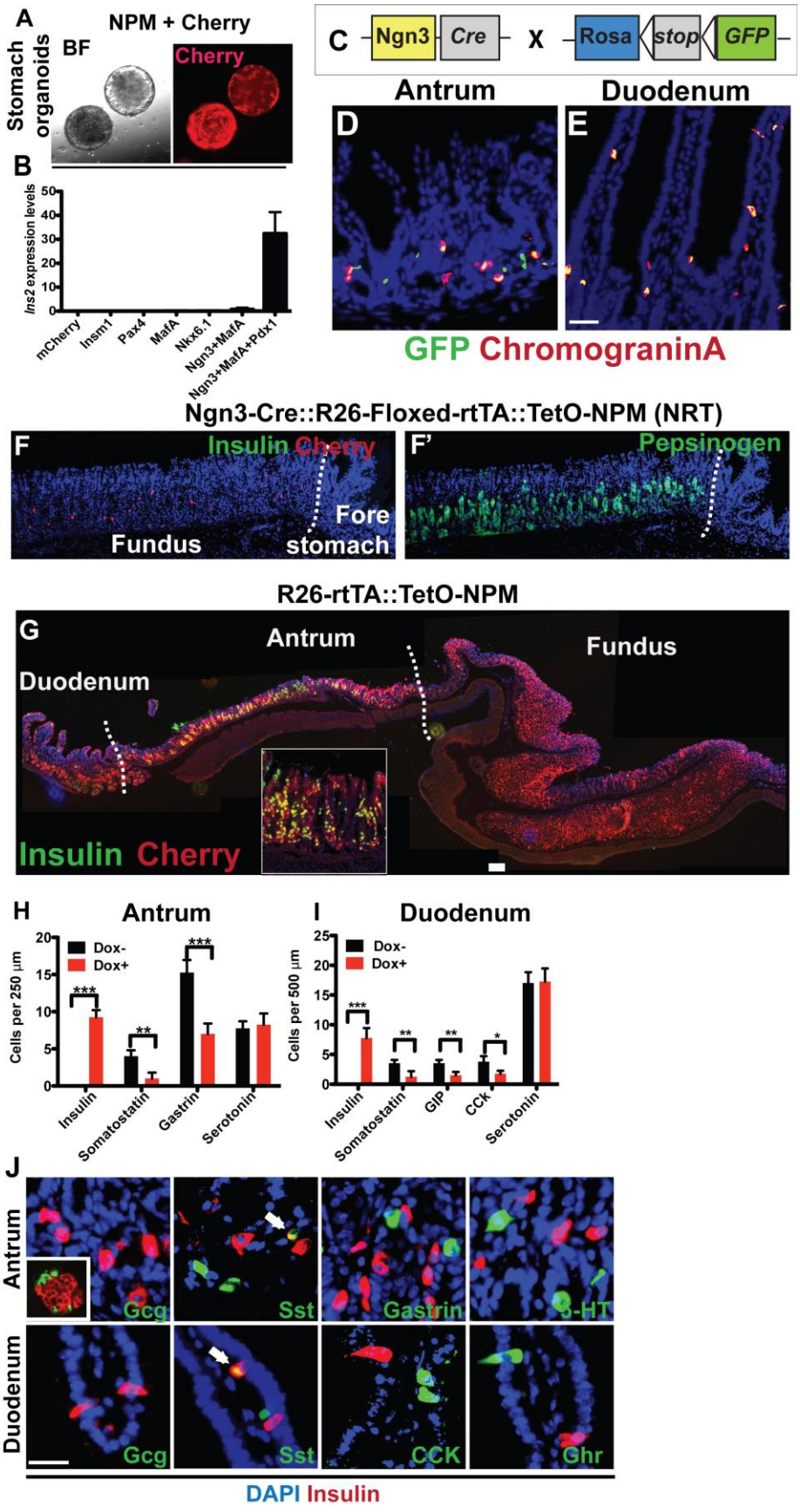

A
NPM + Cherry
Stomach organoids
BF
Cherry
B
Ins2 expression levels
mCherry
Insm1
Pax4
MafA
Nkx6.1
Ngn3+MafA
Ngn3+MafA+Pdx1
C
Ngn3
Cre
X
Rosa
stop
GFP
Antrum
Duodenum
D
E
GFP ChromograninA
Ngn3-Cre::R26-Floxed-rtTA::TetO-NPM (NRT)
F
Insulin Cherry
Fundus
Fore stomach
F'
Pepsinogen
R26-rtTA::TetO-NPM
G
Antrum
Fundus
Duodenum
Insulin Cherry
H
Antrum
Cells per 250 µm
Dox-
Dox+
Insulin
Somatostatin
Gastrin
Serotonin
I
Duodenum
Cells per 500 µm
Dox-
Dox+
Insulin
Somatostatin
GIP
CCk
Serotonin
J
Antrum
Gcg
Sst
Gastrin
5-HT
Duodenum
Gcg
Sst
CCK
Ghr
DAPI Insulin

Supplementary Figure 1 (related to Figure 1). Induction of insulin$^+$ cells from the gastrointestinal tissues by NPM factors (Ngn3, Pdx1 and MafA).

A, B. To evaluate whether gastric tissues can be converted to insulin$^+$ cells by the NPM factors, we cultured gastric organoids from the antral stomach of wild-type animals and transduced with adenoviruses expressing NPM factors along with Cherry. qPCR analysis reveals the highest induction of *Ins2* transcripts from gastric organoids by NPM factors, but not by other putative beta-cell reprogramming factors. Data presented as mean ± s.d, n=3.

