Oogenesis requires germ cell-specific transcriptional regulators Sohlh1 and Lhx8

Stephanie A. Pangas*, Youngsok Choi†, Daniel J. Ballow†, Yangu Zhao‡, Heiner Westphal‡, Martin M. Matzuk*§¶, and Aleksandar Rajkovic†

Departments of *Pathology, †Obstetrics and Gynecology, §Molecular and Cellular Biology, and ¶Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030; and ‡Laboratory of Mammalian Genes and Development, National Institute of Child Health and Human Development, Bethesda, MD 20892

Edited by Brigid L. M. Hogan, Duke University Medical Center, Durham, NC, and approved April 12, 2006 (received for review February 9, 2006)

Mammalian oogenesis requires oocyte-specific transcriptional regulators. The full complement of oocyte-specific transcription factors is unknown. Here, we describe the finding that Sohlh1, a spermatogenesis and oogenesis basic helix–loop–helix transcription factor in females, is preferentially expressed in oocytes and required for oogenesis. Sohlh1 disruption perturbs follicular formation in part by causing down-regulation of two genes that are known to disrupt folliculogenesis: newborn ovary homeobox gene (Nobox) and factor in the germ-line alpha (Figla). In addition, we show that Lhx8 is downstream of Sohlh1 and critical in fertility. Thus, Sohlh1 and Lhx8 are two germ cell-specific, critical regulators of oogenesis.

infertility | oocyte | ovary | reproduction | folliculogenesis

Mou se primordial germ cells, at embryonic days 9.5–11.5 (E9.5–E11.5), migrate to the urogenital ridges from the proximal epiblast. Mitotic division of primordial germ cells, coupled with incomplete cytokinesis, results in oocytes in clusters (also called cysts) at ≈E10.5 (1, 2). Female germ cells, at ≈E13.5, begin entry into prophase I of meiosis and arrest in the diplotene stage of the first meiotic division. Oocytes arrested in meiosis I are thought to remain arrested until the time of ovulation. Germ cell clusters formed in the embryonic gonad break down shortly after birth in the mouse. Breakdown of germ cell cysts results in oocytes enveloped by somatic pregranulosa cells (now called primordial follicles) (1). Primordial follicles represent a reservoir of follicles that are recruited periodically to grow into primary and more advanced follicular structures.

The breakdown of germ cell clusters, formation of primordial follicles, and transition to primary follicles represents a critical period of follicle formation. Transcription of numerous oocyte-specific genes, such as growth and differentiation factor 9 (Gdf9), bone morphogenetic protein 15 (Bmp15), and zona pellucida genes 1–3 (Zp1–3) (3–6), commences during early folliculogenesis. Regulation of these genes is in part due to expression of two known oocyte-specific transcription factors, Figla (7) and Nobox (3). FIGLA is a basic helix-loop–helix transcription factor that regulates expression of the zona pellucida genes (7, 8). Nobox is a homeobox gene that is necessary for expression of several key oocyte-specific genes, including Gdf9, Bmp15, and Pou5f1 but not Figla or Zp1–3 (3). However, other oocyte-specific transcriptional regulators likely exist, because numerous oocyte-specific genes are not affected by the lack of Nobox and Figla. Here, we describe our finding that Sohlh1 and Lhx8, which are both preferentially expressed during oogenesis in females, are critical in early folliculogenesis.

Results and Discussion

We identified Sohlh1 by an in silico subtraction strategy to identify genes that are preferentially expressed during early folliculogenesis (9). Sohlh1 encodes a basic helix-loop–helix transcription factor with homologues in humans and other placental mammals (Fig. 8, which is published as supporting information on the PNAS web site). In females, ovaries preferentially express Sohlh1 transcripts as shown by multitissue RT-PCR analysis (Fig. 9, which is published as supporting information on the PNAS web site). Embryonic ovaries express readily

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: En, embryonic day n; ChIP, chromatin immunoprecipitation; GCNA1, germ cell nuclear antigen 1.

1To whom correspondence should be addressed at: Department of Obstetrics and Gynecology, Baylor College of Medicine, 1709 Dryden Road, Suite 1100, Houston, TX 77030. E-mail: rajkovic@bcm.tmc.edu.

