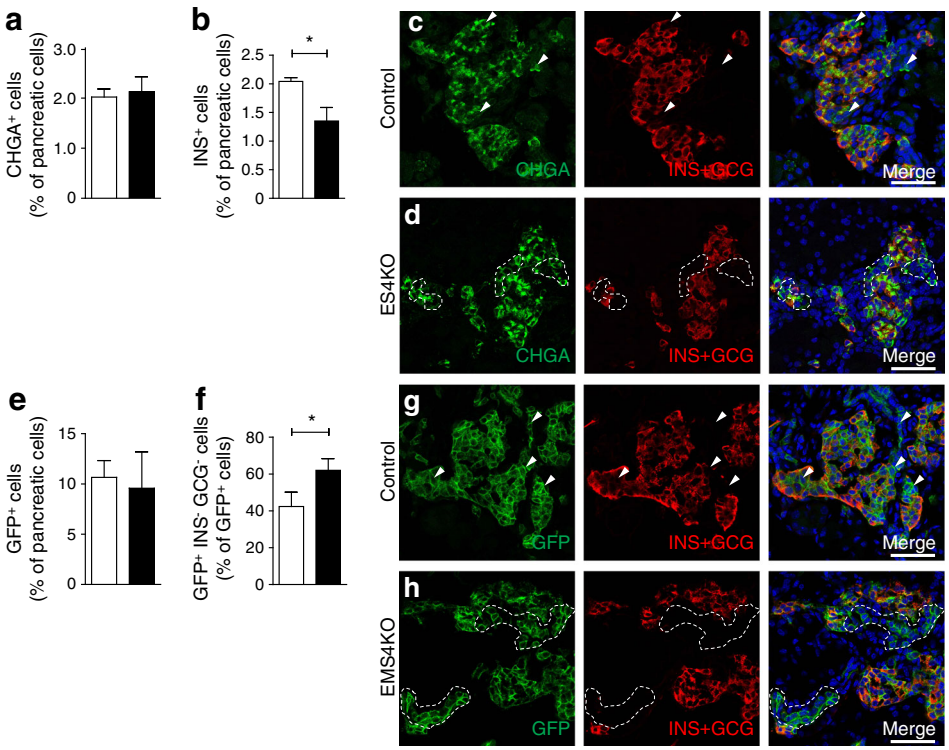

Fig. 6 *Sox4*-null endocrine progenitor cells remain in the endocrine lineage but do not differentiate into islet endocrine cells. CHGA$^+$ (**a**), but not INS$^+$ (**b**), cells are present in normal numbers in the E18.5 ES4KO (black bars). Significant numbers of CHGA$^+$ INS$^-$ GCG$^-$ cells can be appreciated in the ES4KO (**d**, dotted area) but not in control animals (**c**, arrowheads). There is no change in the total number of labelled endocrine lineage GFP$^+$ cells in the E18.5 ES4KO pancreas (**e**). However, significantly increased numbers of GFP$^+$ cells that did not co-stain for either INS or GCG were observed (**f**, **h**, dotted area) compared with littermate controls (white bars, **g**, arrowheads). $n \geq 3$; *$p \leq 0.05$; scale bars, 50 μm; nuclei are stained blue