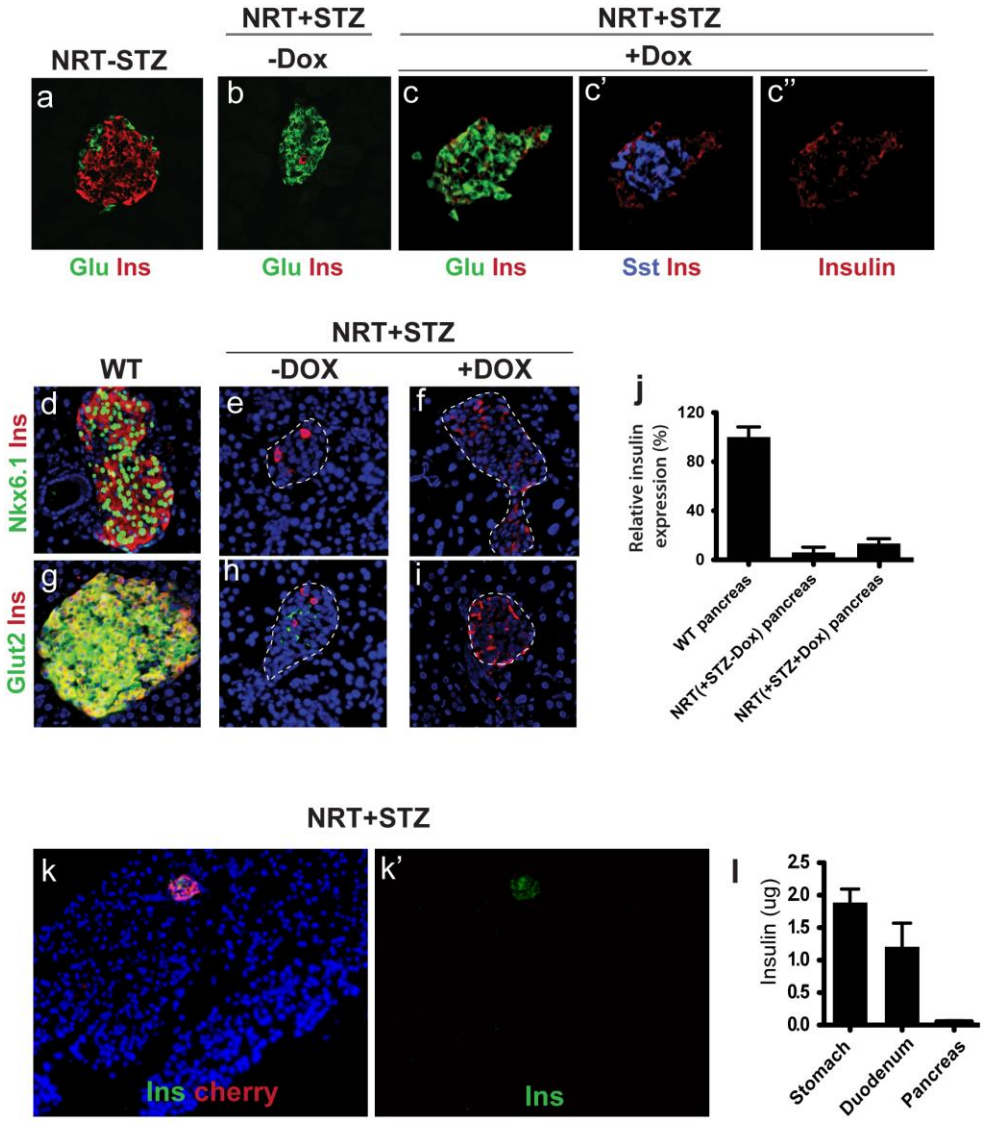
C-E. To enable expression of NPM factors in the enteroendocrine lineage of the GI tract, we used the well-described Ngn3-Cre mouse line, which labels all enteroendocrine cells in the intestine and the majority of enteroendocrine cells in the antral stomach. After crossing with the Rosa-GFP reporter line (**C**), we confirmed that co-expression of GFP$^+$ cells in antral stomach and duodenum with Chromogranin, a marker of mature enteroendocrine cells (**D, E**). Some immature GFP$^+$Chromogranin$^-$ enteroendocrine cells were also present (**D, E**). Scale bar: 50um.

F. In triple transgenic NRT animals, a small number of cherry$^+$ cells were observed in the fundus (corpus) stomach, few of the fundal cherry$^+$ are insulin$^+$. Pepsinogen staining (**F**') was used to confirm the fundal identity.

G. Rosa-rtTA;TetO-NPM double transgenic adult mice were treated with Dox for 7 days and analyzed with immunohistochemistry. Dox treatment longer than 10 days led to animal death due to hypoglycemia. Many insulin$^+$ cells appeared in the antrum whereas very few insulin$^+$ cells were induced from the fundus, indicating that fundal tissue is not amenable for NPM-mediated conversion. Inset: A higher magnification view of the induced antral insulin$^+$ cells. Scale bar: 1mm.

H, I. Quantitation of major endocrine subtypes in antral stomach and duodenum before and after induction of insulin$^+$ cells in NRT animals. Significant reduction of Somatostatin$^+$ and Gastrin$^+$ cells was observed in the antrum (**H**). In duodenum, somatostatin$^+$, GIP$^+$, and CCK$^+$ cell populations all showed significant reduction after Dox treatment and insulin induction (**I**). In contrast, the number of serotonin$^+$ cells in both antrum and duodenum remains unchanged (**H, I**). Data presented as mean ± s.d, n=4 animals. * = $p<0.05$, ** = $p<0.01$, and *** = $p<0.001$, Student's t-test.

J. Immunohistochemistry in the antrum and the duodenum showed that the vast majority of induced insulin$^+$ cells are mono-hormonal. Gcg: glucagon. Sst: somatostatin. 5-HT: serotonin. CCK: cholecystokinin. Ghr: Ghrelin. Arrows indicate a small number of insulin$^+$Sst$^+$ cells in antrum and duodenum (1.1 ± 0.8% of total insulin$^+$ cells, mean ± s.d, n=3 animals). Scale bar, 50 μm.



Supplementary Figure 2 (related to Figure 2). Low level of insulin expression was induced from Glucagon$^+$ cells in the islet remnants of STZ-treated NRT animals.

A. Immunohistochemistry shows that before STZ-mediated ablation, the islets are composed largely of insulin$^+$ cells and a minority of glucagon$^+$ cells.

B. After STZ treatment, the endogenous insulin$^+$ cells disappeared, the islet remnants are now composed largely of glucagon$^+$ cells and some somatostatin$^+$ cells.

C. After NPM induction in islet cells in NRT animals, we observed insulin expression, the majority of which is in glucagon$^+$ cells (**c**). Some somatostatin$^+$ cells also express insulin (**c'**). The insulin channel is shown separately in **c''**.

