HHS Public Access

Author manuscript

FASEB J. Author manuscript; available in PMC 2023 April 25.

Published in final edited form as:

FASEB J. 2023 February ; 37(2): e22747. doi:10.1096/fj.202201481R.

Sumoylation regulates functional properties of the oocyte transcription factors SOHLH1 and NOBOX

Bethany K. Patton[1,2], **Surabhi Madadi**[1,3], **Shawn M. Briley**[1,4], **Avery A. Ahmed**[1,5], **Stephanie A. Pangas**[1,2,4,5,6,*]

[1]Department of Pathology & Immunology, Baylor College of Medicine, Houston, TX 77030

[2]Graduate Program in Molecular & Cellular Biology, Baylor College of Medicine, Houston, TX 77030

[3]Rice University, Houston, TX 77005

[4]Graduate Program in Biochemistry & Molecular Biology, Baylor College of Medicine, Houston, TX 77030

[5]Graduate Program in Development, Disease Models & Therapeutics, Baylor College of Medicine, Houston, TX 77030

[6]Department of Molecular & Cellular Biology, Baylor College of Medicine, Houston, TX 77030

Abstract

SOHLH1 and NOBOX are oocyte-expressed transcription factors with critical roles in ovary development and fertility. In mice, *Sohlh1* and *Nobox* are essential for fertility through their regulation of the oocyte transcriptional network and cross-talk to somatic cells. Sumoylation is a post-translational modification that regulates transcription factor function, and we previously showed that mouse oocytes deficient for sumoylation had an altered transcriptional landscape that included significant changes in NOBOX target genes. Here we show that mouse SOHLH1 is modified by SUMO2/3 at lysine 345 and mutation of this residue alters SOHLH1 nuclear to cytoplasmic localization. In NOBOX, we identify a non-consensus SUMO site, K97, that eliminates NOBOX mono-SUMO2/3 conjugation, while a point mutation at K125 had no effect on NOBOX sumoylation. However, $NOBOX^{K97R/K125R}$ double mutants showed loss of mono-SUMO2/3 and altered higher molecular weight modifications, suggesting cooperation between these lysine's. $NOBOX^{K97R}$ and $NOBOX^{K97R/K125R}$ differentially regulated NOBOX promoter targets, with increased activity on the *Gdf9* promoter, but no effect on the *Pou5f1* promoter. These data implicate sumoylation as a novel regulatory mechanism for SOHLH1 and NOBOX,

[*]**Corresponding Author**: Stephanie A. Pangas, PhD, One Baylor Plaza, Baylor College of Medicine, Houston, TX 77030, spangas@bcm.edu.

Author Contributions: B.K.P and S.A.P designed the study, B.K.P, S.A.P, S.M., S.M.B, A.A.A performed experiments. B.K.P, A.A.A. and S.A.P analyzed the data. B.K.P and S.A.P wrote the manuscript with input from all authors.

Conflict of Interest: The authors declare no conflict of interests

Animal Studies: Experimental animals were used in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals with prior approval of the Institutional Animal Care and Use Committee at Baylor College of Medicine under animal protocol AN-4762.

which may prove useful in refining their roles during oogenesis as well as their function during reprogramming to generate de novo germ cells.

Keywords

PTM; oocyte; sumoylation; transcription factor

Introduction

Spermatogenesis and oogenesis basic helix loop helix 1 (SOHLH1) is a transcription factor expressed in germ cells during embryonic development and in oocytes of primordial and primary follicles[1]. NOBOX is an oocyte-specific homeobox transcription factor expressed in oocytes of primordial follicles with continued expression throughout folliculogenesis until oocytes are fully grown, and is thought to be regulated by SOHLH1[2,3]. Both transcription factors are essential for folliculogenesis, and homozygous deletion of either gene in mice results in female sterility with a rapid postnatal depletion of the ovarian reserve[1,2]. Furthermore, both genes are implicated in reproductive disease in women with *NOBOX* recognized as one of the most highly mutated genes in women diagnosed with premature ovarian insufficiency (POI)[4-6]. *SOHLH1* variants have also been identified in POI patients but less frequently than *NOBOX*[7-9]. More recently, SOHLH1 and NOBOX were included in a set of transcription factors for reprogramming stem cells into oocyte-like cells (OLC)[10]. While SOHLH1 and NOBOX are essential for oocyte development, and potentially useful for therapeutics and technological advancements, nothing is known about their post-translational regulation.