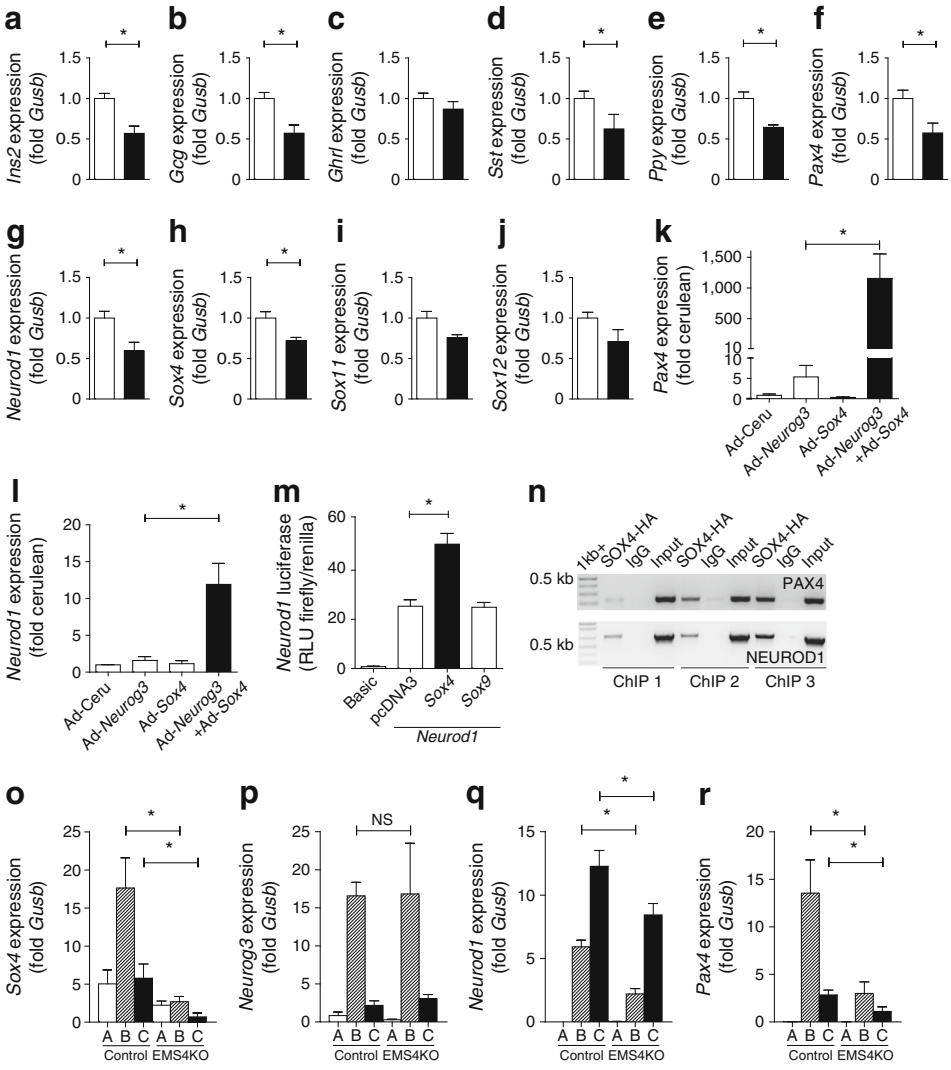

Fig. 7 SOX4 potentiates the induction of pro-beta cell genes during endocrine pancreas development. *Ins2* (**a**), *Gcg* (**b**), *Sst* (**d**), *Ppy* (**e**), *Pax4* (**f**), *Neurod1* (**g**), and *Sox4* (**h**) message levels were significantly reduced, while *Ghrl* (**c**), *Sox11* (**i**), and *Sox12* (**j**) were not in E17.5 ES4KO (black bars). *Pax4* (**k**) and *Neurod1* (**l**) expression was induced following combined Ad-*Neurog3* and Ad-*Sox4* in mPACL20 cells. SOX4 significantly induced *Neurod1* promoter-driven luciferase activity in mPACL20 cells (**m**). ChIP demonstrated that HA-tagged SOX4 was enriched at the *Pax4* and *Neurod1* promoters in the Ad-*Neurog3*-transduced mPAC-*Sox4*-HA cells (**n**; $n=3$). qPCR analyses of *Sox4* (**o**); *Neurog3* (**p**); *Neurod1* (**q**); and *Pax4* (**r**) expression in newborn endocrine cells from E17.5 control and EMS4KO mice. A, Tomato$^+$; B, Tomato$^+$ GFP$^+$ (newborn); and C, GFP$^+$ (*Neurog3* lineage). $n\geq3$; $^*p\leq0.05$. RLU, relative light units

In order to test whether SOX4 binding to the *Pax4* and *Neurod1* promoters was important for beta cell formation, mPACL20 cells were transduced with Ad-*Neurog3* and/or Ad-*Sox4*, and both *Pax4* (Fig. 7k) and *Neurod1* (Fig. 7l) expression was assessed using qPCR. Remarkably, in this system the combination of both Ad-*Neurog3* and Ad-*Sox4* induced *Pax4* expression ~200-fold and *Neurod1* expression ~tenfold compared with *Neurog3* alone. *Neurod1* promoter-driven luciferase transcription [47] was significantly induced with co-transfection of *Sox4* (Fig. 7m). HA-ChIP on Ad-*Neurog3*-transduced mPAC-*Sox4*-HA cells demonstrated interaction of SOX4-HA at potential binding sites within both the *Pax4* and *Neurod1* (Fig. 7n) promoters, providing further evidence for the direct regulation of these beta cell transcription factors by SOX4.