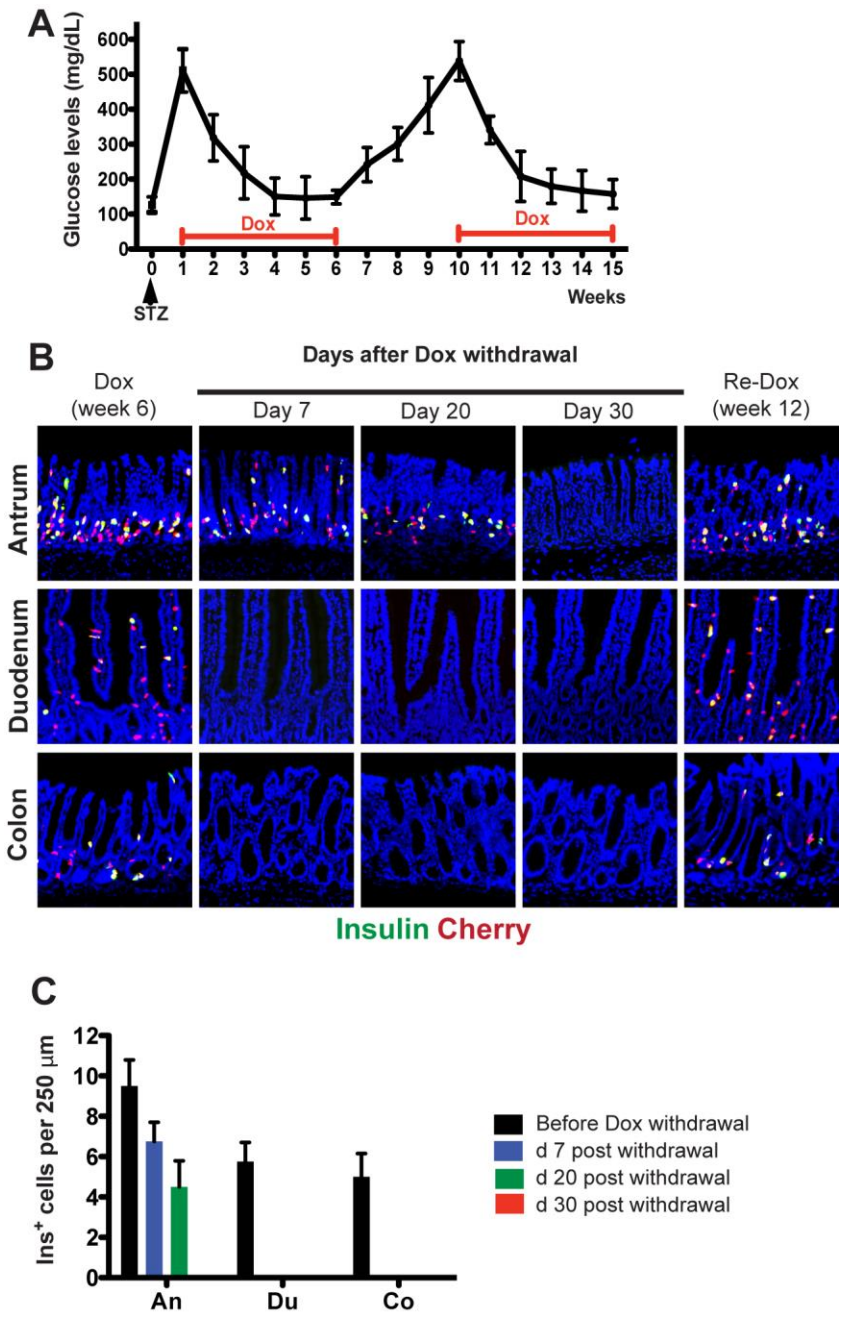
D-I. Consistent with the interpretation that the glucagon$^+$insulin$^+$ cells in the NRT animals after Dox induction represent insulin up-regulation in glucagon$^+$ cells, rather than an conversion towards beta cells, the glucagon$^+$insulin$^+$ cells do not express other beta cell markers Nkx6.1 and Glut2.

J. Compared with endogenous beta cells, the glucagon$^+$Insulin$^+$ cells express insulin at significantly lower levels. Data presented as mean ± s.d, n=3 animals.

K. We did not detect insulin$^+$ expression outside of the islet remnants in the pancreas of NRT animals that have undergone 80% pancreatectomy.

I. There is little insulin content in the pancreatectomized pancreas. In contrast, the stomach and the duodenum contain substantial amount of insulin. Data presented as mean ± s.d, n=3 animals.

To facilitate direct comparison, all pictures presented were taken at the same exposure time and at non-saturating pixel intensity**.**



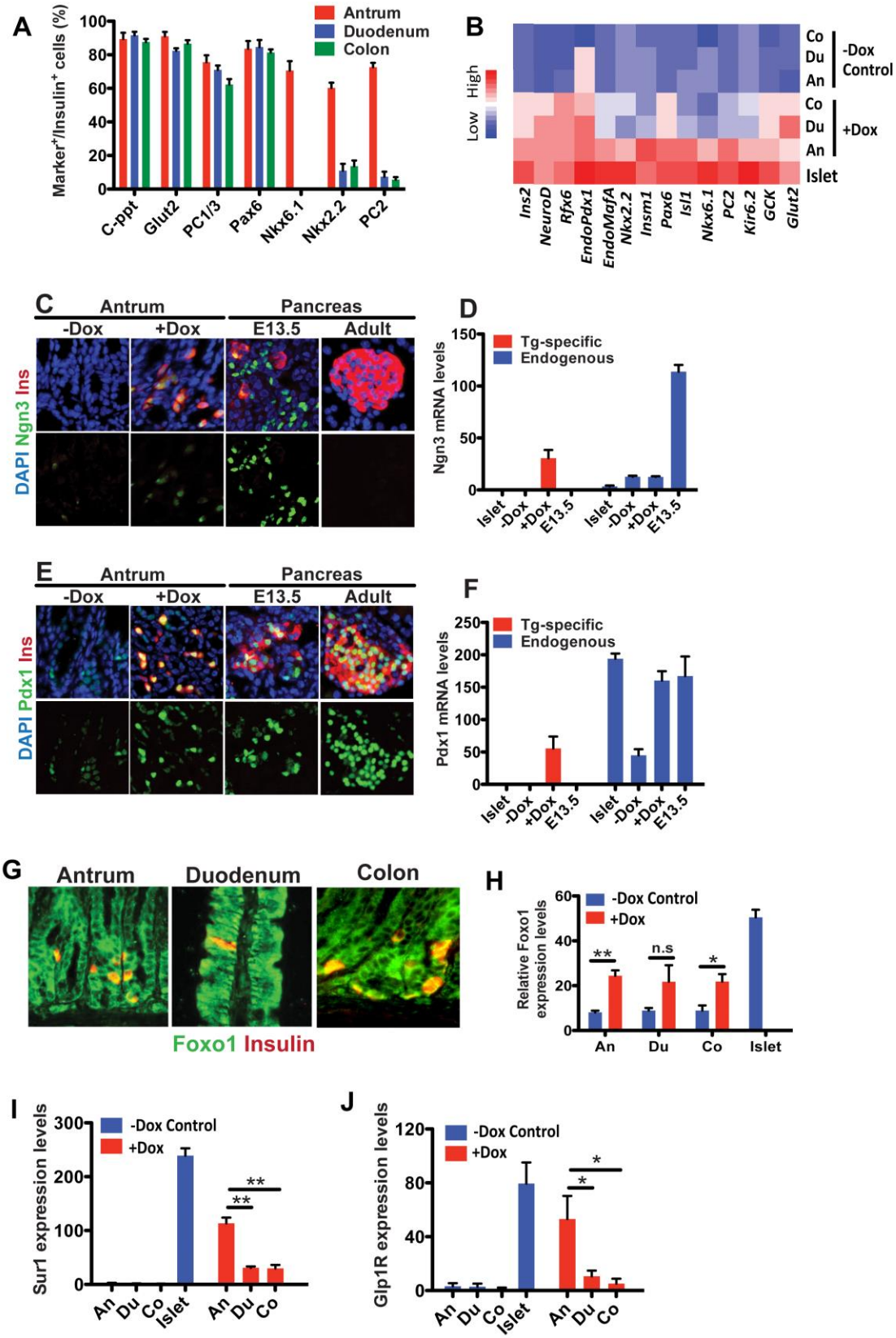
Supplementary Figure 3 (related to Figure 2). Antral insulin+ cells have longer lifespan in vivo compared with intestinal insulin+ cells and can suppress hyperglycemia.

