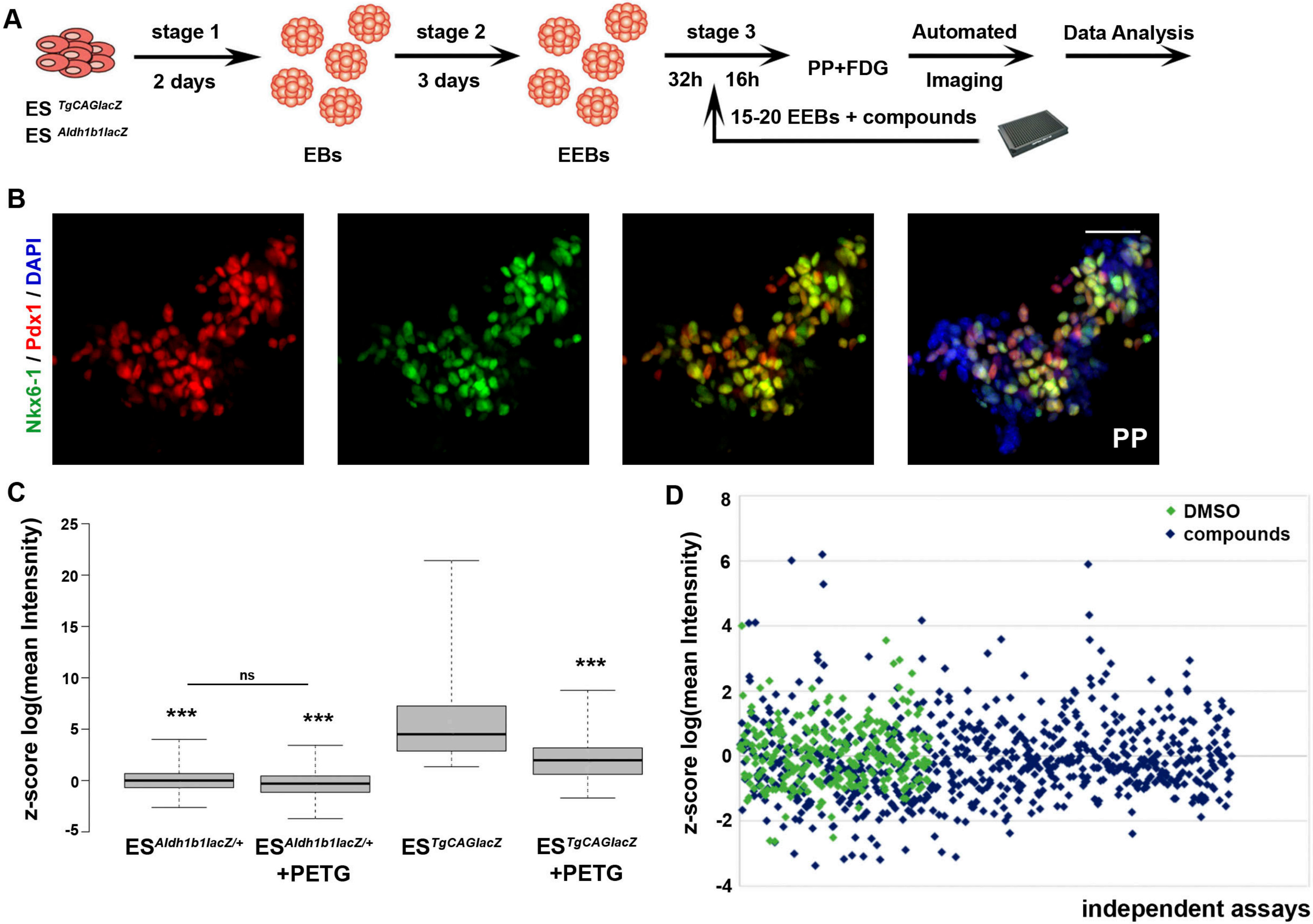
A
ES TgCAGlacZ
ES Aldh1b1lacZ
stage 1
2 days
EBs
stage 2
3 days
EEBs
stage 3
32h
16h
PP+FDG
Automated
Imaging
Data Analysis
15-20 EEBs + compounds
B
Nkx6-1 / Pdx1 / DAPI
PP
C
z-score log(mean Intensnity)
25
20
15
10
5
0
-5
ns
***
***
***
ES Aldh1b1lacZ/+
ES Aldh1b1lacZ/+ +PETG
ES TgCAGlacZ
ES TgCAGlacZ +PETG
D
z-score log(mean Intensnity)
8
6
4
2
0
-2
-4
DMSO
compounds
independent assays

Figure S1

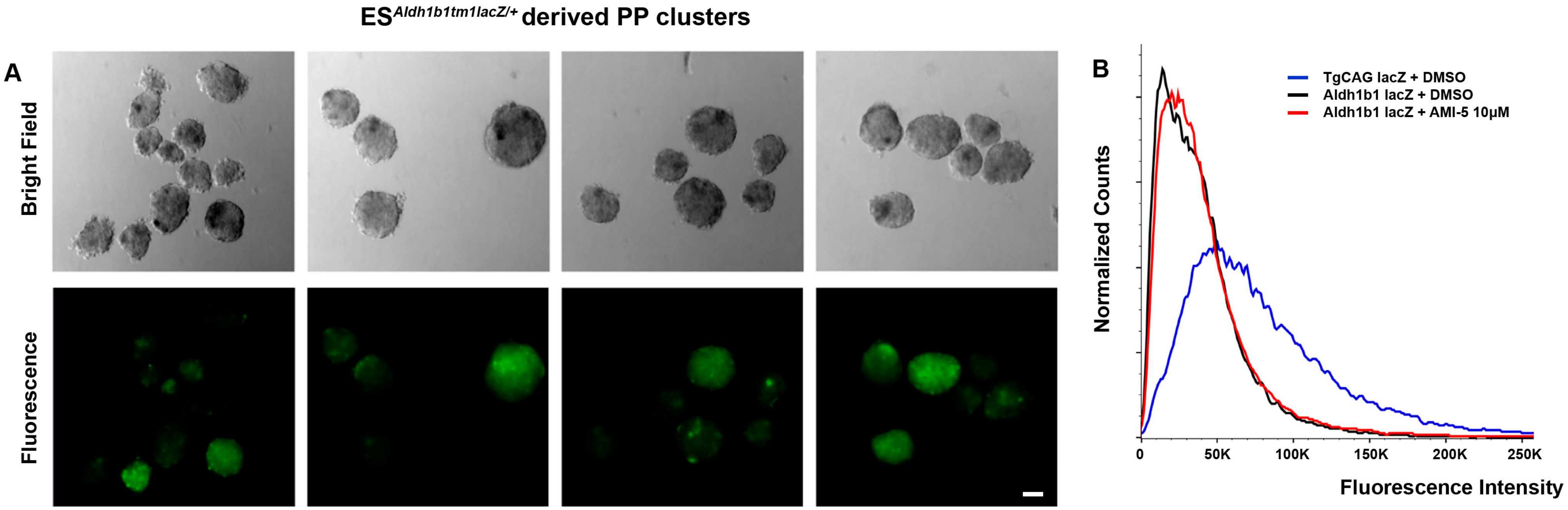
ES$^{Aldh1b1tm1lacZ/+}$ derived PP clusters
A
Bright Field
Fluorescence
B
TgCAG lacZ + DMSO
Aldh1b1 lacZ + DMSO
Aldh1b1 lacZ + AMI-5 10μM
Normalized Counts
0
50K
100K
150K
200K
250K
Fluorescence Intensity

Figure S2

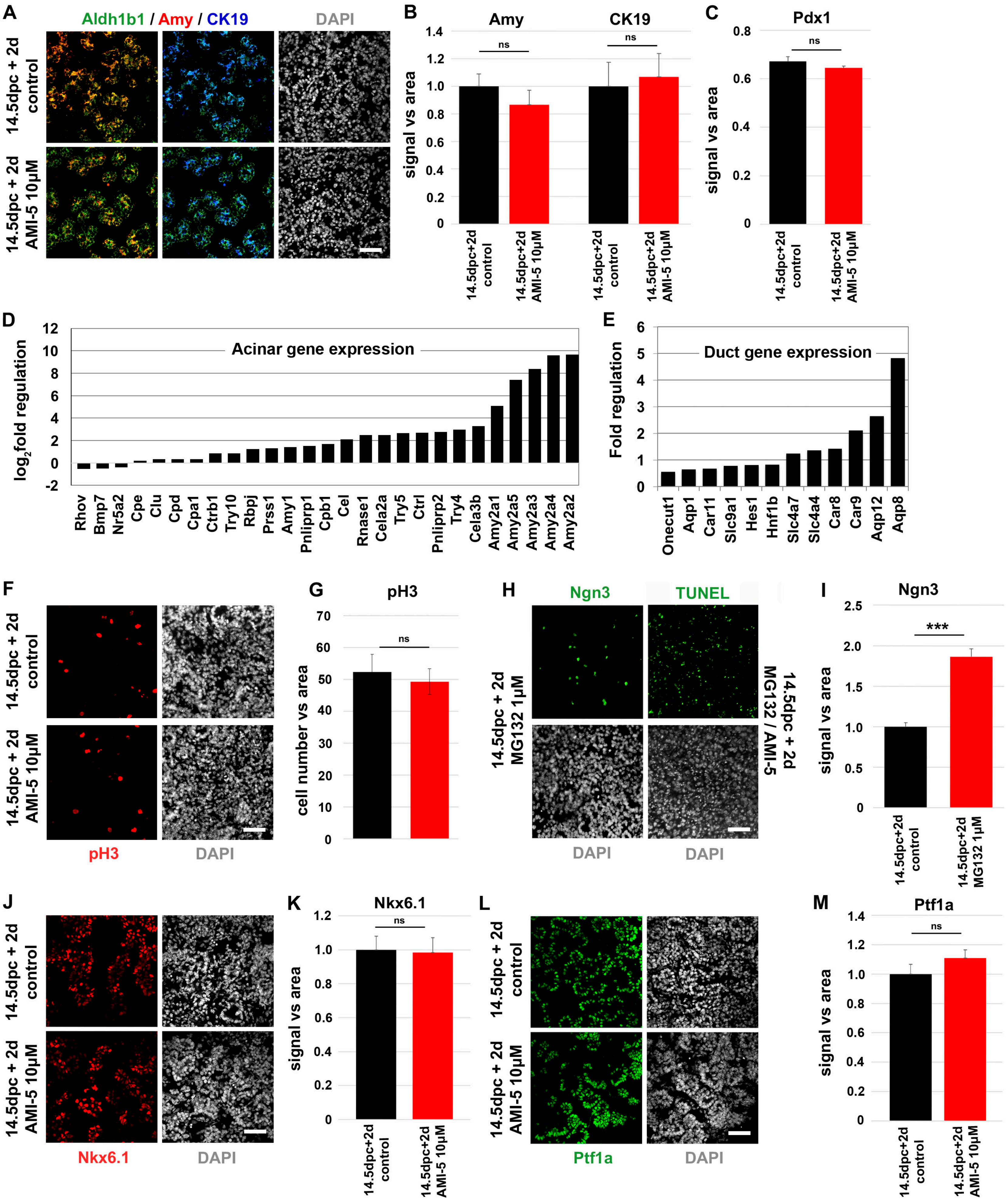
A
Aldh1b1 / Amy / CK19
DAPI
14.5dpc + 2d control
14.5dpc + 2d AMI-5 10µM
B
Amy
CK19
ns
ns
signal vs area
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
14.5dpc+2d control