© 2006 by The National Academy of Sciences of the USA
expression studies show that oning may therefore regulate SOHLH1 function. The above

cated both in the nucleus and the cytoplasm. Nuclear chaper-
signal, and immunohistochemistry shows that SOHLH1 is lo-
expression pattern that differs from other known oocyte-specific
was used to calculate

ferences were observed between the WT and knockout ovaries. Data are repre-

ected in germ cell cysts, primordial follicles, and primary

transcripts are present in oocytes of germ cell

transcripts are preferentially

expressed in primordial oocytes but disappear rapidly as the

ectos are recruited to form primary and secondary (multilayer

ects and perinatal) follicles (Fig. 1 C and D). SOHLH1 protein is
detected in germ cell cysts, primordial follicles, and primary

ects but is undetectable by the secondary follicle stage (Fig. 1 E–H). SOHLH1 protein lacks the classic nuclear localization

signal, and immunohistochemistry shows that SOHLH1 is lo-
cated both in the nucleus and the cytoplasm. Nuclear chaper-
oning may therefore regulate SOHLH1 function. The above

ession studies show that Sohlh1 has a RNA and protein

expression pattern that differs from other known oocyte-specific
detectable levels of Sohlh1 at E15.5, at the time when oocytes
have entered meiosis I, although a low level of Sohlh1 mRNA
expression is detectable at E13.5. Newborn mouse ovaries con-
tain oocyte clusters confined to germ cell cysts and primordial
follicles. Sohlh1 transcripts are present in oocytes of germ cell
cysts as well as primordial follicles in the newborn ovary (Fig. 1
A and B). In adult ovaries, Sohlh1 transcripts are preferentially
expressed in primordial oocytes but disappear rapidly as the
oocytes are recruited to form primary and secondary (multilayer
and preantral) follicles (Fig. 1 C and D). SOHLH1 protein is
detected in germ cell cysts, primordial follicles, and primary
follicles but is undetectable by the secondary follicle stage (Fig. 1
E–H). SOHLH1 protein lacks the classic nuclear localization
signal, and immunohistochemistry shows that SOHLH1 is located
both in the nucleus and the cytoplasm. Nuclear chaperonin may therefore regulate SOHLH1 function. The above
expression studies show that Sohlh1 has a RNA and protein
expression pattern that differs from other known oocyte-specific

transcription regulators, Nobox and Figla, and suggest that
Sohlh1 plays a unique role in early folliculogenesis.

To define the roles of SOHLH1 in early stages of folliculogenesis, we generated a targeted deletion of the first three exons that encode
the SOHLH1 basic helix–loop–helix domain as well as 500 nucleo-

tides upstream of the putative transcription initiation site (10).

heterozygous Sohlh1 +/- matings produced expected Mendelian

tions, and females averaged 8.1 ± 2.0 pups per litter (n = 45
breeding pairs) over a 6-month period and remained fertile for at
least 9 months. Litter sizes were not statistically different from the
WT average (8.4 ± 2.0 pups per litter). In contrast, all Sohlh1 +/-

 homozygous null females were infertile, with atrophic ovaries that,
on histologic examination, lacked oocytes at 10 weeks of age on a
C57BL/6J;129SvEvBrd hybrid background (Fig. 2 A–C). Adult

Sohlh1 +/- ovaries occupied ~1/10 of the WT volume.

To determine the onset of histologic changes and germ cell loss, we
compared ovarian development in WT and Sohlh1 +/- mice.

Newborn WT and Sohlh1 +/- ovaries show no apparent differences
in morphology or histology, and histomorphometric analysis found
no statistically significant differences in the number of oocytes
present (Fig. 2D). We used antibodies against two germ cell-specific
antigens, GCNA1 (germ cell nuclear antigen 1) and MSY2 (germ
cell-specific Y box protein), to study oocyte development in WT
and mutants. GCNA1 is a nuclear marker for the germ cell lineage.
It is expressed in germ cells after their arrival at the gonad until the

Fig. 2. Sohlh1 adult knockout anatomy, histology, and histomorphometric
analysis. (A) Gross reproductive tracts dissected from WT, heterozygous (+/-),
and homozygous (-/-) Sohlh1 mice. Note markedly smaller ovaries in Sohlh1 +/-
mice. (B) WT ovary with advanced antral follicle (AnF) as well as primary follicles
(PrF). (C) Sohlh1 +/- ovary (-/-) lacks germ cells. (D) Five pairs of newborn ovaries
from WT and Sohlh1 knockout (-/-) mice were sectioned, and oocytes within
the germ cell clusters and primordial follicles were counted. No significant dif-
fferences were observed between the WT and knockout ovaries. Data are repre-

characterized as mean values, with error bars representing the SEM. Fisher’s exact t test
was used to calculate P values. (Scale bars: 400 μm.)