In order to understand if *Pax4* and *Neurod1* are SOX4 targets in vivo, FACS purification of 'newborn' *Neurog3*-lineage cells was carried out [32, 48] (Fig. 7o–r; ESM Fig. 8) using the *Neurog3*-Cre mT/mG mouse. We reasoned that newly born cells would possess both RF and GF and FACS purification of these cells demonstrated significant enrichment of both *Neurog3* and *Sox4*, indicating that SOX4 is indeed enriched in the newborn endocrine progenitor pool in vivo (Fig. 7o, p). Assessment by FACS suggests that *Neurog3*-Cre mediated SOX4 inactivation is robust in the EMS4KO line, with over 80% loss of *Sox4* message in the endocrine lineage (Fig. 7o). More robust *Sox4* message detected by whole pancreas qPCR (Fig. 7h) may be due to low-level expression in the exocrine tissues, which form the bulk of the pancreas. In agreement with in vitro studies, significant reductions in expression of *Pax4*, *Neurod1* and *Ins2* (but not aristaless related homeobox [*Arx*]) were observed in the *Neurog3*-lineage cells from the EMS4KO (Fig. 7q, r; ESM Fig. 8c, d). The previous demonstrations that *Pax4* and *Neurod1* are directly downstream of *Neurog3* and the observation that expression of these factors was reduced in the newborn (RF + GF) endocrine cells that expressed normal *Neurog3* levels (Fig. 7p), lends credence to them being direct, biologically important, SOX4 targets in vivo.

Discussion

Many of the developmental transitions that underlie the formation of beta cells are marked by expression of unique SOX proteins. In embryonic stem cells, SOX2 helps maintain the network of factors that are important for the pluripotent state [49]. SOX factors remain important downstream of SOX2 during definitive endoderm formation, when octamer-binding transcription factor 4 (OCT4) complexes with SOX17 to regulate endoderm induction [50]. SOX17 then likely interacts with undefined factors until expression wanes in the pancreatic buds at the time when SOX9 is induced [51]. SOX9 establishes the gene expression programme necessary for pancreatic budding and growth and remains important in the formation of the endocrine progenitor cells [52]; however, it is silenced in the endocrine lineage.

This study provides the first evidence of a cell-autonomous role for SOX4 during normal endocrine cell formation. Based on our data, we propose a model (Fig. 8) where SOX4 within a subset of multipotent epithelial progenitor cells stabilises *Neurog3* transcription and endocrine progenitor fate. SOX4 then couples endocrine progenitor cell differentiation with subsequent development through cooperation with NEUROG3.

SOX4 has been implicated in development of a wide array of mammalian tissues. As such, the previous pancreatic studies did not definitively show that cell-autonomous actions of SOX4 in the pancreatic epithelium were necessary [28, 31, 53]. Furthermore, the early studies did not uncover when SOX4 becomes important for islet cell formation or how it functions to regulate islet cell genesis. Finally, these previous studies did not attempt to assess whether SOX4 expression in the exocrine progenitors is important for their differentiation. Herein, we address all of these previous questions and for the first time show that SOX4 expression within the trunk cells is necessary for endocrine cell development and that loss of SOX4 impacts both formation of progenitors and terminal differentiation of all islet cells but not the exocrine tissue.

Of note, *Danio rerio Sox4b*, an orthologue of *Sox4*, was transiently expressed in endocrine precursor cells and knockdown of SOX4B uncovered a requirement for alpha cell differentiation. Other endocrine cell types, however, were unchanged [29]. The differences in observed phenotypes between our study and this former one may be attributed in part to unique endocrine differentiation pathways present in fish and mice or to the differences in knockout/knockdown approaches used.

In the multipotent pancreatic progenitor cells, *Neurog3* expression is repressed through active NOTCH signalling [54, 55]. Endocrine cell differentiation requires loss of NOTCH signalling and binding of a number of transcription factors to the *Neurog3* promoter including: one cut homeobox 1 (ONECUT1), SOX9, hepatocyte nuclear factor 1β (HNF1B), forkhead box protein A2 (FOXA2) and NEUROG3 itself [15, 42, 54, 56–58]. This study adds to the cadre of factors that regulate NEUROG3 by demonstrating that SOX4 participates in this NEUROG3 regulatory network.