14.5dpc+2d AMI-5 10µM
14.5dpc+2d control
14.5dpc+2d AMI-5 10µM
C
Pdx1
ns
signal vs area
0.8
0.6
0.4
0.2
0
14.5dpc+2d control
14.5dpc+2d AMI-5 10µM
D
Acinar gene expression
log2fold regulation
12
10
8
6
4
2
0
-2
Rhov
Bmp7
Nr5a2
Cpe
Clu
Cpd
Cpa1
Ctrb1
Try10
Rbpj
Prss1
Amy1
Pnliprp1
Cpb1
Cel
Rnase1
Cela2a
Try5
Ctrl
Pnliprp2
Try4
Cela3b
Amy2a1
Amy2a5
Amy2a3
Amy2a4
Amy2a2
E
Duct gene expression
Fold regulation
6
5
4
3
2
1
0
Onecut1
Aqp1
Car11
Slc9a1
Hes1
Hnf1b
Slc4a7
Slc4a4
Car8
Car9
Aqp12
Aqp8
F
14.5dpc + 2d control
14.5dpc + 2d AMI-5 10µM
pH3
DAPI
G
pH3
ns
cell number vs area
70
60
50
40
30
20
10
0
H
Ngn3
TUNEL
14.5dpc + 2d MG132 1µM
14.5dpc + 2d MG132 / AMI-5
DAPI
DAPI
I
Ngn3
***
signal vs area
2.5
2.0
1.5
1.0
0.5
0
14.5dpc+2d control
14.5dpc+2d MG132 1µM
J
14.5dpc + 2d control
14.5dpc + 2d AMI-5 10µM
Nkx6.1
DAPI
K
Nkx6.1
ns
signal vs area
1.2
1.0
0.8
0.6
0.4
0.2
0
14.5dpc+2d control
14.5dpc+2d AMI-5 10µM
L
14.5dpc + 2d control