Sohlh1 or *Nobox* homozygous null female mice are sterile, exhibit blocks in early ovarian folliculogenesis, and lose all oocytes by postnatal day 21 and 14, respectively[1,2,11]. Key oocyte-expressed genes downregulated in *Sohlh1*$^{-/-}$ ovaries include *Nobox*, factor in germline alpha (*Figla*), and LIM homeodomain 8 (*Lhx8*), while *Nobox*$^{-/-}$ ovaries lose expression of key genes such as growth and differentiation factor 9 (*Gdf9*), bone morphogenetic protein 15 (*Bmp15*), and POU domain class 5 transcription factor 1 (*Pou5f1*; also called *Oct4*)[2]. In developing ovaries, embryonic and newborn SOHLH1 localizes to the cytoplasm and nucleus, and SOHLH1 is essential for the proper localization of its paralog, spermatogenesis and oogenesis basic helix loop helix 2 (SOHLH2)[12]. The DNA binding domain of NOBOX has been characterized and NOBOX directly binds the *Gdf9* and *Pou5f1* promoters[13]. While SOHLH1 and NOBOX directly or indirectly mediate transcription of a large subsets of oocyte genes, it is unknown if there are additional regulatory inputs, such as post-translational modifications that allow them to differentially regulate transcriptional programs temporally during oocyte development and folliculogenesis.

Sumoylation is a ubiquitin-like post-translational modification (PTM) in which a small ubiquitin modifier (SUMO) is attached to a lysine within a consensus sequence via an E1-E2-E3 enzyme cascade[14,15]. SUMO has three isoforms expressed in oocytes: SUMO1, SUMO2, and SUMO3[16]. SUMO2 and SUMO3 are indistinguishable from one another and are collectively referred to as SUMO2/3[15]. Sumoylation is highly dynamic and removed

from the substrate by a SUMO specific protease[17]. Female mice with an oocyte-specific deletion of the single essential E2 ligase, ubiquitin-conjugating enzyme E2I (*Ube2i;* also called *Ubc9*), are sterile with defective folliculogenesis and premature depletion of the ovarian reserve[18]. In ovaries of *Ube2i* oocyte-conditional knockout mice (referred to as *Ube2i* cKO), *Nobox* mRNA increases without any change in its protein levels, while several NOBOX target genes such as *Gdf9* and *Pou5f1* are downregulated, suggesting NOBOX is less active in sumoylation-null oocytes. NOBOX is modified by SUMO1 and SUMO2/3 *in vitro*, but the functional significance remains to be determined. While SOHLH1 contains a single consensus sumoylation site, whether SOHLH1 is sumoylated is not known, nor is it known if sumoylation alters SOHLH1 activity[18].

The goals of this study were to determine if SOHLH1 is directly modified by sumoylation, identify the sumoylation sites, and determine if there are functional consequences for SOHLH1 and NOBOX sumoylation. Through *in vitro* approaches and mouse models, we show that both SOHLH1 and NOBOX are sumoylated with SUMO2/3, and sumoylation has differential functional effects on these critical transcription factors.

Materials and Methods

Experimental Mice

Experimental animals were used in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and this study was approved by the Institutional Animal Care and Use Committee at Baylor College of Medicine (animal protocol AN-4762). The $Nobox^{G403A}$ knockin allele that mutated K125 to arginine (R) was generated at the Embryonic Stem Cell and Genetically Engineered Mouse Cores (now called the Genetically Engineered Rat and Mouse Core) at Baylor College of Medicine. A point mutation in *Nobox* was induced using a single sgRNA designed to target Cas9 to the genomic region encoding the SUMO consensus sequence creating a single base pair change from G to A at amino acid 403. An additional silent, EcoRI digest site was added for genotyping purposes. 47 pups were born and 40 contained the EcoRI restriction digest site. Samples were sent for sequencing to confirm the presence of the mutation. One founder was selected and crossed to F1 mixed hybrid strain *C57BL/6/129S7/SvEvBrd*, which is the wild type genetic background of our mouse colony and that of the original *Nobox* null allele. Backcrossing was performed to ensure germline transmission of the mutation and remove off-target mutations identified in the initial sequencing. After two generations of back cross, pups were sequenced and found to have the correct mutation and reduce the probability of mice carrying off-target mutations. Subsequent genotypes were performed by PCR (forward primer 5'-CTGCCTGGTCTTCCCTTCAG-3', reverse primer 5'AGAATCGCTGAAGCCATCCA-3') and validated by genomic DNA sequencing for each offspring from ear punches. Wild type and experimental mice $Nobox^{G403A/G403A}$ were generated from heterozygous crosses of $Nobox^{G403A/+}$.

