

Sox9 coordinates a transcriptional network in pancreatic progenitor cells

F. C. Lynn*, S. B. Smith*, M. E. Wilson*[†], K. Y. Yang*, N. Nekrep*, and M. S. German*^{‡§}

*Diabetes Center, Hormone Research Institute, and [‡]Department of Medicine, University of California, San Francisco, CA 94143

Communicated by William J. Rutter, Synergenics, LLC, San Francisco, CA, May 4, 2007 (received for review September 29, 2006)

During pancreas development, both the exocrine and endocrine lineages differentiate from a common pool of progenitor cells with similarities to mature pancreatic duct cells. A small set of transcription factors, including Tcf2, Onecut1, and Foxa2, has been identified in these pancreatic progenitor cells. The Sry/HMG box transcription factor Sox9 is also expressed in the early pancreatic epithelium and is required for normal pancreatic exocrine and endocrine development in humans. In this study, we found Sox9 in mice specifically expressed with the other progenitor transcription factors in both pancreatic progenitor cells and duct cells in the adult pancreas. Sox9 directly bound to all three genes *in vitro* and in intact cells, and regulated their expression. In turn, both Foxa2 and Tcf2 regulated Sox9 expression, demonstrating feedback circuits between these genes. Furthermore, Sox9 activated the expression of the proendocrine factor Neurogenin3, which also depends on the other members of the progenitor transcription network. These studies indicate that Sox9 plays a dual role in pancreatic progenitor cells: both maintaining a stable transcriptional network and supporting the programs by which these cells differentiate into distinct lineages.

The distinct cell types that populate the mature pancreas derive from a common pool of undifferentiated progenitor cells that form the early pancreatic bud. Although recent work has begun to define the gene-expression cascades that drive the differentiation of the mature pancreatic cell lineages (1), the gene-expression programs that maintain the progenitor cells and also permit their proper differentiation are not well understood.

The pancreas first appears as a cluster of cells budding from the dorsal aspect of the gut endoderm at embryonic day (e)9.5 in the mouse embryo, and a ventral bud appears approximately a day later. The endoderm cells that make up the buds rapidly divide and form a branching epithelium surrounded by mesenchyme. As this epithelial progenitor cell population expands during the first 3 days of pancreatic growth, a small number of cells delaminate from this epithelium and stop dividing. Most of these postmitotic cells express the *glucagon* gene, and the remainder express *insulin*, but they lack many key characteristics of mature pancreatic endocrine cells (2–5).

Starting at e13 in the mouse, the embryonic pancreas undergoes a dramatic transformation in a synchronized wave of differentiation termed the secondary transition. The branching epithelium organizes into ducts, with the ends of the ducts differentiating into exocrine cells, whereas the cells lining the ducts retain progenitor cell characteristics and the ability to generate endocrine cells. In scattered progenitor cells along the ducts, notch signaling is inactivated, allowing the transient expression of the proendocrine bHLH transcription factor Neurogenin3 that initiates endocrine differentiation (6, 7).

The epithelial progenitor cells that populate the pancreatic buds before the secondary transition differ from the progenitor cells along the ducts from which the major wave of differentiated cells originates after e13. Both cell types express the transcription factors Tcf2, Onecut1, and Foxa2 (8–11), but the early population of progenitor cells also express a number of transcription factors that later become restricted to the endocrine lineage, including Hb9, Pdx1, Nkx2.2, Nkx6.1, and Sox4 (4,

12–17) and at least one factor that is later restricted to the exocrine cells, Ptf1a (18). Interestingly, Tcf2, Onecut1, and Foxa2 are simultaneously expressed in several regions in the early endoderm, including the developing gut and liver, and persist in the duct cells of the adult pancreas.

Recently, several members of the SRY/HMG box (Sox) family of transcription factors (19–21) have been detected during pancreatic development (14, 17, 22, 23). Among these, Sox9 is of particular interest because of its high level of expression in the early pancreatic epithelium (14). In addition, mice carrying *lacZ* inserted into the *Sox9* gene have demonstrated that Sox9 is expressed in the adult pancreatic ductal epithelium, and lineage tracing in the same study revealed that pancreatic cells from all lineages derive from Sox9-expressing precursors (24). Haploinsufficiency of *SOX9* causes campomelic dysplasia (CD) in humans, which is characterized by a severe skeletal dysplasia, male (XY) sex reversal (25), and defects in the development of both the exocrine and endocrine pancreas (26).