14.5dpc + 2d AMI-5 10µM
Ptf1a
DAPI
M
Ptf1a
ns
signal vs area
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
14.5dpc+2d control
14.5dpc+2d AMI-5 10µM

Figure S3

Figure S1. A screen using mES cell derived pancreas progenitors to identify small molecules that induce *Aldh1b1* expression. (A) Schematic diagram of the screening protocol for the $ES^{Aldh1b1lacZ}$ and $ES^{TgCAGlacZ}$ cell lines. (B) Immunofluorescence analysis of PP clusters showed that $Pdx1^{+}$ cells comprise the large majority of PP cells while the numerous $Nkx6.1^{+}$ cells are a subset of the former. **(C)** To confirm the specificity of the screening assay, the β-gal inhibitor PETG was added during the screen. The activity of the $ES^{TgCAGlacZ}$ decreased to background levels whereas the activity of the $ES^{Aldh1b1lacZ}$ remained always negligible ($ES^{Aldh1b1lacZ}$ n=268, $ES^{Aldh1b1lacZ}$+PETG n=86, $ES^{TgCAGlacZ}$ n=86, $ES^{TgCAGlacZ}$+PETG n=86), samples were compared to the $ES^{TgCAGlacZ}$ sample. **(D)** Plot of the z-score of the log(mean fluorescence intensity) for each of the assays using compounds of the StemSelect™ Small Molecule Regulators Library (blue) or DMSO controls (green). Abbreviations: FDG, fluorescein-di-beta-D-galactopyranoside. Scale bar: 25 μm. *** $p < 0.001$.

Figure S2. Heterogeneity of AMI-5 induced *Aldh1b1* expression in mouse ES cell derived pancreatic progenitors. (A) Cellular heterogeneity of $ES^{Aldh1b1lacZ}$ derived PP clusters as revealed by the fluorescence β-galactosidase assay. **(B)** Fluorescence intensity distribution of PP clusters generated from the $ES^{Aldh1b1lacZ}$ and $ES^{TgCAGlacZ}$ lines after treatment for 16 h with DMSO or 10μM AMI-5 respectively. The Y-axis represents events normalized for the area under each curve (percentage of cells found at a given bin) (n=3). scale bar: 50μm.