A. In a pulse-chase experiment, we first used STZ to ablate pancreatic beta-cells in NRT animals, leading to hyperglycemia. Subsequent Dox treatment led to formation of GI insulin^{+} cells and normoglycemia in 5 weeks. After Dox withdrawal, the animals

progressively returned to hyperglycemia in a span of 4 weeks (Week 6-10). Re-administration of Dox normalized the hyperglycemia again.

B. Immunohistology showed an abundance of insulin$^+$ cells in the antrum, duodenum, and colon at the end of the first Dox period (first panel on the left). 7 days after Dox withdrawal, intestine insulin$^+$ cells became undetectable whereas plenty of antral insulin$^+$ cells were still present (second panel). Antral insulin$^+$ cells were detectable on Day 20 (third panel) but disappeared at Day 30 (fourth panel). Re-administration of Dox re-induced insulin$^+$ cells from all GI regions (last panel). Given that antral insulin$^+$ cells persisted for at least 2 weeks after the disappearance of intestinal insulin$^+$ cells, suppression of hyperglycemia during weeks 7-9 in **A** reflect suppression of hyperglycemia by antral cells alone.

C. quantitation of insulin$^+$ cells from different GI regions in the pulse-chase experiment. Data presented as mean ± s.d, n=3 animals.


A
Marker+/Insulin+ cells (%)
Antrum
Duodenum
Colon
C-ppt
Glut2
PC1/3
Pax6
Nkx6.1
Nkx2.2
PC2
B
High
Low
Co
Du
An
-Dox Control
+Dox
Islet
Ins2
NeuroD
Rfx6
EndoPdx1
EndoMafA
Nkx2.2
Insm1
Pax6
Isl1
Nkx6.1
PC2
Kir6.2
GCK
Glut2
C
Antrum
Pancreas
-Dox
+Dox
E13.5
Adult
DAPI Ngn3 Ins
D
Tg-specific
Endogenous
Ngn3 mRNA levels
Islet
-Dox
+Dox
E13.5
E
DAPI Pdx1 Ins
F
Pdx1 mRNA levels
G
Antrum
Duodenum
Colon
Foxo1 Insulin
H
-Dox Control
+Dox
Relative Foxo1 expression levels
n.s
An
Du
Co
Islet
I
Sur1 expression levels
J
Glp1R expression levels

Supplementary Figure 4 (related to Figure 3). Characterization of insulin$^+$ cells from the antral stomach, the duodenum, and the colon.

A. Quantitation of immunohistochemistry staining data showed that the majority of induced insulin$^+$ cells from the antrum and duodenum express c-peptide, glucose transporter 2 (Glut2), Prohormone convertase 1/3 (PC1/3), and Pax6. However, the expression of Nkx6.1, Nkx2.2 and PC2 are largely restricted to antrum insulin$^+$ cells. Data presented as mean ± s.d, n=3 animals. A total of at least 1,000 insulin$^+$ cells were counted.

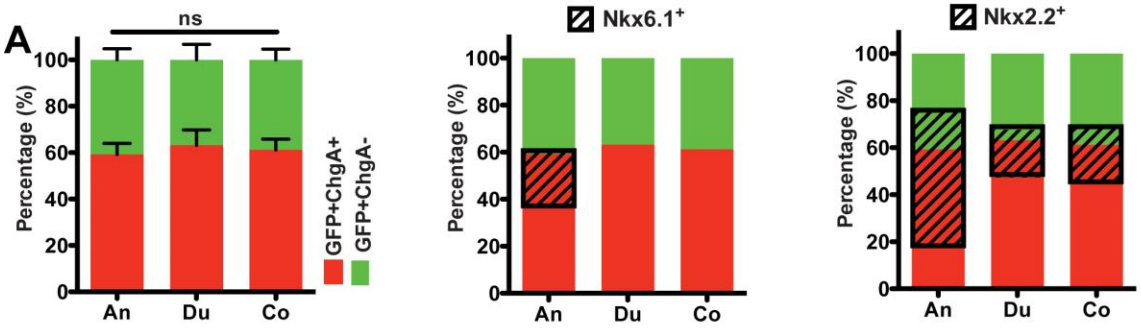
B. Heatmap representation of beta-cell genes in Colon (Co), duodenum (Du), and antrum (An) in control and Dox-treated samples. qPCR data were generated from control samples (total epithelial population from NRT animals without Dox treatment) and FACS-purified Cherry+ cells isolated from NRT animals 10 days after Dox treatment. The beta-cell factors are robustly induced in antral tissue. In contrast, many beta-cell genes are not strongly induced or not induced in the duodenal and colonic tissues. Islet: purified pancreatic islets.

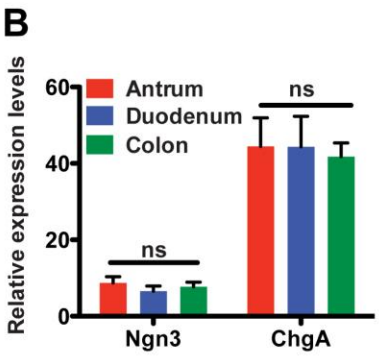
C, D. Ngn3 protein is expressed transiently in a small number of endocrine progenitors of the Antrum (**C**, first panel). After Dox induction in NRT animals, Ngn3 is expressed at a relatively low level in the induced insulin$^+$ cells (**C**, second panel). In contrast, Ngn3 expression in the pancreatic endocrine progenitors during embryogenesis is much higher (**C**, third panel). No Ngn3 expression is detectable in adult islet cells by immunohistochemistry (**C**, last panel). qPCR analysis indicates that NPM factors do not activate endogenous Ngn3 expression (**D**). Data presented as mean ± s.d, n=3 animals.

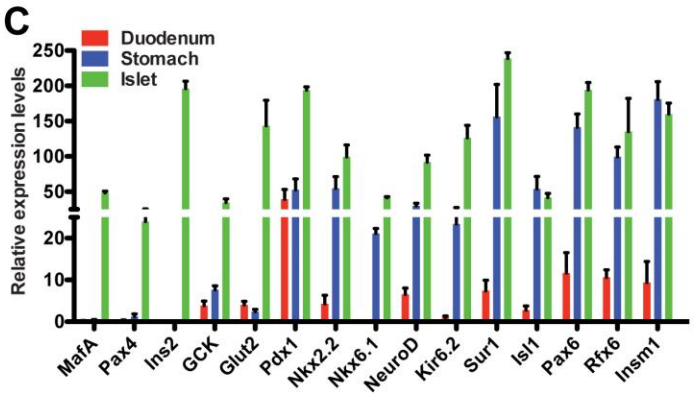
E, F. Pdx1 is expressed at a low level in antral epithelial cells (**E**, first panel). Strong Pdx1 expression was observed in the induced insulin$^+$ cells (**E**, second panel), comparable to Pdx1 expression in embryonic pancreas (**E**, third panel) and adult islet beta-cells (**E**, last panel). qPCR analysis indicates activation of endogenous Pdx1 expression in the induced insulin+ cells (**F**). Data presented as mean ± s.d, n=3 animals.