SOX4 remains important for endocrine differentiation downstream of its initial role in activating NEUROG3 protein expression. *Pax4* and *Neurod1*, which lie directly downstream of *Neurog3* and are essential during beta cell formation [24, 59], were significantly reduced in the ES4KO pancreas. Null mutation of *Neurod1* leads to reductions in alpha, beta and delta cells [20] while pancreatic *Pax4*-knockouts have reductions of both beta and delta cells and increases in the other endocrine cells including $GHRL^+$ epsilon cells [24, 60, 61]. Notably, a dramatic increase in epsilon cells was not observed following loss of SOX4, suggesting that residual *Pax4* expression is sufficient to suppress epsilon cell formation. Interestingly, the highly abundant epsilon cells in the *Pax4*-null pancreas also express CHGA [60]; thus, the possibility remains that loss of SOX4 drives endocrine differentiation to an unappreciated terminally differentiated non-islet endocrine cell type.

A number of studies have characterised transcription factors and their regulatory networks in pancreatic development. This knowledge has catalysed advancements of directed differentiation protocols and has brought cell-based strategies for the generation of functional INS-producing cells to clinical trials. In conclusion, this work expands the network of transcription factors that are directly important for beta cell formation, and future studies that aim to understand how SOX4 is induced within endocrine progenitors may refine protocols for differentiation of beta cells from hESCs.

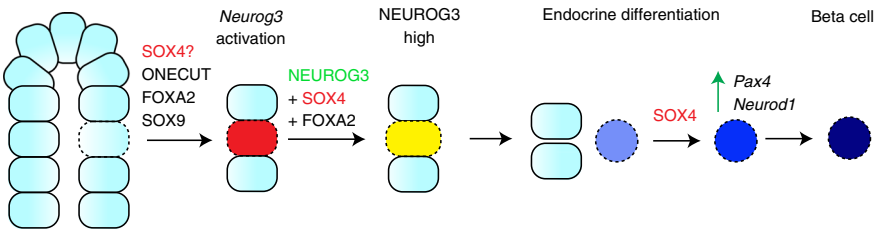

Fig. 8 Model highlighting a role for SOX4 during pancreas development. SOX4-expressing bipotent progenitor cells (red) activate high levels of NEUROG3 expression (yellow). SOX4 and NEUROG3 cooperate following *Neurog3* activation to increase NEUROG3 levels and also to increase expression of downstream targets *Pax4* and *Neurod1*, genes necessary for mature endocrine cell formation (light and dark blue)

Acknowledgements We thank A. Grapin-Botton (University of Copenhagen, Copenhagen, Denmark), M. S. German (University of California San Francisco, San Francisco, CA, USA), S. B. Smith (University of California San Francisco, San Francisco, CA, USA), L. Sussel (Columbia University, New York, NY, USA), D. Melton (Harvard University, Cambridge, MA, USA) and D. Hanahan (École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland) for their generous contributions of reagents that made this work possible. We thank all members of the Lynn laboratory as well as R. Gasa (Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain) and T. Miyatsuka for their thoughtful reading of the manuscript.

Duality of interest The authors declare that there is no duality of interest associated with this manuscript.

Funding This work was supported by grants to FCL from the JDRF (2-2011-91) and the Canadian Institutes of Health Research (CIHR) (MOP 102628). The Michael Smith Foundation for Health Research, the Canadian Diabetes Association and the Child and Family Research Institute (CFRI) provided salary support for FCL. EEX and NAJK were recipients of CIHR-BC Transplantation Trainee Fellowships and EEX was a recipient of a CFRI graduate studentship.

Contribution statement EEX, NAJK, ST, SZC, MT, CN, and FCL designed, acquired and analysed data. FCL and EEX drafted the manuscript and all authors subsequently revised the manuscript for important intellectual content. All authors approve the version of the manuscript to be published. FCL is the guarantor of the integrity of this work.