Fig. 3. Sohlh1 knockout histology and immunohistochemistry. WT and

knockout (-/-) data are shown. (A and B) Newborn ovaries stained with
antibodies against GCNA1 show no difference in primordial follicles (PrF) or
germ cell cysts (GCC) between WT (A) and knockout (B). (C and D) Three-day
ovaries stained anti-MSY2. Primary follicles (PrF) are seen in WT (G) but not
knockout (D) ovaries. (E and F) Periodic acid/Schiff reagent (PAS) staining of
7-day WT (E) and knockout (F) ovaries show fewer follicle types and empty
follicles (EF) in the knockout (F). (G and H) PAS staining of 3-week ovaries
shows no remaining oocytes in the knockout (H) but all stages of development
in the WT (G). (Scale bars: A–G, 50 μm; H, 400 μm.)

Pangas et al.
diplotene/dictyate stage of the first meiotic division (11). Oocytes in primordial follicles and germ cell cysts from WT and Sohlh1−/− newborn ovaries stained similarly with anti-GCNA1 antibodies (Fig. S4 and B). These data indicate that embryonic germ cell migration and proliferation is grossly normal in newborn Sohlh1−/− females. MSY2 is a cytoplasmic marker for oocytes that have entered the diplotene stage and persists in dictyate stages (12, 13). By postnatal day 3, MSY2 immunoreactivity is present in oocytes of primary follicles of WT ovaries (Fig. S3C). In contrast, Sohlh1−/− ovaries lack primary follicles on postnatal day 3 (Fig. S3D), and anti-MSY2 staining is present only in oocytes of primordial follicles. Thus, there appears to be a defect in follicle development during the primordial-to-primary follicle transition.

Secondary follicles develop in WT ovaries by postnatal day 7 (Fig. S3E), but, by this time, Sohlh1−/− ovaries are significantly smaller than WT, and few oocytes are observed (Fig. S3F). Most oocytes in the 7-day-old Sohlh1−/− ovaries are still enveloped by flat somatic cells similar to primordial follicles but now also contain multiple empty follicles in the central portion of the ovary (Fig. S3F). By 3 weeks of age, Sohlh1−/− ovaries contained few germ cells (Fig. S3H), although a secondary follicle was occasionally seen, and Sohlh1−/− ovaries beyond 7 weeks lacked by histology germ cells and follicular structures. Thus, the time frame for oocyte and follicle loss in Sohlh1−/− mutant females is reminiscent of Figla−/− and Noobox−/− ovaries (3, 7), which also exhibit early postnatal oocyte loss. Similar to Noobox−/− ovaries, Sohlh1−/− ovaries do not misexpress meiotic genes Mlh1 and Msh5 or apoptosis genes Bax, Bcl2, Casp2, and Bcl2L2 (data not shown). Figla, Noobox, and Sohlh1 oocyte-specific pathways may therefore overlap.

We reported previously that the Noobox deficiency did not affect expression of Figla and FIGLA’s presumed targets, the zona pellucida genes Zp1, Zp2, and Zp3 (3). However, Sohlh1−/− ovaries contain significantly lower amounts of Figla transcripts, as shown by in situ hybridization (Fig. S4A–D) and quantitative RT-PCR (Fig. S10, which is published as supporting information on the PNAS web site). Ovaries that lack Figla form very few primordial follicles (7). Sohlh1−/− ovarian pathology is less severe than reported for the Figla knockouts, and persistent low levels of Figla expression in Sohlh1−/− animals may account for this difference in pathology.