Fertility Studies

Continuous breeding pairs were established between sexually mature (6-week-old) wild type or $Nobox^{G403A/G403A}$ female mice with sexually mature, 8-week old wild type males (F1

C57BL/6;129S7/SvEvBrd) and housed for seven months. The number and date of newborn pups was monitored for each cage with litters weaned at 21 days.

Histologic Analysis

Mice were weighed, anesthetized by isoflurane (Abbott Laboratories, Abbott Park, IL), and euthanized by cervical dislocation. Ovaries were fixed in 10% neutral buffered formalin (Electron Microscopy Sciences, Hatfield, PA) and processed and embedded at the Baylor College of Medicine Human Tissue Acquisition and Pathology Core. 6-week old ovaries were sectioned at 5-μm and representative sections from the center of each ovary were stained. Slides were stained in periodic acid-Schiff (Sigma, St. Louis, MO). Histology was analyzed by light microscopy using a digital camera, AxioCam 105 (Zeiss, Oberkochen, Germany), at 5x and 40x magnification.

Cell Culture and Plasmids

HEK-293T cells were purchased from ATCC and cultured in Dulbecco's modified Eagle's medium (Invitrogen, Waltham, MA) supplemented with 10% fetal bovine serum (Gibco, Waltham, MA). Expression plasmids for Myc-SUMO1ρ and Myc-SUMO2ρ, which are mostly resistant to desumoylation, were kindly provided by Dr. Deborah Johnson (Baylor College of Medicine, Houston, TX). Full-length FLAG-tagged mouse *Nobox* (provided by Aleksandar Rajkovic, University of California, San Francisco) was cloned into pCMV-Tag2A and verified by DNA sequencing. Mouse Myc-DDK-tagged YY1 was purchased from Origene (MR206531, Rockville, MD) and subcloned into pcDNA3.1(+) to remove the DDK tag. Plasmids for transfection were prepared using a maxiprep kit (Qiagen, Hilden, Germany) and verified by DNA sequencing (Genewiz, South Plainfield, NJ). Mutagenesis was performed using KOD DNA polymerase (Sigma-Aldrich, St. Louis, MO) and mutation specific primers (Supporting Table S1). Mutagenized plasmids were transformed into DH5α cells for preparation and sequencing as described above. The *Nobox-Luc* construct was made by cloning 2KB upstream of the murine *Nobox* start codon in front of the luciferase gene through the cloning service, Vector Builder (Santa Clara, CA).

Transient Transfection and Immunoprecipitation

HEK-293T cells were grown to 80% confluence, then transfected with JetPrime reagent following the manufacturer's recommendations (Polypus, Illkirch-Graffenstaden, France). Cells were grown for 48 h then harvested in 1 mL of Mammalian Protein Extraction Reagent (M-PER) (Thermo Fisher Scientific, Waltham, MA) with protease inhibitor cocktail for general use (Sigma-Aldrich, St. Louis, MO). Cell lysate (2 mg) was immunoprecipitated using 40 μL of anti-FLAG-M2 beads (Sigma-Aldrich, St. Louis, MO). The lysate/bead mixture was incubated overnight at 4°C. Beads were washed three times in TBS and eluted in 100 μL of 1× LDS sample buffer with β-mercaptoethanol. For input lanes, 10 μg of protein was loaded with 1× LDS sample buffer (Invitrogen, Waltham, MA) and sample reducing reagent (Invitrogen, Waltham, MA) and processed as above.