To better understand the mechanisms by which Sox9 regulates pancreatic development, we explored its role in the transcriptional network that maintains gene expression in the pancreatic progenitors and its role in the transcriptional cascade that drives the progenitor cells to differentiate into endocrine cells. Here, we demonstrate that Sox9 regulates a network of transcription factors that both controls progenitor cell identity and supports endocrine cell differentiation within the developing organ.

Results

Pancreatic Expression of Sox9. Immunohistochemical staining of sections from developing mouse pancreas detected Sox9 immunoreactivity at e10.5 and e12.5 in the dorsal (Fig. 1 *A–D*) and ventral (data not shown) pancreas along with Pdx1 in primary pancreatic progenitor cells but not in differentiated glucagon-expressing cells or in Pdx1-expressing cells in the adjacent duodenum. At e14.5, after the secondary transition, Sox9 persisted along with lowered levels of Pdx1 in the secondary pancreatic progenitor cells along the ducts but not in the intensely Pdx1-staining differentiated cells (Fig. 1 *E*). At e15.5, the proendocrine transcription factor Neurogenin3 was detected in a subset of Sox9-positive cells scattered along the ducts, as well as Sox9-negative cells adjacent to the ducts (Fig. 1 *G* and *H*). Confocal microscopy verified the coexpression of Sox9 with Neurogenin3 in a subset of pancreatic duct cells at e14.5 (Fig. 2) but not with more mature islet markers Nkx2.2 and Isl1 (data not shown).

Author contributions: F.C.L. and S.B.S. contributed equally to this work; F.C.L., S.B.S., M.E.W., and M.S.G. designed research; F.C.L., S.B.S., M.E.W., K.Y.Y., and N.N. performed research; F.C.L., S.B.S., M.E.W., K.Y.Y., and M.S.G. analyzed data; and F.C.L. and M.S.G. wrote the paper.

The authors declare no conflict of interest.

Abbreviation: e(*n*), embryonic day (*n*).

[†]Present address: Metabolex, Inc, 3876 Bay Center Place, Hayward, CA 94545.

[§]To whom correspondence should be addressed at: University of California Diabetes Center, 513 Parnassus Avenue, San Francisco, CA 94143-0534. E-mail: mgerman@diabetes.ucsf.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0704054104/DC1.