G, H. Immunohistochemistry and qPCR data showing expression of Foxo1 in the GI insulin$^+$ cells. qPCR data were generated from control samples (total epithelial population from NRT animals without Dox treatment) and FACS-purified Cherry+ cells isolated from NRT animals 10 days after Dox treatment. Islet: purified pancreatic islets. Data presented as mean ± s.d, n=3 animals. Student t-test.

I, J. qPCR analysis of FACS-purified Cherry$^+$ cells from Dox-treated NRT animals showed that Sur1 and Glp1R expression is substantially higher in antral insulin$^+$ cells versus duodenal and colonic insulin$^+$ cells. Control samples are total epithelial populations from NRT animals without Dox treatment. Islet: purified pancreatic islets. Student t-test. Data presented as mean ± s.d, n=4 animals.



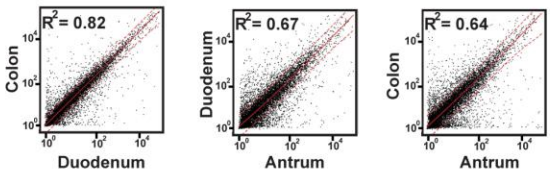




D **Spearman Correlation Coefficient (8563 genes)**

	Beta	Colon	Antrum	Duodenum
Beta	1	0.57	0.72	0.57
Colon		1	0.80	0.91
Antrum			1	0.82
Duodenum				1

E **Scatter plots of 8563 genes**



F

GeneOntology analysis

Group1

	P values
Organophosphate metabolic process	2.51E-08
Protein transport	1.11E-07
Protein catabolic process	4.29E-05
Modification-dependent protein catabolic process	4.39E-05
Proteolysis involved in cellular protein catabolic process	5.39E-05
Modification-dependent macromolecule catabolic process	6.21E-05
Single-organism biosynthetic process	1.01E-04
Cellular protein catabolic process	1.24E-04
Nucleobase-containing small molecule metabolic process	2.25E-04
Macromolecule catabolic process	2.57E-04
Ubiquitin-dependent protein catabolic process	2.76E-04
Organonitrogen compound metabolic process	3.25E-04
Cellular macromolecule catabolic process	4.53E-04
Carboxylic acid metabolic process	5.96E-04
Organophosphate biosynthetic process	8.54E-04

Group2

	P value
regulation of secretion by cell (GO:1903530)	8.8E-03
regulation of G-protein coupled receptor signaling (GO:0008277)	9.8E-03
regulation of hormone secretion (GO:0046883)	1.5E-02
regulation of peptide hormone secretion (GO:0090276)	2.0E-02
regulation of secretion (GO:0051046)	2.3E-02
regulation of peptide secretion (GO:0002791)	2.4E-02
pancreatic A cell differentiation (GO:0003310)	3.0E-02
cell-cell signaling (GO:0007267)	3.6E-02

Group3

Vesicle-mediated transport	1.07E-11
Intracellular transport	2.13E-08
Cellular protein localization	5.95E-06
Cellular macromolecule localization	8.29E-06
Single-organism intracellular transport	1.20E-05
Protein transport	2.73E-05
Single-organism carbohydrate metabolic process	2.83E-05
Intracellular protein transport	9.66E-05
Golgi vesicle transport	1.34E-04
Regulation of cellular localization	1.86E-04
Carbohydrate metabolic process	3.10E-04
Glycoprotein biosynthetic process	6.99E-04

Group4

None

Group5

Regulation of secretion by cell	1.93E-04
Carbohydrate homeostasis	4.30E-04
Glucose homeostasis	4.30E-04
Regulation of secretion	6.78E-04

Supplementary Figure 5 (related to Figure 4). Characterization of purified GI enteroendocrine populations

A. Quantitation of GFP^+ cells from the GI tract of the Ngn3-GFP reporter mice (bar graph on left). The percentile of chromogranin$^+$ ($Chga^+$, red bar) and $Chga^-$ (green bar) populations are comparable in antrum, duodenum, and colon. The bar graphs in the middle and the right represent distribution of $Nkx6.1^+$ and $Nkx2.2^+$ cells in the GFP^+Chga^+ and GFP^+Chga^- populations. Statistical analysis by student t-test. Data presented as mean ± s.d, n=3 animals.

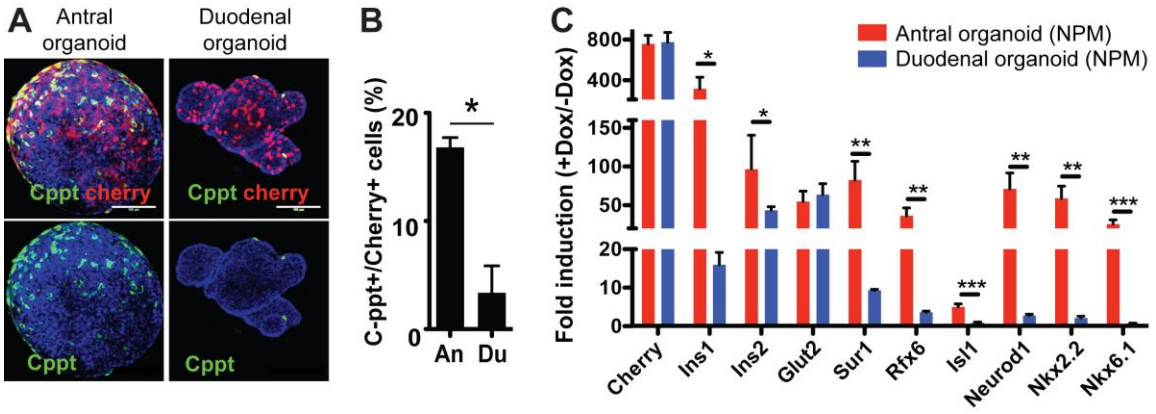
B. qPCR analysis of FACS-purified GFP^+ cells from the Ngn3-GFP mice showed similar expression levels of the Ngn3 and Chga genes in Antrum, duodenum, and colon. Data presented as mean ± s.d, n=3 animals. Statistical analysis by student t-test.

C. Expression of beta-cell factors in the Ngn3-GFP^+ enteroendocrine populations from the antral stomach, the duodenum, and comparison with pancreatic beta-cells. Data presented as mean ± s.d, n=3 animals.