By sequence analysis, the Figla promoter保守了E box elements and therefore is not likely a direct transcriptional target of Sohlh1. FIGLA’s target genes, Zp1 and Zp3, are drastically down-regulated in Sohlh1−/− ovaries (Figs. S4 E–H and M–P and 10) as compared with Zp2 expression (Figs. S1 I–L and 10). It is possible that SOHLH1 positively cooperates with FIGLA in the transcriptional regulation of Zp1 and Zp3 but not Zp2 and that, in the absence of SOHLH1, FIGLA is insufficient to activate transcription of Zp1 and Zp3. Alternatively, the 4-fold reduction in Figla expression may be more detrimental for transcription of Zp1 and Zp3 as compared with Zp2. Studies have shown that FIGLA-binding sites in the Zp2 promoter differ from the Zp1 and Zp3 promoter sites (8) and that FIGLA transactivates the Zp2 promoter 2-fold higher than the promoters of either Zp1 or Zp3 (8).

Noobox transcripts are also reduced ~4-fold in Sohlh1−/− ovaries (Fig. S11, which is published as supporting information on the PNAS web site), whereas Sohlh1 transcripts (AW554400) are not significantly affected in Noobox−/− ovaries (3). Oocyte-specific genes that are down-regulated in Noobox−/− ovaries, such as Gdf9, Pou5f1, Zar1, Mos, and Hif1α, are also down-regulated in Sohlh1−/− ovaries, consistent with the reduction of Noobox (Fig. S11 B, C, and G) and confirming that the NOBOX pathway is compromised in Sohlh1−/− ovaries. Therefore, NOBOX functions downstream of SOHLH1, and loss of the NOBOX pathway likely contributes to the Sohlh1−/− phenotype. Not all germ cell-specific genes are down-regulated; Noahma transcript levels are not significantly different in WT and Sohlh1−/− newborn ovaries (Fig. S11D). Stra8, an early molecular marker for female germ cell differentiation (14), disappears in the WT at ~E16.5, at the time when meiosis is ongoing in female germ cells. Sohlh1−/− ovaries may not be functional primordial oocytes because of misexpression of multiple oocyte-specific genes.

Because the newborn Sohlh1−/− ovarian histology was grossly similar to WT but demonstrated molecular defects, we chose this...
Preferentially expressed in testes and ovaries (Fig. 5) was drastically down-regulated in Lhx8 and Lhx8 riboprobe to WT newborn (E and F). Lhx8 encodes a LIM homeodomain protein (15–17), and because the Lhx8 role is critical in oocyte development. Our data indicate that SOHLH1 acts upstream of Lhx8. Figla, and Nobox. Of these, Lhx8 is a direct target gene, as are Zp1 and Zp3. In the embryonic period, the SOHLH1 and Lhx8 mRNA expression patterns in oocytes are similar, consistent with Lhx8 regulation by SOHLH1. However, SOHLH1 expression postnatally is confined to oocytes of follicles up to the primary follicle stage. This temporally restricted expression pattern is in stark contrast to the postnatal and adult expression patterns of Lhx8, Figla, Zp1–3, and Nobox, which are maintained in oocytes of more advanced follicles. Thus, differentiation of primordial oocytes may depend on the transient expression of SOHLH1 to direct expression of Lhx8 and perhaps others, which in turn regulate Nobox and Figla (Fig. 7). These transcription factors may then serve as self-sustaining determinants of oocyte-specific gene expression throughout the remainder of folliculogenesis. In total, our data indicate that SOHLH1 is an integral part of a genetic program that is required for germ cell-specific gene expression.