© 2007 by The National Academy of Sciences of the USA

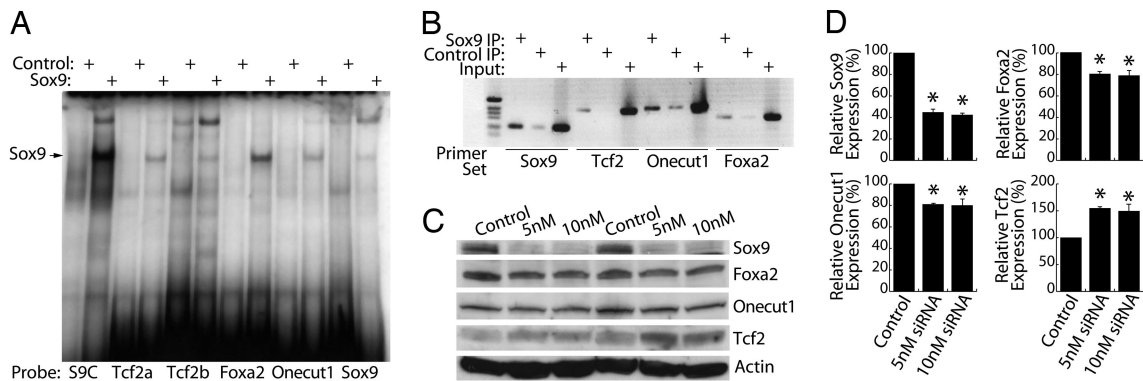


Fig. 4. Sox9 regulates the other progenitor cell transcription factors. (A) DNA binding of *in vitro*-translated Sox9 or Luciferase (Control) protein was tested by EMSA. Double-stranded, radiolabeled oligonucleotide probes contained binding sites from the indicated promoters and a consensus Sox9-binding site (S9C). (B) Chromatin IP studies were performed by immunoprecipitating cross-linked chromatin with antiserum against Sox9 or with control IgG. 4 fragments of the indicated promoters were amplified by PCR from the precipitates or the input DNA. (C) mPAC L20 cells were transfected with synthetic double-stranded siRNA directed against Sox9 or with a control siRNA. Protein levels were assessed by Western blot 48 h after transfection. (D) Quantification is shown for Western blots. Data represent the mean \pm SEM of three independent experiments; statistical analyses were carried out by using one-way ANOVA, followed by the Newman–Keuls post hoc test. Asterisks indicate significant difference ($P < 0.001$) from control conditions.

against Sox9, selective immunoprecipitation of regions of the *Neurog3* promoter containing putative Sox9-binding sites was demonstrated by PCR (Fig. 3B). The Sox9 antiserum did not immunoprecipitate the insulin promoter, the upstream stimulatory factor (USF) promoter, or the phosphoenolpyruvate carboxykinase promoter (data not shown).

EMSA mapped these binding sites in more detail. *In vitro*-translated mouse Sox9 bound to sequences at four distinct sites within the *NEUROG3* promoter with affinities similar to a consensus Sox9-binding site (28) (Fig. 3C). It should be noted that some lanes also contain a slower mobility band. This larger complex may result from dimerization of Sox9 on the DNA, which has been shown to affect the function of Sox9 (29).

Finally, we used RNA interference technology to test the importance of Sox9 in normal Neurogenin3 expression. By stably expressing an shRNA targeting Sox9, we established mPAC L20 mouse pancreatic cells with reduced Sox9 expression. Because the mPAC L20 cell line does not normally express Neurogenin3, an adenovirus expressing the bHLH protein Mash1 was used to induce Neurogenin3 expression (30). In mPAC L20 cells containing normal levels of Sox9, Mash1 activated Neurogenin3 expression as described (30); however, Mash1 could not induce Neurogenin3 expression in the Sox9-deficient cells (Fig. 3D).

Sox9 Regulates a Network of Factors Upstream of Neurogenin3.

Because Sox9 expression parallels that of other pancreatic progenitor and ductal cell transcription factors Tcf2, Oncut1, and Foxa2, all of which have been implicated in the expression of Neurogenin3, we tested the hypothesis that Sox9 coordinates the expression of these factors and establishes an environment that is permissive for Neurogenin3 induction in pancreatic progenitor cells. Scanning the upstream sequences of these genes identified conserved consensus Sox9-binding sequences in each. Binding to each of these sites was confirmed by EMSA (Fig. 4A) and ChIP from mPAC L20 cells (Fig. 4B) and embryonic pancreas [Oncut1 only, supporting information (SI) Fig. 7]. Interestingly, a conserved consensus binding site in the *Sox9* gene itself was also verified by EMSA and ChIP (Fig. 4A and B), suggesting that Sox9 may also autoregulate its own expression in the pancreas.

To test whether Sox9 regulates these genes in intact cells, Sox9 expression was reduced by RNAi with synthetic double-stranded RNA oligonucleotides transfected into mPAC L20 cells. Transient knockdown of Sox9 levels by $\approx 50\%$ resulted in a 20–30%

reduction in the expression of Foxa2 and Oncut1 and a 50% increase in the expression of Tcf2 (Fig. 4C and D).

Finally, we used the same siRNA approach to test the role of Tcf2 and Foxa2 in regulating Sox9 and the other members of the network (Fig. 5). Targeted reduction of Tcf2 expression by siRNA using transfected double-stranded RNA oligonucleotides in mPAC L20 cells caused modest reductions in the expression of Foxa2 but no significant change in Oncut1 expression. Foxa2 knockdown produced a more dramatic decrease in Sox9 expression but did not affect Oncut1 or Tcf2 expression. These experiments demonstrate that Foxa2 is necessary for the maintenance of Sox9 expression in these cells and that Sox9 participates in a network of factors including Foxa2, Tcf2, and Oncut1 that are required for normal pancreatic organogenesis.