D. Spearman correlation coefficient indicates that endocrine cells from duodenum and colon are closely related to each other (0.91), whereas they are more dissimilar from antrum endocrine cells (0.82 and 0.80). Among the 3 GI endocrine populations, antrum cells share greater transcriptome similarity with pancreatic beta cells (0.72) compared with duodenum and colon endocrine cells (0.57 and 0.57). Spearman coefficient calculated with 8563 genes after excluding commonly expressed housekeeping genes (see Methods for details).

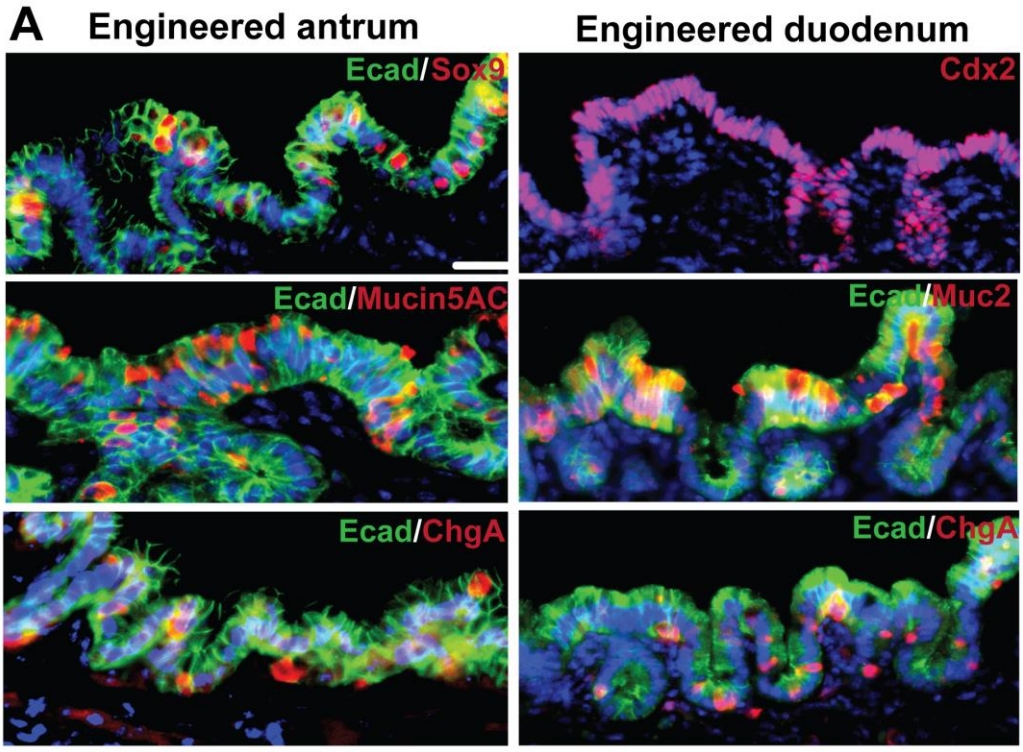
E. Scatter plots of comparison between the three GI endocrine populations.

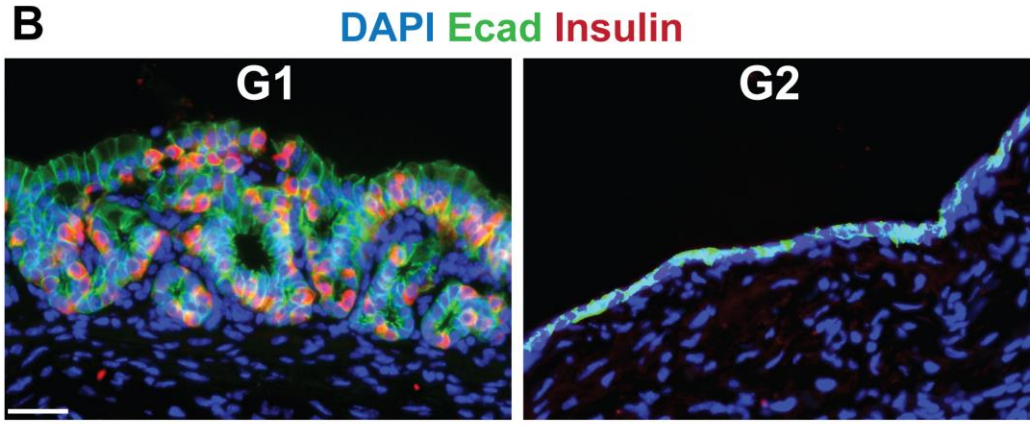
F. Gene ontology analyses of groups of genes shown in Fig. 4D.

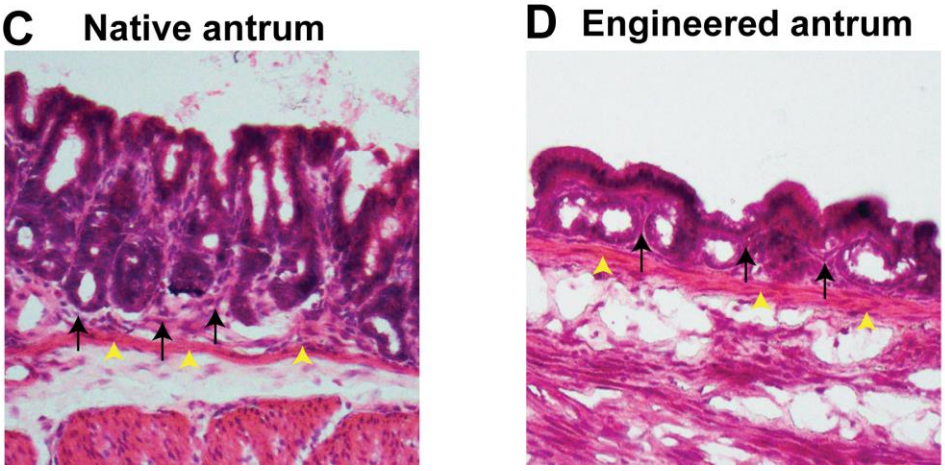


Supplementary Figure 6 (related to Figure 5). Cultured antral epithelial organoids produce more c-peptide^{+} cells than intestinal organoids.

Epithelial organoids were established from the antrum or duodenum of the double transgenic Rosa-rtTA;Teto-NPMcherry (Rosa-NPM) animals. After Dox induction for 7 days, c-peptide staining revealed an abundance of c-peptide^{+} (Cppt^{+}) cells from the antral organoids and relatively few Cppt^{+} cells from the intestinal organoids (**A**). Quantitation showed higher conversion efficiency for antral organoids (**B**). qPCR data indicate that antral cells express higher levels of many beta-cell factors compared with intestinal cells after NPM induction (**C**). Scale bar, 100 μm. Data presented as mean ± s.d. N=3. Student t-test was used to evaluate statistical significance.