**Materials and Methods**

**Targeting Construct.** SOHLH1 genomic clones were isolated from a 129S6/SvEv genomic library, and a targeting construct was generated (10). The SOHLH1 targeting construct was electroporated into AB2.2 ES cells to mutate the WT SOHLH1 locus by homologous recombination as described in ref. 19. The mutant SOHLH1 allele replaces exons 2–8 with the Pgk1-HPRT cassette.
Breeding, Histology, Histomorphometric Analysis, and Immunohistochemistry. All mouse experiments were carried out on a C57BL/6J 129S5/SvEvBrd hybrid background. Litters were weaned at 3 weeks, and breeding pairs were set up at 6 weeks of age. One mating pair was placed per cage and inspected every morning for the presence of litters. For histological analysis, ovaries were placed in 10% buffered formalin, processed, embedded in paraffin, serially sectioned (5 µm), and stained with hematoxylin and eosin or with periodic acid/Schiff reagent and hematoxylin. Five pairs of ovaries of each genotype were subjected to gross and microscopic analysis for each time point. For histomorphometric analysis, every fifth section was derived from the long axis of the ovary and photographed, and oocytes containing nuclei were scored. The total numbers of oocytes from all of the sections were summed, and the mean number per ovary was determined. No correction factor was used. Fisher’s exact test was used to calculate P values. Adult ovarian volume was calculated by using the ellipse formula, A × B × C = 0.5233, where A is the long axis, B is the largest transverse diameter, and C is the largest transverse diameter.

Germ cell cysts were defined as two or more oocytes that were not individually separated by stromal cells. Primordial follicles were defined as small oocytes (<20 µm) surrounded by flat epithelial cells. Primary follicles were defined as having larger oocytes (>20 µm) surrounded by a single layer of cuboidal granulosa cells; secondary follicles were defined as larger oocytes surrounded by two or more layers of granulosa cells.

For immunohistochemistry, we used antibodies against GCNA1, MSY2, and SOHLH1 proteins. Anti-GCNA1 rat monoclonal antibody (11) was kindly provided by George C. Enders (University of Pennsylvania, Philadelphia). Polyclonal rabbit antibodies against SOHLH1 (COOH terminus; amino acids 121–357) were generated by using the pET-23 system (Novagen) and immunizing goats at Cocalico Biologicals (Reamstown, PA). The anti-SOHLH1 antibodies were immunopurified and used in immunohistochemistry as described in ref. 20.
Microarray Analysis. Total RNA isolated from WT and Sohlh1<sup>−/−</sup> ovaries was used for oligonucleotide microarray analysis by using mouse genome 430 2.0 Array (Affymetrix, Santa Clara, CA) at the Baylor College of Medicine Microarray Core Facility with established protocols. Data were analyzed by using GENESPRING software (Silicon Graphics, Mountain View, CA).

RNA Isolation, RT-PCR, Quantitative Real-Time PCR, and In Situ Hybridization. Multissite RT-PCR was performed as described in ref. 21. Oligonucleotides corresponding to Sohlh1, Lhx8, Gdf9, Pou5f1, and Rpfl4 were selected by using Primer 3 software to generate an ~500-bp nucleotide fragment that is interrupted by an intron within the mouse genome. The sequences of these primers are available on request from A.R. Mouse actin-specific primers were used to verify cDNA synthesis from RNA isolated from each tissue. PCR was carried out for 28 cycles on three independently collected pools of newborn ovaries.

Quantitative real-time PCR was performed on the Prism 7500 Sequence Detection System (Applied Biosystems) by using Assays-On-Demand PCR primer (Applied Biosystems) and probe sets for each gene and mouse Gapd (VIC-labeled MGD probe, primer limited, Applied Biosystems) as the endogenous control. RT-PCR was performed by using the TaqMan Universal PCR Master Mix (Applied Biosystems) in 20 μl. Each sample was analyzed in duplicate on at least three independent newborn WT and Sohlh1<sup>−/−</sup> cDNA samples. Two nontemplate control (RNase-free water) samples were included on each plate for each primer–probe set. The relative amount of transcript was calculated by the ΔΔCT method as described by Applied Biosystems by using 7500 SYSTEM 1.2.3 software (Applied Biosystems) and normalized to the endogenous reference (Gapd). One WT sample was randomly chosen to serve as the reference sample, to which all other samples were normalized. The average and standard error was calculated for the triplicate measurements, and the relative amount of target gene expression for each sample was plotted. Significance was performed by using Student’s t test with EXCEL (Microsoft).

Mouse cDNA fragments corresponding to Sohlh1, Lhx8, Zp1, Zp2, Zp3, and Figla were subcloned into pGEM-T Easy vectors (Promega) and used to generate anti-sense and sense strands by labeling with [α-35S]UTP using the Riboprobe T7/SP6 Combination System (Promega) (22). In situ hybridization was carried out as described in ref. 21. In situ hybridizations were performed on ovarian sections derived from two different animals.