References

1. Seymour PA, Freude KK, Tran MN et al (2007) SOX9 is required for maintenance of the pancreatic progenitor cell pool. Proc Natl Acad Sci U S A 104:1865–1870
2. Kawaguchi Y, Cooper B, Gannon M et al (2002) The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. Nat Genet 32:128–134
3. Jonsson J, Carlsson L, Edlund T, Edlund H (1994) Insulin-promoter-factor 1 is required for pancreas development in mice. Nature 371: 606–609
4. Lammert E, Cleaver O, Melton D (2001) Induction of pancreatic differentiation by signals from blood vessels. Science 294:564–567
5. Kim SK, Melton DA (1998) Pancreas development is promoted by cyclopamine, a hedgehog signaling inhibitor. Proc Natl Acad Sci U S A 95:13036–13041
6. Deutsch G, Jung J, Zheng M et al (2001) A bipotential precursor population for pancreas and liver within the embryonic endoderm. Development 128:871–881
7. Hebrok M, Kim SK, St Jacques B et al (2000) Regulation of pancreas development by hedgehog signaling. Development 127:4905–4913
8. Kim SK, Hebrok M, Li E et al (2000) Activin receptor patterning of foregut organogenesis. Genes Dev 14:1866–1871
9. Pictet RL, Clark WR, Williams RH, Rutter WJ (1972) An ultrastructural analysis of the developing embryonic pancreas. Dev Biol 29: 436–467
10. Shih HP, Wang A, Sander M (2013) Pancreas organogenesis: from lineage determination to morphogenesis. Annu Rev Cell Dev Biol 29:81–105
11. Gouzi M, Kim YH, Katsumoto K et al (2011) Neurogenin3 initiates stepwise delamination of differentiating endocrine cells during pancreas development. Dev Dyn 240:589–604
12. Gradwohl G, Dierich A, LeMeur M, Guillemot F (2000) Neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. Proc Natl Acad Sci U S A 97:1607–1611
13. Zhou Q, Law AC, Rajagopal J et al (2007) A multipotent progenitor domain guides pancreatic organogenesis. Dev Cell 13:103–114
14. Kopp JL, Dubois CL, Hao E et al (2011) Progenitor cell domains in the developing and adult pancreas. Cell Cycle 10:1921–1927
15. Jensen J, Pedersen EE, Galante P et al (2000) Control of endodermal endocrine development by Hes-1. Nat Genet 24:36–44
16. Murtaugh LC, Stanger BZ, Kwan KM, Melton DA (2003) Notch signaling controls multiple steps of pancreatic differentiation. Proc Natl Acad Sci U S A 100:14920–14925
17. Gu G, Dubauskaite J, Melton DA (2002) Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. Development 129:2447–2457
18. Miyatsuka T, Li Z, German MS (2009) Chronology of islet differentiation revealed by temporal cell labeling. Diabetes 58:1863–1868
19. Ahlgren U, Pfaff SL, Jessell TM et al (1997) Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. Nature 385:257–260
20. Naya FJ, Huang HP, Qiu Y et al (1997) Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice. Genes Dev 11:2323–2334
21. Sander M, Neubüser A, Kalamaras J et al (1997) Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. Genes Dev 11:1662–1673
22. Smith SB, Qu H-Q, Taleb N et al (2010) Rfx6 directs islet formation and insulin production in mice and humans. Nature 463:775–780
23. Sussel L, Kalamaras J, Hartigan-O'Connor DJ et al (1998) Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells. Development 125:2213–2221
24. Sosa-Pineda B, Chowdhury K, Torres M et al (1997) The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas. Nature 386:399–402
25. Matsuoka T, Zhao L, Artner I et al (2003) Members of the large Maf transcription family regulate insulin gene transcription in islet beta cells. Mol Cell Biol 23:6049–6062
26. Wegner M (2010) All purpose Sox: the many roles of Sox proteins in gene expression. Int J Biochem Cell Biol 42:381–390
27. Kamachi Y, Kondoh H (2013) Sox proteins: regulators of cell fate specification and differentiation. Development 140:4129–4144
28. Lioubinski O, Müller M, Wegner M, Sander M (2003) Expression of Sox transcription factors in the developing mouse pancreas. Dev Dyn 227:402–408
29. Mavropoulos A, Devos N, Biemar F et al (2005) Sox4b is a key player of pancreatic alpha cell differentiation in zebrafish. Dev Biol 285:211–223
30. McDonald E, Li J, Krishnamurthy M et al (2012) SOX9 regulates endocrine cell differentiation during human fetal pancreas development. Int J Biochem Cell Biol 44:72–83
31. Wilson ME, Yang KY, Kalousova A et al (2005) The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas. Diabetes 54:3402–3409
32. Muzumdar MD, Tasic B, Miyamichi K et al (2007) A global double-fluorescent Cre reporter mouse. Genesis 45:593–605
33. Schonhoff SE, Giel-Moloney M, Leiter AB (2004) Neurogenin 3-expressing progenitor cells in the gastrointestinal tract differentiate into both endocrine and non-endocrine cell types. Dev Biol 270:443–454
34. Penzo-Méndez A, Dy P, Pallavi B, Lefebvre V (2007) Generation of mice harboring a Sox4 conditional null allele. 780:776–780
35. Sabatini PV, Krentz NAJ, Zarrouki B et al (2013) Npas4 is a novel activity-regulated cytoprotective factor in pancreatic β-cells. Diabetes 62:2808–2820