Immunoblotting

Samples were prepared by loading 20uL of IP eluant or 10 μg input protein with LDS sample buffer (Invitrogen, Waltham, MA) and sample reducing agent (Invitrogen, Waltham, MA), then incubated at 70°C for 10 min. Samples were cooled to room temperature and loaded into NuPAGE 4-12%bis-tris protein gels (Invitrogen, Waltham, MA) and run at 200 V for 50 minutes in MOPS SDS running buffer (Invitrogen, Waltham, MA) under reducing conditions. Proteins were transferred to nitrocellulose membranes (Thermo Fisher Scientific, Waltham, MA) in NuPAGE transfer buffer (Invitrogen, Waltham, MA) with 10% methanol (VWR, Radnor, PA) at 30 V for 1 hour. After transfer, membranes were allowed to dry. Non-specific antigen blocking was carried out by incubating membranes in 5% non-fat milk in TBS-T for at least 1 hour. Primary antibodies (Rabbit Anti-Myc Tag 1:1000, 2278, Cell Signaling, Danvers, MA) were diluted in 5% milk in TBS-T and incubated on with rocking at 4°C overnight. Membranes were washed three times with TBS-T for 10 min. Secondary antibodies (HRP-conjugated donkey anti-rabbit, 1:10,000 Jackson ImmunoResearch, West Grove, PA) were diluted in 5% milk in TBS-T and incubated on the membrane for 1 hour at room temperature with rocking, and then washed in TBS-T as described above. Chemiluminescent detection was performed with SuperSignal West PicoPLUS chemiluminescent substrate (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's directions. Signal detection was performed on HyBlot CL autoradiography film (Thomas Scientific, Swedesboro, NJ) developed using a Konica SRX-101A medical film processor (Konica Minolta Medical and Graphic, Wayne, NJ). Blots were stripped for 15 minutes in stripping buffer (Thermo Fisher Scientific, Waltham, MA) and reprobed using rabbit anti-FLAG (1:1000, F7425 Sigma-Aldrich, St. Louis, MO) and HRP-conjugated donkey anti-rabbit secondary (1:10,000 Jackson Immunoresearch, West Grove, PA)

Immunofluorescent Microscopy

Tissue sections were antigen-retrieved in 0.01 M citric acid and 0.1% Triton X (Sigma, St. Louis, MO), blocked with 3% BSA in Tris-buffered saline, 0.1% Tween 20 (TBS-T), then incubated overnight at 4°C with either rabbit anti-FLAG (1:250, Sigma Aldrich, St. Louis, MO) or rabbit anti-SOHLH1 (1:100, a gift from Dr Aleksander Rajkovic, University of California, San Francisco, USA). Slides were washed in TBS-T, incubated at room temperature with goat anti-rabbit Alexa Fluor 594 (1:250, A32740, Invitrogen, Waltham, MA) for 1 hour, washed, incubated in 4'6'-diamidino-2-phenylindole (DAPI) (1:1000) in TBS for 5 minutes and mounted in ProLong Diamond antifade mount (Thermo Fisher Scientific, Waltham, MA). Fluorescent images were captured using a Zeiss LSM 880 with Airyscan FAST confocal microscope at the BCM Optical Imaging and Vital Microscopy Core and processed with the Zeiss Zen Blue software (Zeiss, Oberkochen, Germany). Exposure times were held constant between wild type and mutant samples. Representative populations of transfected HEK-293T cells were imaged for FLAG-NOBOX slides. Representative follicles within ovary sections for SOHLH1 slides were analyzed using ImageJ software (ImageJ 1.52a Wayne Rasband, National Institutes of Health, USA http://imagej.nih.gov/ij) to measure the mean fluorescence intensity of the total area of the oocyte. Nuclear intensity was measured by quantifying the intensity of the oocyte that overlapped with the DAPI stain and cytoplasmic intensity was calculated by subtracting

nuclear fluorescent intensity from total oocyte fluorescent intensity. Only oocytes with visible nuclei were measured. Follicles were classified as described.

Luciferase Assay

HEK-293T cells were transfected with the appropriate plasmids in 48-well plates using Lipofectamine 2000 per manufacturer's instructions (Thermo Fisher Scientific, Waltham MA). 2ng of renilla expression plasmid, pRL-TK (Promega, Madison, WI), was transfected in all wells as an internal normalization control. 48 hours after transfection cells were harvested using the Dual Luciferase kit from following the manufacturer's protocol (BPS Bioscience, City, State). Measurements were made using a Tecan Infinite M1000 Pro plate reader (Tecan, Männedorf, Switzerland) The Renilla reagent was then added to all wells and an additional reading was taken. Results are reported as a fold change of the luciferase to renilla ratio normalized to NOBOX alone wells.