Plasmids. Luciferase reporter vectors carrying mouse partial promoter sequences were constructed by introducing the PCR-amplified promoter into the vector of pGL4 (Promega). PCR amplification was performed by using oligonucleotides pLhx8-1 and pLhx8-2 for Lhx8, pNobox-1 and pNobox-2 for Nobox, pZp1-1 and pZp1-2 for Zp1, Zp2-1 and pZp2-2 for Zp2, and Zp3-1 and pZp3-2 for Zp3 (Table 1, which is published as supporting information on the PNAS web site). To generate mutant E boxes, we used the QuikChange Multi Site-Directed Mutagenesis Kit (Stratagene) with the oligonucleotides listed in Table 1. An expression vector for mouse Sohlh1 was constructed by cloning the full-length Sohlh1 cDNA into pcDNA3 (Invitrogen).

Cell Culture and Reporter Assays. Human embryonic kidney cells (HEK293) were grown in DMEM with 10% FCS. For transient transfection, FuGENE6 (Roche Applied Science, Indianapolis) was used according to the manufacturer’s instructions. After transfection, cells were cultured for 48 h before harvest. For each transfection, 200 ng of reporter construct, 200 ng of the indicated expression plasmid, and 20 ng of prfRT-TK normalization plasmid (Promega) were used per well in a 12-well plate. Dual luciferase assays were carried out with total cell extracts as recommended by Promega. All transfection experiments were performed in triplicate, and results were normalized to the expression of Renilla luciferase. Data are relative to the mock transfection of empty parent vector, pGL4. Statistical analysis was performed by using one-way ANOVA followed by the Tukey–Kramer honestly significant difference test for multiple comparisons (JMP 5.1; JMP, Cary, NC). P < 0.05 was considered statistically significant.

ChIP. Mouse newborn ovaries and liver were collected and fixed directly in formaldehyde. The fixed ovaries were washed, homogenized, and prepared for immunoprecipitation by using a modified protocol of the ChIP assay kit (Upstate Biotechnology, Lake Placid, NY). The samples were precleared with salmon sperm DNA/protein A agarose (Upstate Biotechnology) and incubated with affinity-purified anti-SOHLH1 antibody or purified IgG at 4°C overnight. Chromatin samples immunoprecipitated by salmon sperm DNA/protein A agarose were washed two times with 1 ml of low-salt immune complex wash buffer (0.1% SDS/1% Triton X-100/2 mM EDTA/20 mM Tris-HCl, pH 8.1/150 mM NaCl), once with 1 ml of high-salt immune complex wash buffer (0.1% SDS/1% Triton X-100/2 mM EDTA/20 mM Tris-HCl, pH 8.1/500 mM NaCl), once with 1 ml of LiCl immune complex wash buffer (0.25 M LiCl/Igepal-CA630/1% sodium deoxycholate/1 mM EDTA/10 mM Tris-HCl, pH 8.1), and with two times 1 ml of TE buffer (10 mM Tris-HCl, pH 8.0/1 mM EDTA). Immune complexes were eluted from the antibody by adding 500 μl of elution buffer (0.1% SDS, 0.1M NaHCO<sub>3</sub>). The complex-DNA cross-links were reversed by adding 20 μl of M NaCl and heating at 65°C for 4 h and then adding 10 μl of 0.5 M EDTA, 20 μl of 1 M Tris-HCl (pH 6.5), and 2 μl of proteinase K (10 mg/ml) for 1 h at 45°C. DNA was recovered by phenol/chloroform extraction and ethanol precipitation, and resuspended in TE buffer. The supernatant of an immunoprecipitation reaction performed in the absence of the anti-SOHLH1 antibody was purified and used as a control. PCR analysis used primers from promoter regions of Lhx8, Nobox, Zp1, Zp2, or Zp3, shown in Table 2, which is published as supporting information on the PNAS web site.

This work was supported National Institutes of Health Grant HD44858 and March of Dimes Basil O'Connor Award 5-FY02-206 (to A.R.), National Institutes of Health Grant HD42500 (to M.M.M.), and National Institutes of Health/National Research Service Award 5F32 HD46335-01A1 (to S.A.P.).