A
Engineered antrum
Engineered duodenum
Ecad/Sox9
Cdx2
Ecad/Mucin5AC
Ecad/Muc2
Ecad/ChgA
Ecad/ChgA


B
DAPI Ecad Insulin
G1
G2


C
Native antrum
D
Engineered antrum

Supplementary Figure 7 (related to Figures 6 and 7). Histology and molecular marker analysis of the engineered stomach and intestine mini-organs.

A.. Immunohistochemistry of an engineered stomach construct (left panel) and an engineered intestine construct (right panel) reveals the presence of one or multiple layers of Ecadherin^{+} cells that contain Sox9^{+} stem/progenitor cells, Mucin^{+} secretory cells, and Chromogranin^{+} enteroendocrine cells. Cdx2 is expressed in the intestine tissues. Scale bar, 100 μm.

B. Histology comparison between G1 and G2 stomach spheres. Engineered stomach from a G1 animal show robust epithelia reconstitution and abundant insulin^{+} cells (left panel). In contrast, stomach spheres from G2 animals have poor epithelial structure and few or no insulin^{+} cells (right panel). These differences reflect inefficient reconstitution of epithelium in some of the engineered stomach spheres. Scale bar: 50um.

C, D. H&E histology pictures from the native antral stomach and engineered antral stomach. The epithelial layer is thinner in the engineered constructs but the major non-epithelial, or mesenchymal, components, are preserved in the engineered stomach. Black and yellow arrows indicate two types of mesenchymal cells that have distinct location and morphology.

Table1 qPCR primers (related to Figures 1, 4, and 5)

Genes	primer sequences (5' to 3') (forward; reverse)
Insulin1	CCAGCTATAATCAGAGACCA; GTGTAGAAGAAGCCACGCT
Insulin2	TCCGCTACAATCAAAAACCAT; GCTGGGTAGTGGTGGGTCTA
Transgene (NPM)	AGCATGGTGTTCTCAGAC; GTACTGCTCCTCACTGTTC
Glut2	CTCTTCCAACGTGGTCCCTA; GAGCCCCTCGTAGGTTTTTC
GCK	ATAAGCCAGTGGTGGAGTGG; CCAACACAGGTGTGAAGTGG
Sur1	GAATGATGACAGCTGCTCCA; TCTTCTTATGCCCAAACCTCTG
Kir6.2	TTGGAAGGCGTGGTAGAAAC; CCCCATAGAATCTCGTCAGC
Nkx6.1	TCCGAGTCCTGCTTCTTCT; CACGCTTGGCCTATTCTCTG
Nkx2.2	CTTATCCAATCGCTCCACCTT; TCCAGAACCATCGCTACAAG
Pax4	AGAAGCTGAAATGGGAAGCA; GGGGACTGTGCAGAGATGAT
Isl1	GCCTGTAAACCACCATCATGT; GTGCAAGGACAAGAAACGC
Rfx6	CAATAAATGCCTCCACTGTAGC; TCTTACCATATCCAACAGCACA
NeuroD	ACACTCATCTGTCCAGCTTG; AGATCGTCACTATTCAGAACCTTT
Insm1	GGTTTGCTCTGCCTACCAAT; TCACCCAAAACAACCCGTA

Foxo1	CTTCAAGGATAAGGGCGACA; GACAGATTGTGGCGAATTGA
EndoNgn3	ACTGCTGCTTGTCACTGAC; TCTCTGGGGACACTTGGATG
EndoPdx1	TAAGGCCTGGCTTGTAGCTC; AATGCGCACGGGTCCTTGTAG
EndoMafA	TCTTTCTGTGAGCGCGG; TCAGAGTCCGAACCGAGG
Cdx2	CACCATCAGGAGGAAAAGTGA; CTGCGGTTCTGAAACCAAAT
ChgA	CGATCCAGAAAGATGATGGTC; CGGAAGCCTCTGTCTTTCC
GLP1R	CTGCCCAGCAACACCAGT; CAGTCGGCAGCCTAGAGAGT
RPL19	CTCGTTGCCGGAAAAACA; TCATCCAGGTCACCTTCTCA

Tables 2-4
There tables are created to show detailed information of the gene-profiling data from purified Ngn3-GFP^{+} cells and the comparison with pancreatic beta-cells. Numbers in the tables represent Log2 transformed average signal for each gene from the Illumina arrays.

Tables 2 (Related to Figure 4) includes 5 spread sheets that list all Group1, 2, 3, 4, 5 endocrine genes as shown in Figure 4D.

Table 3 (Related to Figure 4) lists 1,470 genes that are differentially expressed between duodenum and colon samples. Comparison was performed with the entire transcriptome.

Table 4 (Related to Figure 4) lists 121 genes that show higher expression in antrum and duodenum tissues and lower expression in colon.

Supplemental Materials and Methods

Mouse strains and doxycycline treatment R26-floxed-rtTA, R26-floxed-GFP, R26-rtTA, CAG-rtTA, Ngn3-Cre, and NSG mouse strains were obtained from the Jackson laboratories. Ngn3-GFP(Gu et al., 2002; Jenny et al., 2002; Lee et al., 2002) and $Cdx2^{fl/fl}$ (Silberg et al., 2002; Verzi et al., 2013) mice have been described previously. The TetO-NPMcherry mouse lines were made by standard pronuclear injection at the Harvard Genomic Modification Facility (Cambridge, MA) using a polycistronic construct that co-expresses Ngn3, Pdx1, Mafa, and the Cherry fluorescent protein. We validated the transgenic lines by crossing the founders with R26-rtTA, applying Doxycycline (2mg/ml in drinking water) for 10 days, and analyzing tissues for Cherry fluorescence. Two independent transgenic lines with robust global Cherry induction were identified and used in our experiments, which yielded similar results.

Antibodies and Immunofluorescence Tissues were processed as previously described(Li et al., 2014b). The following primary antibodies were used: rabbit anti-chromogranin A (Immunostar: 1:200), guinea pig anti-insulin (Dako; 1:200), guinea pig anti-glucagon (Linco; 1:2000), goat anti-somatostatin (Santa Cruz; 1:500), goat anti-gastrin antibody (Santa Cruz; 1:100), rabbit anti-serotonin (Immunostar; 1:2000), goat anti-CCK (Santa Cruz; 1:100), goat anti-ghrelin (Santa Cruz; 1:100), goat anti-Glut2 (Santa Cruz; 1:500), rabbit anti-PC1/3 (Chemicon; 1:500), rabbit anti-PC2 (Santa cruz; 1:200), rabbit anti-Cppt (AbCam; 1:200), anti-Ngn3 (BCBC; 1:1,1000) rabbit anti-Nkx6.1 (BCBC; 1:1,000), mouse anti-Nkx6.1 (DSHB; 1:50), mouse anti-Nkx2.2 (DSHB; 1:20), rabbit anti-Pax6 (Chemicon; 1:500), rat anti-Ecadherin (Zymed; 1:800), rat anti-Pecam (Pharmingen; 1:200), rabbit anti-Foxo1 (Santa Cruz; 1:100), rabbit anti-Ki67 (AbCam; 1:500), goat anti-Sox9 (R&D; 1:500), mouse anti-Cdx2 (Biogenex, 1:200), Rabbit anti-Pepsinogen (Santa cruz; 1:50), mouse anti-Mucin5AC (AbCam: 1:200), rabbit anti-GFP (Life Technologies; 1:500), rabbit anti-RFP (AbCam; 1:500) and rat anti-RFP (Antibodies-online; 1:500). Secondary antibodies were obtained from the Jackson Immunoresearch laboratories. Pictures were taken with a Zeiss LSM 510 META confocal microscope. Hematoxylin and Eosin (H&E) staining was performed with a standard protocol. For quantification of $marker^+$ cells such as $insulin^+$ cells, a total of at least 1000 $Marker^+$ were analyzed from tissues harvested from three different animals. Typically, at least 10 randomly selected sections were counted per animal.