Statistical Analyses

GraphPad Prism 5 (GraphPad Software, La Jolla, CA) was used for statistical analysis. Two-tailed unpaired Student's *t*-test or the non-parametric Mann-Whitney U was used for single comparisons. One-way analysis of variance (ANOVA) with post-hoc analysis by Fisher's Least Significant Difference test or Bonferroni's Multiple Comparison test was used for multiple comparisons. For data presented as proportions, the average of each portion was calculated, and Fisher's Exact test was used to compare proportions. A power analysis was performed for all experimental methods, and sample sizes are indicated in the text and figure legends; a minimum of three independent experiments was carried out at all times, with $P<0.05$ considered statistically significant.

Results

SOHLH1 is sumoylated by SUMO2/3

Our previous study showed that while *Nobox* mRNA was upregulated in *Ube2i* cKO ovaries compared to controls, the known NOBOX downstream transcriptional targets were downregulated even though NOBOX protein levels were unchanged[18]. We hypothesized that loss of sumoylation potentially caused overactivation of SOHLH1, because sumoylation often results in transcriptional repression[19-22]. First, we determined if SOHLH1 was a direct target of sumoylation. Using SUMO-GPS software[23,24], we identified a single predicted SUMO consensus site ("ψ-K-x-E") on mouse SOHLH1 lysine 345; though not directly conserved, human SOHLH1 also contains a single predicted SUMO consensus lysine near the carboxyl-terminus. We then generated a K345R mutation of SOHLH1 expression plasmid ("3xFLAG-SOHLH1^{K345R}") by site-directed mutagenesis using a previously constructed 3x-FLAG tagged SOHLH1[25]. SOHLH1 expression constructs were transiently transfected with non-deconjugatable Myc-tagged SUMO1 or SUMO2/3 (termed "Myc-SUMO1ρ" or "Myc-SUMO2/3ρ")[18,26] into HEK-293T cells, cell lysates were immunoprecipitated against FLAG, then immunoblotted for Myc. Immunoblotting of IP lysates with an anti-SOHLH1 antibody confirmed that the FLAG tag protein corresponded to SOHLH1 (Supporting Fig S1A).

IP-immunoblotting of lysates from cells co-transfected with wild type 3x-FLAG-SOHLH1 and Myc-SUMO2/3ρ showed a 65 kDA molecular weight (MW) band corresponding to mono-SUMO2/3 modified SOHLH1 (Fig. 1A, right panel, arrow). Higher molecular weight forms of SUMO2/3 modified SOHLH1 were also identified. In contrast, no similar sized MW band that would correspond to mono-SUMO1 conjugation was detectable when 3x-FLAG-SOHLH1 was co-transfected with Myc-SUMO1ρ; however, higher molecular weight forms were observed. Co-transfection of 3xFLAG-SOHLH1^{K345R} with Myc-SUMO2/3ρ almost fully eliminated the 65kDA band (Fig 1B, right panel, arrow), while co-transfection of 3xFLAG-SOHLH1^{K345R} with Myc-SUMO1ρ did not alter the banding pattern by immunoblotting. These data indicate that consensus sumoylation site at K345 in SOHLH1 is utilized by SUMO2/3 but not SUMO1.

Loss of K345-SUMO2/3 on SOHLH1 leads to increased nuclear localization.

Sumoylation has target-specific functional effects that include altering subcellular localization[15], and SOHLH1 is known to shuttle from the nucleus to the cytoplasm[12]. We first validated if loss of sumoylation in oocytes causes changes to SOHLH1 localization in *Ube2i* cKO ovaries compared to littermate controls. Histology sections of 14-day old ovaries were analyzed by confocal microscopy by indirect immunofluorescence to SOHLH1. Total fluorescent signal and nuclear fluorescent signal for individual oocytes was quantified using ImageJ. Nuclear signal was subtracted from the total signal to obtain the cytoplasmic intensity. Only oocytes with visible nuclei (identified by DAPI) were quantified. In control ovaries, SOHLH1 immunoreactivity was found predominantly in the cytoplasm of oocytes in both primordial (Prf) and primary follicles (Pf) (Fig. 2A, upper row). However, in *Ube2i* cKO ovaries, SOHLH1 distribution was significantly greater in the oocyte nucleus for both the primordial and primary follicles (Fig 2A,B). While localization differed in the knockout, the known downregulation of SOHLH1 in oocytes of secondary follicles was similar in *Ube2i* cKO and controls (Fig 2A).