Figure S3. AMI-5 does not affect specification of the acinar or duct lineages. (A) Immunofluorescence analysis of 14.5dpc pancreata after 2 days in ALI cultures shows no difference in the expression of amylase and CK19 in pancreata treated with 10 μM AMI-5. **(B)** Relative quantitation of the Amylase and CK19 fluorescence signal in 14.5dpc pancreata cultured in ALI for 2 days in the absence or presence of 10 μM AMI-5 (n=4). **(C)** Quantitation of $Pdx1^{+}$ cells following immunofluorescence in 14.5dpc pancreata after 2 days in ALI in the absence or presence of 10 μM AMI-5. **(D, E)** Fold regulation of acinar (D) and duct (E) markers at 14.5 + 2 days in ALI culture in the presence of 10 μM AMI-5 in relation to untreated controls. Only significantly regulated genes are shown ($padj \leq 0.05$). **(F, G)** Quantitation of cells in mitosis using pH3 immunofluorescence showed that mitosis rates did not change upon AMI-5 treatment (n=3). **(H)** Immunofluorescence analysis of Ngn3 expression in 14.5dpc pancreata after 2 days in ALI cultures in pancreata treated with 1 μM MG132 and TUNEL analysis following exposure to a combination of 10 μM AMI-5 and 1 μM MG132 for two days. **(I)** Relative quantitation of the number of $Ngn3^{+}$ cells in 14.5dpc pancreata cultured in ALI for 2 days in the presence of 1 μM MG132 (n=3). **(J)** Immunofluorescence analysis of 14.5dpc pancreata after 2 days in ALI cultures shows no effect in the expression of the bipotent progenitor marker Nkx6.1 in pancreata treated with 10 μM AMI-5. **(K)** Relative quantitation of the Nkx6.1 fluorescence signal in 14.5dpc pancreata cultured in ALI for 2 days in the absence or presence of 10 μM AMI-5 (n=4). **(L)** Immunofluorescence analysis of 14.5dpc pancreata after 2 days in ALI cultures shows no effect in the expression of the acinar progenitor marker Ptf1a in pancreata treated with 10 μM AMI-5. **(M)** Relative quantitation of the Ptf1a fluorescence signal in 14.5dpc pancreata cultured in ALI for 2 days in the absence or presence of 10 μM AMI-5 (n=4). Error bars show SEM; ns, not significant. Scale bar: 50μm.

Table S1 Expression of lineage specific genes in 14.5dpc + 2 pancreatic explants

Ensembl ID	Gene Symbol	Fold Regulation (+AMI-5 / CONTROL)	padj (+AMI-5 / CONTROL)	CONTROL 14.5 dpc + 2	AMI-5 14.5 dpc + 2	FULL GENE NAME
ENDOCRINE CELL MARKERS						
ENSMUSG00000047591	Mafa	0,384	1,26178E-22	418,3	157	v-maf musculoaponeurotic fibrosarcoma oncogene family, protein A
ENSMUSG00000029706	Pax4	0,390	2,66349E-07	100,6	36,3	paired box 4
ENSMUSG00000035277	Arx	0,406	2,00763E-12	255,7	100,4	aristaless related homeobox
ENSMUSG00000055197	Fev	0,483	9,60573E-10	278,5	131	FEV (ETS oncogene family)
ENSMUSG00000001517	Foxm1	0,507	8,51374E-24	1057,9	531,2	forkhead box M1
ENSMUSG00000034771	Tle2	0,518	3,60418E-18	1676	862,4	transducin-like enhancer of split 2