FACS isolation of GFP^+ cells and gene profiling To dissociate tissues into single cells for FACS purification, we modified published protocols (Habib et al., 2012; Reimann et al., 2008; Talchai et al., 2012a). Briefly, duodenum and colon tissues were cut into small pieces, incubated with 10 mM EDTA for 30 min, and mechanically dissociated by pipetting. The released epithelial clusters were digested with dispase for 15 minutes into single cells. For antral stomach, the minced tissues were incubated in 10mM EDTA for 2 hours, followed by pipetting to release the glands. The glands were further incubated with Accutase® for 30 minutes into single cells. GFP^+ cells were isolated by fluorescent activated cell sorting (FACS) with FACSaria (BD Bioscience). Transcriptome data were generated Illumina microarrays. Briefly, RNA was extracted with TriZol reagent (Invitrogent). Biotin-labeled cRNA probes were synthesized with Illumina TotalPrep RNA Amplificatin kit (Ambion).

Gene profiling was performed with Sentrix BeadChip MouseRef-8 v2 Arrays (Illumina) that contain probes for ~19,000 genes. We used *GenomeStudio* to extract raw gene

expression data from the BeadChip arrays. We then used the *lumi* R package(Du et al., 2008) to transform the data using the variance-stabilizing transformation (VST) algorithm and to perform quantile-normalization. We performed pairwise comparisons between genes in each tissue pair using the *limma* R package, using empirical Bayes smoothing of standard errors. We selected 8,563 genes displaying significantly differential expression (p<0.05, fold-change≥1.5) between at least one tissue pair to calculate the Pearson correlation coefficients for each tissue-pair (Fig. S1). We also selected 2,398 genes with significantly higher expression (p<0.05, fold-change≥1.5) in beta cells compared with acinar cells, and we clustered these genes using *K*-means (*K*=5, Euclidean distance function) in *GeneCluster 3.0(Eisen et al., 1998)* (Fig. 1g). The heatmap in Fig. 1g was plotted using *TreeView*. Gene ontology analysis of biological process term enrichment was performed using geneontology.org (Gu et al.) (p<0.05, GO terms with >1,000 genes excluded).

Transcriptional Analysis RNA isolation from primary tissues and cDNA synthesis were performed using Qiagen RNA extraction kit and iScript™ cDNA Synthesis (Biorad). For organoids, cDNA was synthesized using High Capacity cDNA Reverse Transcriptase kit (Applied Biosystems) and pre-amplified using TaqMan PreAmp Master mix (Applied Biosystems). To assess levels of transgene expression and transcriptional program of gut insulin^{+} cells, Cherry^{+} cells from each GI region were FACS-sorted as above described, and RNA isolation was performed by Qiagen RNeasy Micro Kit. Real time PCR was carried out using an Applied Biosystems 7900 Thermal Cycler with TaqMan or SYBR green PCR assays. Expression levels were normalized to the housekeeping ribosomal protein L19 (RPL19) gene. All primers used are listed in Supplemental Table 1.

Glucose stimulated insulin secretion (GSIS) assay and drug treatments Isolation of gastrointestinal epithelial cells was carried out as previously described with slight modifications (Habib et al., 2012; Reimann et al., 2008; Talchai et al., 2012a). Briefly, antrum tissue was cut into small pieces and incubated in 10 mM EDTA for 30 minutes, followed by mechanical dissociation to extract cell clusters. To collect duodenal samples, the duodenum tissue was first cut into small pieces, incubated in 10mM EDTA for 5 minutes, and mechanically stripped to release the villi. Each cell fraction was incubated with Krebs Ringer buffer supplemented with 1.7 mM or 20.2 mM glucose, glibenclamide (10 nM; Tocris), diazoxide (0.5 mM; Sigma), or exendin-4 (100 nM; Sigma). Released insulin amounts were determined by ELISA (Alpco) at the Joslin Specialized Assay core (Joslin Diabetes Center, Boston). Data from different tissue sources were standardized according to basal insulin release level. Each standardized sample is equivalent to 1/15 of antrum or 5mm of duodenum, or 10mm of colon in a single animal.

Physiological Studies Diabetic animals were produced with intraperitoneal injection of streptozotocin (STZ; 150-170mg/kg) in 6- to 8-week-old animals. Animals that displayed >400 mg dl^{-1} blood glucose levels for 2 consecutive days after STZ administration were used for experiments. Blood glucose was measured with an Ascensia Elite glucometer (Bayer). For blood glucose monitoring, a short 2-hour fasting precedes glucose measurements. Glucose tolerance test was performed with intraperitoneal injection of 1mg/g body weight of glucose, preceded by 4hr fasting. Blood insulin was collected from tail vein blood sampling and measured by ELISA. Tissue insulin was extracted with acid-ethanol solutions and measured by ELISA.

Organoid culture and adenoviral infection Antral stomach and duodenal organoids were derived from young adult mice (1-2 months) and cultured using standard growth media, essentially as described(Barker et al., 2010). To test beta-cell reprogramming factors, 4-7 day cultures of antral or duodenal organoids were recovered from the Matrigel with the recover solution (Corning), infected with purified adenovirus at 4x 10^7 pfu in 100ul medium for 1 hour at 37 °C and re-embedded in Matrigel. qPCR analysis was performed 6-8 days after infection.

Generation of bioengineered stomach and intestine Generation of bioengineered stomach and stomach spheres was performed essentially as described(Maemura et al., 2004; Speer et al., 2011). Briefly, antrum and duodenal tissues were minced and digested by a mixture of collagenase and dispase for 20 minutes at room temperature. Disk-shaped scaffolds (6mm in diameter, 3mm thick) were formed from an 80% porous, non-woven mesh of PGA (Biomedical Structures). The dissociated tissues were collected, fractioned, washed multiple times, mixed with Matrigel, and loaded onto a PGA scaffold, allowing the matrigel to distribute within the pores of the scaffold. After Matrigel had solidified, the loaded scaffold was wrapped inside the omentum flap and secured in place with a suture. Dox induction of insulin^{+} cells was carried out 3-5 weeks after transplantation