To directly confirm if sumoylation affects SOHLH1 localization, HEK-293T cells were transiently transfected with the 3xFLAG-SOHLH1 or 3xFLAG-SOHLH1^{K345R} expression constructs, and then cytoplasmic and nuclear fractions were isolated. Relative proportions of SOHLH1 in the cytoplasm and nucleus were quantified via immunoblotting and measuring the band intensity of each fraction. The intensity of the nuclear and cytoplasmic fraction were then added together to calculate the total intensity of FLAG protein in each cell lysate. Nuclear and cytoplasmic proportions were then calculated by dividing the intensity of the band for each individual fraction by the total intensity (Fig 2C). The 3xFLAG-SOHLH1^{K345R} mutant showed significantly greater levels in the nuclear fraction as compared to the wild type (Fig 2D). As this site is sumoylated by SUMO2/3 (Fig. 1B), these data show that sumoylation of SOHLH1 by SUMO2/3 is a mechanism for regulating its nuclear export. In addition, based on histologic analysis of *Ube2i* cKO ovaries, it is unlikely that sumoylation plays a role in the downregulation of SOHLH1 in oocytes of secondary follicles.

SOHLH1 regulates the transcription of the *Nobox* gene

After confirming that sumoylation regulates SOHLH1 localization, we sought to determine if this increased SOHLH1 activity specifically on the NOBOX promoter. Oocytes lacking SOHLH1 lose *Nobox* transcripts[1], but it is not known if SOHLH1 directly regulates the *Nobox* promoter. To test this, cloned 2KB upstream of the mouse *Nobox* transcription start site as a putative *Nobox* promoter luciferase construct. 3x-FLAG-SOHLH1 was transiently co-transfected with the *Nobox* reporter construct along with the renilla luciferase construct as an internal control. When compared to the empty vector control, expression of SOHLH1 caused a two-fold increase in luciferase activity (Fig 2E). Co-transfection of 3xFLAG-SOHLH1^{K345R} with the *Nobox* luciferase reporter showed a similar upregulation when compared to the empty vector control with WT SOHLH1. These data show that SOHLH1 is a direct regulator of the *Nobox* promoter, but sumoylation at K345 is not required for its regulation of *Nobox*.

K125 of NOBOX is dispensable for female fertility in mice

We previously showed that NOBOX was sumoylated *in vitro*[18] and similar to SOHLH1, NOBOX contains a single sumoylation consensus motif centered at K125. To determine if K125 modifications to NOBOX are required for female fertility, we used CRISPR-Cas9 gene editing to create a single base pair change from G to A at amino acid 403. Following validation of allele sequence, one founder line was selected for further study and backcrossed to a similar genetic background as the original *Nobox* null allele[2]. Homozygous pups (*Nobox*$^{G403A/G403A}$) were generated at normal Mendelian ratios. To assess reproductive defects, 6-week old wild type and *Nobox*$^{G403A/G403A}$ littermates females (n=5 per genotype) were continuously mated to 8-week old wild type males for eight months. There were no differences in the average number of pups per litter or average number of litters per month between genotypes (Fig 3B, C). There was also no difference when cumulative pups per litter with age was analyzed for a similar number of control and *Nobox*$^{G403A/G403A}$ females (n=4 females per genotype) (Fig 3A). Ovarian histology at 6-weeks of age (Fig 3E, F) and 12-weeks of age (data not shown) showed no histologic differences with all stages of follicles present in both genotypes. Thus, sumoylation at K125 of NOBOX is not required for female fertility in mice.

Mouse and human NOBOX have multiple predicted PTMs

After observing no difference in the reproductive phenotype of the *Nobox*$^{G403A/G403A}$ mouse model, we performed an *in silico* analysis of human and mouse NOBOX to identify possible non-consensus sumoylation sites and map them to key structural domains. Human and mouse NOBOX share roughly 50% amino acid sequence homology but have similar structural elements. Both proteins contain a homeodomain required for DNA binding and co-factor interactions[13] with a putative nuclear localization signal (NLS) responsible for the exclusively nuclear localization of NOBOX[27]. Additionally, the carboxyl-terminus of both proteins are proline-rich (PR) and each protein contains a putative Src-homology 3 (SH3) binding domain. The function of this domain is unknown in NOBOX but predicted to be involved in protein-protein interaction[28].