Discussion

During embryonic development, the differentiated cells of the pancreas derive from a transient pool of stable, replicating progenitor cells. The data presented here demonstrate that Sox9 plays a central role in both maintaining a gene network that

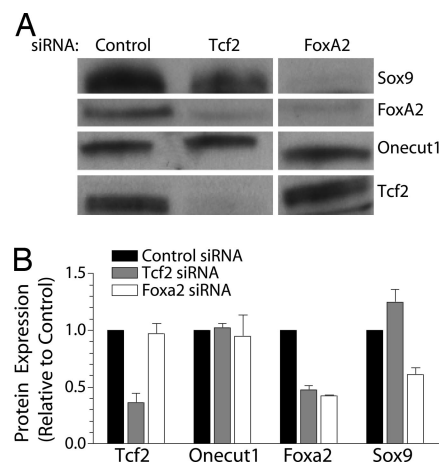


Fig. 5. Foxa2 and Tcf2 regulate Sox9 expression. mPAC L20 cells were transfected with a control siRNA or siRNAs against Foxa2 and Tcf2 at a concentration of 10 nM. Protein levels were assessed by Western blot 48 h after transfection. (B) Quantification for Western blots shown in A. Data represent the mean \pm SEM of four independent experiments.

pSilencerTM-4.1-CMV vector (Ambion, Austin, TX). Transient knockdown was performed by transfecting small interfering RNA duplexes targeting Sox9 (AAAGAGATGTGAGUCU-GUCCGGGGAUC), Foxa2 (GGUCUCGGGUCUGA-UUUAATT), Tcf2 (GAGAGUAUGGAAAACCGGCUU), or Control siRNA #1 (Ambion) into mPAC L20 cells by using HiPerfect (Qiagen, Valencia, CA). Low concentrations (5–10 nM) of siRNA duplexes were used to minimize potential off-target effects.

EMSA. Synthetic DNA oligonucleotides were labeled, and EMSAs were performed in the presence of 500 ng of poly(deoxyinosine-deoxycytidine) per 10 μ l of binding mix as described (6). Protein was generated *in vitro* with the TNT-coupled Reticulocyte Lysate System (Promega) in 50 μ l of total volume from 1 μ g of DNA; 1 μ l (\approx 5 ng of protein) of the reaction mix was then used per binding reaction.

The oligonucleotides sequences are in SI Table 1.

ChIP Assays. Mouse mPAC L20 cells grown to 70% confluence were fixed at room temperature with 1% formaldehyde for 10 min. Cross-linking was quenched by the addition of glycine (125 mM). Cells were then washed in ice-cold PBS, scraped from the growth surface, and pelleted. The pellets were washed and lysed as described (37). Chromatin was then sheared to \approx 1-kb fragments by using sonication and cleared by centrifugation. Antibody binding was carried out overnight at 4°C by using 200 μ g of chromatin and normal rabbit IgG (sc-2027; Santa Cruz Biotechnology) or rabbit anti-Sox9 antibody (sc-20095; Santa Cruz Biotechnology). Antibody-bound complexes were coupled for 1 h to previously blocked (1 mg/ml BSA, 0.1 mg/ml salmon sperm DNA) protein A-Sepharose beads (GE Healthcare/Amersham, Piscataway, NJ). Coupled beads were washed in 1.5 ml each of: TSE1 [0.1% SDS/1% Triton X-100/2 mM EDTA/20 mM Tris-HCl (pH 8.1)/150 mM NaCl], TSE2 (same as TSE1 but with 500 mM NaCl), ChIP Buffer 3 [0.25M LiCl/1% Nonidet P-40/1% sodium deoxycholate/1 mM EDTA/10 mM Tris-HCl (pH 8.1)] and TE, all containing protease inhibitors (Roche, Indianapolis, IN) and

PMSF. Bound material was then eluted from the beads in 50 mM NaHCO₃ containing 1% SDS; cross-linking was reversed by incubation at 67°C for 5 h; and DNA was precipitated in ethanol. DNA pellets were then resuspended in TE and treated successively with 200 μ g/ml RNase A and with 75 μ g/ml proteinase K for 2 h each at 37°C and 55°C, respectively. DNA was then phenol/chloroform/isoamyl alcohol-extracted and precipitated with 5 μ g of glycogen and then resuspended in 20 μ l of water.