ENSMUSG00000001566	Mnx1	0,532	5,10738E-07	225,1	117,9	motor neuron and pancreas homeobox 1
ENSMUSG00000035187	Nkx6-1	0,559	1,31108E-21	2459,1	1368,4	NK6 homeobox 1
ENSMUSG00000004099	Dnmt1	0,570	8,98186E-25	3936,5	2239,1	DNA methyltransferase (cytosine-5) 1
ENSMUSG00000068154	Insm1	0,607	1,69432E-12	990,5	596,3	insulinoma-associated 1
ENSMUSG00000032280	Tle3	0,633	3,49833E-20	2736,5	1725,2	transducin-like enhancer of split 3
ENSMUSG00000064177	Ghrl	0,642	2,75095E-09	1371,1	872,4	ghrelin
ENSMUSG00000029644	Pdx1	0,646	1,06343E-11	1549,7	1000	pancreatic and duodenal homeobox 1
ENSMUSG00000029556	Hnf1a	0,664	8,67669E-07	507,7	335,9	HNF1 homeobox A
ENSMUSG00000041798	Gck	0,669	3,25688E-09	892,4	596,6	glucokinase
ENSMUSG00000038894	Irs2	0,678	3,38945E-07	897,9	608,2	insulin receptor substrate 2
ENSMUSG00000028800	Hdac1	0,683	2,90604E-05	415,8	282,4	histone deacetylase 1
ENSMUSG00000025880	Smad7	0,717	0,000452304	587,2	419,6	SMAD family member 7
ENSMUSG00000020661	Dnmt3a	0,722	8,76991E-10	6319,3	4565,3	DNA methyltransferase 3A
ENSMUSG00000001504	Irx2	0,732	0,02407326	232,5	168,7	Iroquois related homeobox 2
ENSMUSG00000027350	Chgb	0,742	4,70042E-05	7563,9	5592,9	chromogranin B
ENSMUSG00000000184	Ccnd2	0,768	1,20113E-09	7758,9	5948,5	cyclin D2
ENSMUSG00000006728	Cdk4	0,785	3,27135E-06	8249,3	6446,1	cyclin-dependent kinase 4
ENSMUSG00000070348	Ccnd1	0,785	8,17319E-05	1812,5	1420,2	cyclin D1
ENSMUSG00000022346	Myc	0,796	0,001341315	1768,4	1404,6	myelocytomatosis oncogene
ENSMUSG00000027715	Ccna2	0,812	0,000959199	1489,4	1209,5	cyclin A2
ENSMUSG00000037025	Foxa2	0,824	0,005218337	1294,9	1060,8	forkhead box A2
ENSMUSG00000027434	Nkx2-2	0,835	0,01122961	1165,7	972,1	NK2 homeobox 2
ENSMUSG00000017950	Hnf4a	0,835	0,02103923	1116,5	929,9	hepatic nuclear factor 4, alpha
ENSMUSG00000033016	Nfatc1	0,841	0,0326822	760,4	637,1	nuclear factor of activated T cells, cytoplasmic, calcineurin dependent 1
ENSMUSG00000027690	Slc2a2	1,197	0,006792527	1718,7	2058	solute carrier family 2 (facilitated glucose transporter), member 2
ENSMUSG00000022315	Slc30a8	1,310	0,001213559	592,3	780,2	solute carrier family 30 (zinc transporter), member 8
ENSMUSG00000027801	Tm4sf4	1,347	2,47186E-05	1877	2533,8	transmembrane 4 superfamily member 4
ENSMUSG00000035804	Ins1	1,424	2,05098E-07	20812	29682,6	insulin I
ENSMUSG00000017007	Rbpjl	0,712	0,09463305	5451	3798,5	recombination signal binding protein for immunoglobulin kappa J region-like
ENSMUSG00000060969	Irx1	0,790	0,1069848	176,9	139,1	Iroquois related homeobox 1
ENSMUSG00000012520	Phox2b	0,877	0,3971957	115,8	101	paired-like homeobox 2b
ENSMUSG00000021194	Chga	0,889	0,05045055	3418,7	3029,1	chromogranin A
ENSMUSG00000052942	Glis3	0,914	0,3770357	977,5	893,7	GLIS family zinc finger 3
ENSMUSG00000026204	Ptprn	0,927	0,5513814	1315,5	1221,1	protein tyrosine phosphatase, receptor type, N