Mouse NOBOX contains 26 lysines while human NOBOX contains 32 lysines. The predictive tool, SUMO-GPS, was used to identify which lysines were most likely to be sumoylated [23,24]. A medium stringency cutoff was used in the predictions for both mouse and human proteins. This program identified the single sumoylation consensus site we had previously described at K125 in mice. Two additional non-consensus sumoylation sites were identified: one near the homeodomain at K97, and a second predicted in the carboxyl-terminus at K526 (Fig 4A). Human NOBOX contains a consensus site near the homeodomain and an additional consensus and a non-consensus site near the carboxyl-terminus (Fig 4B). Though not directly conserved, the predicted sites are situated in proximity to the same domains on each respective protein. Using a second, broad PTM prediction tool, MuSiteDeep, we additionally mapped predicted ubiquitin and acetylation sites as these are both prominent lysine modifications [29-31]. No competition between PTMs was predicted for any lysine in mouse, but human NOBOX did have a carboxyl-terminus lysine predicted to be modified by both sumoylation and ubiquitylation. In addition to predicted elements of NOBOX structure, we determined how these elements overlapped with known *NOBOX* POI variants [4-6]. Table 1 lists currently published POI variants in NOBOX that have been experimentally validated to reduce transcription in *in vitro* assays or have been predicted by multiple programs to be benign. One variant affects lysine 371 and although K371 is not predicted to be a site of post-translational modification in NOBOX, it has been shown to reduce *Gdf9-luc* reporter transcription *in vitro* [5].

K97 and K125 work cooperatively for conjugation of SUMO2/3 on NOBOX

As only 60% of validated SUMO sites fall within the consensus sequence, we wanted to determine which NOBOX lysines of the predicted non-consensus sites were sumoylated [23]. We use site-directed mutagenesis to induce point mutations (lysine to arginine) in full-length FLAG-tagged NOBOX at non-consensus sumoylation consensus sites K97 ("FLAG-NOBOX^{K97R}") and K526 ("FLAG-NOBOX^{K526R}") as well as the originally identified consensus site, K125 (FLAG-NOBOX^{K125R}"). HEK-293T cells were transiently co-transfected with wild type and mutated FLAG-tagged NOBOX expression vectors along with the non-deconjugatable Myc-tagged SUMO plasmids Myc-SUMO1ρ and Myc-SUMO2/3ρ [18,26]. Cell lysates were analyzed by co-immunoprecipitation with antibodies against FLAG and immunoblotting with antibodies to Myc-tag (IP-IB) (Fig 5A)[18]. WT FLAG-NOBOX IP lysates were also immunoblotted with an anti-NOBOX antibody to confirm the FLAG protein was NOBOX (Supporting Fig S1B).

In cells transfected with FLAG-NOBOX^{WT} without the co-expression of Myc-SUMO1ρ or Myc-SUMO2/3ρ , no NOBOX sumoylation was detected (Fig 5A). Similar to our prior publication[18], overexpression of FLAG-NOBOX^{WT} with either Myc-SUMO1ρ or Myc-SUMO2/3ρ showed both SUMO isoforms sumoylated NOBOX (Fig 5A); mono-sumoylated NOBOX is predicted to be the 85kDa MW band (Fig. 5A, arrow). Overexpression of FLAG-NOBOX^{K97R} with Myc-SUMO1ρ or Myc-SUMO2/3ρ showed loss of mono-SUMO2/3 conjugation and no change in Myc-SUMO1ρ conjugation (Fig 5B, arrow), demonstrating that the non-consensus K97 site is sumoylated by SUMO2/3 but not required for SUMO1 conjugation. Higher bands also were seen in the FLAG-NOBOX^{K97R} + Myc-SUMO2/3ρ lane, suggesting the presence of multiple PTMs or chain formation (Fig 5B). In contrast,

overexpression of FLAG-NOBOX^{K125R} with Myc-SUMO1ρ or Myc-SUMO2/3ρ showed no difference in sumoylation patterns compared to wild-type NOBOX (Fig 5C), indicating this predicted consensus site is not sumoylated, and corroborates the lack of phenotype of *Nobox*^{G403A} female mice. FLAG-NOBOX^{K526R} and an additional double mutant containing for the consensus site K125 a neighboring lysine at K126 (FLAG-NOBOX$^{K125R/K126R}$), showed bands similar to wild type (data not shown).

Loss of mono-SUMO2/3 conjugation but retention of SUMO1 conjugation in the FLAG-NOBOX^{K97R} mutant could indicate cooperativity between lysines leading to sumoylation. Therefore, we generated a double NOBOX mutant containing both the K97R and K125R mutation. Co-expression of FLAG-NOBOX$^{K97R/K125R}$ with the Myc-SUMO1ρ or the Myc-SUMO2/3ρ constructs followed by IP-IB showed loss of mono-SUMO2/3 as well as loss of the next higher molecular weight band, establishing cooperativity between the two lysines (Fig 5D). As SUMO1 conjugation remained unchanged in all mutants, additional lysines are likely modified by SUMO1 but are not predictable by current algorithms.