Five to 15 ng of immunoprecipitated DNA per reaction was assayed by PCR with the primers in SI Table 1 to test for the precipitation of specific promoter fragments.

Western Blotting. Whole-cell lysates were harvested with 100°C SDS/PAGE loading buffer [62.5 mM Tris (pH 6.8)/1 mM sodium vanadate/1 mM sodium fluoride/2% SDS/10% glycerol], incubated for 10 min at 100°C, and sonicated by using an ultrasonic needle tip processor (Tekmar, Cincinnati, OH) for 1 min at 70% maximal power (4–6 W). Cellular debris was removed from the lysate by centrifugation at 10,000 \times g for 10 min at 20°C. Supernatants were collected, and proteins were separated by electrophoresis through 10% Tris-HCl polyacrylamide gels. Proteins were transferred to PVDF membranes (Hybond-P; Amersham Biosciences, Piscataway, NJ) and incubated with rabbit anti-neurogenin3 [1:1,000; (41)], rabbit anti-sox9 (1:500; Santa Cruz Biotechnology), rabbit anti-onecut1 (1:500; Santa Cruz Biotechnology), and rabbit anti-tcf2 (1:1,000, Chemicon, Temecula, CA) for 3 h at room temperature. After washing, membranes were incubated with horseradish peroxidase-conjugated secondary antibodies for 1 h at room temperature. Finally, membranes were imaged by using the enhanced chemiluminescence kit (ECL; Amersham).

We thank members of the M.S.G. and Hebrok laboratories, John Rubenstein, and Didier Stainier for helpful advice and criticism. This work was supported by National Institutes of Health (NIH) Grants DK21344 and DK61245, cores from NIH Grant P30 DK063720, Larry L. Hillblom Foundation Grant 2002/1(E), the Nora Eccles Treadwell Foundation, and Juvenile Diabetes Research Foundation Fellowship Awards 3-98-341 (to M.E.W.) and 3-2004-276 (to F.C.L.).