ENSMUSG00000074622	Mafb	0,946	0,4163351	1877,5	1781,2	v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B
ENSMUSG00000010505	Myt1	0,946	0,5190577	635,3	601	myelin transcription factor 1
ENSMUSG00000019900	Rfx6	0,966	0,660487	390,9	378,3	regulatory factor X, 6
ENSMUSG00000000215	Ins2	0,979	1	33058,2	32289,7	insulin II
ENSMUSG00000042258	Isl1	1,000	0,8966366	794,3	794,6	ISL1 transcription factor, LIM/homeodomain
ENSMUSG00000034701	Neurod1	1,014	0,806127	734,2	744,6	neurogenic differentiation 1
ENSMUSG00000021587	Pcsk1	1,035	0,6144358	1363,1	1413,3	proprotein convertase subtilisin/kexin type 1
ENSMUSG00000027168	Pax6	1,035	0,8305871	960,4	995,7	paired box 6
ENSMUSG00000027419	Pcsk2	1,050	0,7591615	5764,9	6059,1	proprotein convertase subtilisin/kexin type 2
ENSMUSG00000017316	Ppy	1,057	0,4488349	440	465,8	pancreatic polypeptide
ENSMUSG00000004366	Sst	1,087	0,2832594	3709,8	4027,9	somatostatin
ENSMUSG00000031393	Mecp2	1,110	0,2442072	1023,7	1137	methyl CpG binding protein 2
ENSMUSG00000041681	Iapp	1,149	0,1215843	12634	14500,6	islet amyloid polypeptide
ENSMUSG00000000394	Gcg	1,338	0,1826873	13333,7	18115,2	glucagon

ACINAR CELL MARKERS

ENSMUSG00000096569	Amy2a2	809,002	0	13,4	20614,1	amylase 2a2
ENSMUSG00000096770	Amy2a4	770,686	6,0561E-275	4,8	9235,9	amylase 2a4
ENSMUSG00000093931	Amy2a3	333,144	1,1025E-234	7,2	4567	amylase 2a3
ENSMUSG00000074268	Amy2a5	171,255	0	29,7	6234,4	amylase 2a5
ENSMUSG00000070360	Amy2a1	34,297	2,33893E-86	11,7	557,3	amylase 2a1
ENSMUSG00000023433	Cela3b	9,714	3,83286E-23	2312,5	31315,8	chymotrypsin-like elastase family, member 3B
ENSMUSG00000054106	Try4	7,781	9,49112E-29	682,2	6341,7	trypsin 4
ENSMUSG00000025091	Pnliprp2	6,774	7,19519E-33	266,1	2034,2	pancreatic lipase-related protein 2
ENSMUSG00000031896	Ctrl	6,498	2,13088E-08	139,2	2330,8	chymotrypsin-like
ENSMUSG00000036938	Try5	6,364	4,255E-160	12906,4	84069,5	trypsin 5
ENSMUSG00000058579	Cela2a	5,657	9,87845E-16	12035,9	83825,3	chymotrypsin-like elastase family, member 2A
ENSMUSG00000035896	Rnase1	5,618	1,50524E-16	2427,7	16497,6	ribonuclease, RNase A family, 1
ENSMUSG00000026818	Cel	4,287	5,02758E-26	32007,9	146406,5	carboxyl ester lipase
ENSMUSG00000011463	Cpb1	3,182	3,26583E-05	4651,9	18657,5	carboxypeptidase B1
ENSMUSG00000042179	Pnliprp1	2,848	2,7938E-12	280429	838879,4	pancreatic lipase related protein 1
ENSMUSG00000074264	Amy1	2,657	1,49864E-19	157,2	428,2	amylase 1, salivary
ENSMUSG00000062751	Prss1	2,445	0,000907372	8425,2	23879,5	protease, serine 1 (trypsin 1)
ENSMUSG00000039191	Rbpj	2,378	2,38702E-19	180,1	436,1	recombination signal binding protein for immunoglobulin kappa J region
ENSMUSG00000031957	Ctrb1	1,828	2,77518E-06	655091,7	1217226,5	chymotrypsinogen B1
ENSMUSG00000071521	Try10	1,828	0,001322013	6590,3	12530,9	trypsin 10
ENSMUSG00000054446	Cpa1	1,275	0,002694321	112256,5	143171,5	carboxypeptidase A1, pancreatic