36. Carpenter AE, Jones TR, Lamprecht MR et al (2006) Cell profiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7:R100
37. Luo J, Deng Z-L, Luo X et al (2007) A protocol for rapid generation of recombinant adenoviruses using the AdEasy system. Nat Protoc 2: 1236–1247
38. Seymour PA, Freude KK, Dubois CL et al (2008) A dosage-dependent requirement for Sox9 in pancreatic endocrine cell formation. Dev Biol 323:19–30
39. Schaffer AE, Freude KK, Nelson SB, Sander M (2010) Nkx6 transcription factors and Ptf1a function as antagonistic lineage determinants in multipotent pancreatic progenitors. Dev Cell 18:1022–1029
40. Yoshida T, Hanahan D (1994) Murine pancreatic ductal adenocarcinoma produced by in vitro transduction of polyoma middle T oncogene into the islets of Langerhans. Am J Pathol 145:671–684
41. Gasa R, Mrejen C, Leachman N et al (2004) Proendocrine genes coordinate the pancreatic islet differentiation program in vitro. Proc Natl Acad Sci U S A 101:13245–13250
42. Lynn FC, Smith SB, Wilson ME et al (2007) Sox9 coordinates a transcriptional network in pancreatic progenitor cells. Proc Natl Acad Sci U S A 104:10500–10505
43. Lee JC, Smith SB, Watada H et al (2001) Regulation of the pancreatic pro-endocrine gene neurogenin3. Diabetes 50:928–936
44. Rukstalis JM, Habener JF (2009) Neurogenin3: a master regulator of pancreatic islet differentiation and regeneration. Islets 1:177–184
45. Bhattaram P, Penzo-Méndez A, Sock E et al (2010) Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. Nat Commun 1:9
46. Talchai C, Xuan S, Kitamura T et al (2012) Generation of functional insulin-producing cells in the gut by Foxo1 ablation. Nat Genet 44: 406–412
47. Anderson KR, Torres CA, Solomon K et al (2009) Cooperative transcriptional regulation of the essential pancreatic islet gene NeuroD1 (beta2) by Nkx2.2 and neurogenin 3. J Biol Chem 284:31236–31248
48. Xiao X, Chen Z, Shiota C et al (2013) No evidence for β cell neogenesis in murine adult pancreas. J Clin Invest 123:2207–2217
49. Boyer LA, Lee TI, Cole MF et al (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122:947–956
50. Aksoy I, Jauch R, Chen J et al (2013) Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. EMBO J 32:938–953
51. Choi E, Kraus MRC, Lemaire LA et al (2012) Dual lineage-specific expression of Sox17 during mouse embryogenesis. Stem Cells 30: 2297–2308
52. Lynn FC, Skewes-Cox P, Kosaka Y et al (2007) MicroRNA expression is required for pancreatic islet cell genesis in the mouse. Diabetes 56:2938–2945
53. Goldsworthy M, Hugill A, Freeman H, Horner E (2008) Role of the transcription factor sox4 in insulin secretion and impaired glucose tolerance. Diabetes 57:2234–2244
54. Apelqvist A, Li H, Sommer L et al (1999) Notch signalling controls pancreatic cell differentiation. Nature 400:877–881
55. Qu X, Afelik S, Jensen JN et al (2013) Notch-mediated post-translational control of Ngn3 protein stability regulates pancreatic patterning and cell fate commitment. Dev Biol 1–12
56. Schwitzgebel VM, Scheel DW, Conners JR et al (2000) Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. Development 127:3533–3542
57. Wang S, Jensen JN, Seymour PA et al (2009) Sustained Neurog3 expression in hormone-expressing islet cells is required for endocrine maturation and function. Proc Natl Acad Sci U S A 106:9715–9720
58. Lee C, Sund N, Vatamaniuk M (2002) Foxa2 controls Pdx1 gene expression in pancreatic beta cells in vivo. Diabetes 51:2546–2551
59. Smith SB, Gasa R, Watada H et al (2003) Neurogenin3 and hepatic nuclear factor 1 cooperate in activating pancreatic expression of Pax4. J Biol Chem 278:38254–38259
60. Prado CL, Pugh-Bernard AE, Elghazi L et al (2004) Ghrelin cells replace insulin-producing beta cells in two mouse models of pancreas development. Proc Natl Acad Sci U S A 101:2924–2929
61. Wang Q, Elghazi L, Martin S et al (2008) Ghrelin is a novel target of Pax4 in endocrine progenitors of the pancreas and duodenum. Dev Dyn 237:51–61