- Wilson ME, Scheel D, German MS (2003) *Mech Dev* 120:65–80.
- Wilson ME, Kalamaras JA, German MS (2002) *Mech Dev* 115:171–176.
- Teitelman G, Alpert S, Polak JM, Martinez A, Hanahan D (1993) *Development (Cambridge, UK)* 118:1031–1039.
- Sander M, Sussel L, Connors J, Scheel D, Kalamaras J, Dela Cruz F, Schwitzgebel V, Hayes-Jordan A, German M (2000) *Development (Cambridge, UK)* 127:5533–5540.
- Oster A, Jensen J, Serup P, Galante P, Madsen OD, Larsson LI (1998) *J Histochem Cytochem* 46:707–715.
- Lee JC, Smith SB, Wataha H, Lin J, Scheel D, Wang J, Mirmira RG, German MS (2001) *Diabetes* 50:928–936.
- Apelqvist A, Li H, Sommer L, Beatus P, Anderson DJ, Honjo T, Hrabe de Angelis M, Lendahl U, Edlund H (1999) *Nature* 400:877–881.
- Coffinier C, Barra J, Babinet C, Yaniv M (1999) *Mech Dev* 89:211–213.
- Jacquemin P, Lemaigre FP, Rousseau GG (2003) *Dev Biol* 258:105–116.
- Jacquemin P, Durvieux SM, Jensen J, Godfraind C, Gradwohl G, Guillemot F, Madsen OD, Carmeliet P, Dewerchin M, Collen D, et al. (2000) *Mol Cell Biol* 20:4445–4454.
- Lee CS, Sund NJ, Behr R, Herrera PL, Kaestner KH (2005) *Dev Biol* 278:484–495.
- Sussel L, Kalamaras J, Hartigan-O'Connor DJ, Meneses JJ, Pedersen RA, Rubenstein JL, German MS (1998) *Development (Cambridge, UK)* 125:2213–2221.
- Ahlgren U, Jonsson J, Edlund H (1996) *Development (Cambridge, UK)* 122:1409–1416.
- Lioubinski O, Muller M, Wegner M, Sander M (2003) *Dev Dyn* 227:402–408.
- Li H, Arber S, Jessell TM, Edlund H (1999) *Nat Genet* 23:67–70.
- Harrison KA, Thaler J, Pfaff SL, Gu H, Kehrl JH (1999) *Nat Genet* 23:71–75.
- Wilson ME, Yang KY, Kalousova A, Lau J, Kosaka Y, Lynn FC, Wang J, Mrejen C, Episkopou V, Clevers HC, German MS (2005) *Diabetes* 54:3402–3409.
- Kawaguchi Y, Cooper B, Gannon M, Ray M, MacDonald RJ, Wright CV (2002) *Nat Genet* 32:128–134.
- Wegner M, Stolt CC (2005) *Trends Neurosci* 28:583–588.
- Wegner M (2005) *Pigment Cell Res* 18:74–85.
- Wegner M (1999) *Nucleic Acids Res* 27:1409–1420.
- Mavropoulos A, Devos N, Biemar F, Zecchin E, Argenton F, Edlund H, Motte P, Martial JA, Peers B (2005) *Dev Biol* 285:211–223.
- Lee YH, Saint-Jeannet JP (2003) *Int J Dev Biol* 47:459–462.
- Akiyama H, Kim JE, Nakashima K, Balmes G, Iwai N, Deng JM, Zhang Z, Martin JF, Behringer RR, Nakamura T, de Crombrughe B (2005) *Proc Natl Acad Sci USA* 102:14665–14670.
- Wagner T, Wirth J, Meyer J, Zabel B, Held M, Zimmer J, Pasantes J, Bricarelli FD, Keutel J, Hustert E, et al. (1994) *Cell* 79:1111–1120.
- Piper K, Ball SG, Keeling JW, Mansoor S, Wilson DI, Hanley NA (2002) *Mech Dev* 116:223–226.
- Mertin S, McDowall SG, Harley VR (1999) *Nucleic Acids Res* 27:1359–1364.
- Harley VR, Lovell-Badge R, Goodfellow PN (1994) *Nucleic Acids Res* 22:1500–1501.
- Bernard P, Tang P, Liu S, Dewing P, Harley VR, Vilain E (2003) *Hum Mol Genet* 12:1755–1765.
- Gasa R, Mrejen C, Leachman N, Otten M, Barnes M, Wang J, Chakrabarti S, Mirmira R, German M (2004) *Proc Natl Acad Sci USA* 101:13245–13250.
- Maestro MA, Boj SF, Luco RF, Pierreux CE, Cabedo J, Servitja JM, German MS, Rousseau GG, Lemaigre FP, Ferrer J (2003) *Hum Mol Genet* 12:3307–3314.
- Lee CS, Sund NJ, Vatamaniuk MZ, Matschinsky FM, Stoffers DA, Kaestner KH (2002) *Diabetes* 51:2546–2551.
- Lantz KA, Kaestner KH (2005) *Clin Sci (London)* 108:195–204.

34. Sun Z, Hopkins N (2001) *Genes Dev* 15:3217–3229.
35. Horikawa Y, Iwasaki N, Hara M, Furuta H, Hinokio Y, Cockburn BN, Lindner T, Yamagata K, Ogata M, Tomonaga O, *et al.* (1997) *Nat Genet* 17:384–385.
36. Zaret KS (2002) *Nat Rev Genet* 3:499–512.
37. Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK, *et al.* (2004) *Science* 303:1378–1381.
38. Kyrmizi I, Hatzis P, Katrakili N, Tronche F, Gonzalez FJ, Talianidis I (2006) *Genes Dev* 20:2293–2305.
39. Costa RH, Kalinichenko VV, Holterman AX, Wang X (2003) *Hepatology* 38:1331–1347.
40. Friedman JR, Kaestner KH (2006) *Cell Mol Life Sci* 63:2317–2328.
41. Schwitzgebel VM, Scheel DW, Conners JR, Kalamaras J, Lee JE, Anderson DJ, Susse L, Johnson JD, German MS (2000) *Development (Cambridge, UK)* 127:3533–3542.