ENSMUSG00000020841	Cpd	1,257	0,001719374	1990,4	2513,4	carboxypeptidase D
ENSMUSG00000022037	Clu	1,248	0,002382454	30231	37919	clusterin
ENSMUSG00000037852	Cpe	1,149	0,03180342	7853,7	9016	carboxypeptidase E
ENSMUSG00000026398	Nr5a2	0,768	0,000268362	1838,4	1406,9	nuclear receptor subfamily 5, group A, member 2
ENSMUSG00000008999	Bmp7	0,707	2,94497E-05	1292,4	910,1	bone morphogenetic protein 7
ENSMUSG00000034226	Rhov	0,693	5,08307E-06	1478	1016,5	ras homolog family member V
ENSMUSG00000062478	Ctrc	1,840	0,09297903	175,8	966,2	chymotrypsin C
ENSMUSG00000023031	Cela1	1,613	0,05101512	6432,6	10941,4	chymotrypsin-like elastase family, member 1
ENSMUSG00000004266	Ptpn6	1,157	0,08275645	447,8	517,6	protein tyrosine phosphatase, non-receptor type 6
ENSMUSG00000071553	Cpa2	1,149	0,1232148	72914	83945,7	carboxypeptidase A2, pancreatic
ENSMUSG00000052271	Bhlha15	1,102	0,3549535	2128,5	2348,5	basic helix-loop-helix family, member a15
ENSMUSG00000020787	P2rx1	1,014	1	1465,3	1490,3	purinergic receptor P2X, ligand-gated ion channel, 1
ENSMUSG00000026735	Ptf1a	0,953	0,7577984	1282,9	1221,9	pancreas specific transcription factor, 1a
ENSMUSG00000017007	Rbpjl	0,712	0,09463305	5451	3798,5	recombination signal binding protein for immunoglobulin kappa J region-like

DUCT CELL MARKERS

ENSMUSG00000030762	Aqp8	4,823	6,31203E-10	170,8	1106	aquaporin 8
ENSMUSG00000045091	Aqp12	2,639	0,000124426	142,3	436,9	aquaporin 12
ENSMUSG00000028463	Car9	2,114	7,33132E-20	554,2	1187,7	carbonic anhydrase 9
ENSMUSG00000041261	Car8	1,424	0,004028652	219,1	315	carbonic anhydrase 8
ENSMUSG00000060961	Slc4a4	1,357	8,72356E-05	496,6	676,6	solute carrier family 4 (anion exchanger)
ENSMUSG00000021733	Slc4a7	1,240	0,007014825	612,2	759,7	solute carrier family 4, sodium bicarbonate cotransporter, member 7
ENSMUSG00000020679	Hnf1b	0,829	0,006654343	1107,1	914,3	HNF1 homeobox B
ENSMUSG00000022528	Hes1	0,812	0,001459381	1265,5	1028,7	hairy and enhancer of split 1
ENSMUSG00000028854	Slc9a1	0,785	0,000586766	1011,4	790,7	solute carrier family 9 (sodium/hydrogen exchanger)
ENSMUSG00000003273	Car11	0,674	0,03759337	136,8	89,9	carbonic anhydrase 11
ENSMUSG00000004655	Aqp1	0,651	4,85829E-09	1287,2	836,3	aquaporin 1
ENSMUSG00000043013	Onecut1	0,559	1,35816E-10	542,1	299	one cut domain, family member 1
ENSMUSG00000024411	Aqp4	1,072	0,403821	823,4	882,2	aquaporin 4
ENSMUSG00000020911	Krt19	1,028	0,9379075	2462,5	2534,8	keratin 19
ENSMUSG00000023259	Slc26a6	1,007	0,7159408	301,4	302,8	solute carrier family 26, member 6
ENSMUSG00000027765	P2ry1	0,966	1	168,3	162,3	purinergic receptor P2Y, G-protein coupled 1
ENSMUSG00000027562	Car2	0,959	1	177,6	170,1	carbonic anhydrase 2
ENSMUSG00000010175	Prox1	0,927	0,2608753	1344,2	1240,2	prospero homeobox 1
ENSMUSG00000000567	Sox9	0,914	0,19618	2976,5	2719,3	SRY (sex determining region Y)-box 9
ENSMUSG00000052942	Glis3	0,914	0,3770357	977,5	893,7	GLIS family zinc finger 3