Localization of NOBOX is unaltered with sumoylation mutations

Our mutation analysis validated one non-consensus SUMO2/3 conjugation site on mouse NOBOX (K97). We next tested if loss of SUMO2/3 conjugation K97 had any functional consequence. We first analyzed if localization was altered in the NOBOX SUMO mutants. NOBOX contains an NLS in the homeodomain of the protein around amino acid 186-193, and it has been observed in oocytes and transfected cells that NOBOX constitutively localizes to the nucleus[2]. However, published studies of other transcription factors indicate sumoylation can regulate protein localization despite the presence of a putative NLS [32-34]. Therefore, we tested localization in HEK-293T cells transiently transfected with FLAG-NOBOX^{WT}, FLAG-NOBOX^{K97R}, or FLAG-NOBOX$^{K97R/K125R}$ and analyzed by indirect immunofluorescence using an anti-FLAG antibody. FLAG-NOBOX and both mutant NOBOX proteins localized exclusively to the nucleus (Fig 6), indicating loss of SUMO2/3 conjugation at K97 or K97/K125 has no effect on subcellular location of NOBOX.

NOBOX or its mutants do not physically interact with its proposed co-factor, YY1.

In addition to subcellular localization, sumoylation can enhance or inhibit transcription factor interaction with cofactors[35]. We next determined if loss of sumoylation could modify NOBOX interactions with other transcriptional cofactors. To date, no cofactors of NOBOX have been experimentally validated, but a few have been proposed. One potential candidate is the ubiquitous transcription factor, Yin Yang 1 (YY1). Oocyte specific loss of *Yy1* results in similar gene expression changes to that of *Nobox*$^{-/-}$ mice[36]. To test for potential interactions, HEK-293T cells were transiently transfected with FLAG-NOBOX^{WT}, FLAG-NOBOX^{K97R}, or FLAG-NOBOX$^{K97R/K125R}$, and cell lysates immunoprecipitated for FLAG. Immunoblotting with an anti-YY1 antibody showed high concentrations of endogenously-expressed YY1 in the input lanes and supernatant lanes from the IP, but no bands in any of the washes or final elution for WT NOBOX (Fig 7A). Likewise, IB-IP with FLAG-NOBOX^{K97R} and FLAG-NOBOX$^{K97/125R}$ also showed no interaction (Fig 7B). A similar experiment was conducted using a mouse Myc-tagged YY1 expression vector transiently transfected alongside FLAG-NOBOX, FLAG-NOBOX^{K97R}, or FLAG-

NOBOX$^{K97/125R}$, but similarly, no interaction was observed (data not shown). These data suggest that NOBOX and YY1 do not physically interact in an *in vitro* model system and loss of sumoylation on NOBOX does not lead to increased interaction.

Mutation of K97 and K97/125 leads to altered NOBOX transcriptional activity

The lack of altered localization or interaction with YY1 led us to investigate the effect of sumoylation mutants on NOBOX transcriptional activity with its known target promoters. HEK-293T cells were transiently co-transfected with wild-type NOBOX or one of the NOBOX sumoylation mutants along with luciferase reporter constructs containing regions of the *Pou5f1 ("Pou5f1-luc")* or *Gdf9 ("Gdf9-luc")* promoters previously shown to be directly regulated by NOBOX [13]. Results were normalized to an internal control (renilla luciferase) and are reported as a fold-change in luciferase activity relative to the empty vector control. *Pou5f1-luc* activity was similar between wild-type and mutant NOBOX expression constructs (Fig 7C). However, *Gdf9-luc* activity was significantly upregulated by both NOBOX^{K97R} and NOBOX$^{K97R/K125R}$ (Fig 7D). There was no significant difference in luciferase activity between NOBOX^{K97R} and NOBOX$^{K97/125R}$, which may indicate that the poly-modification in IP-IB experiments (Fig 5) are an artifact because of the use of non-deconjugatable SUMO plasmids and not a biologically relevant source of regulation on NOBOX. Still, loss of SUMO2/3 modification of NOBOX leads to higher transcriptional activity for a subset of NOBOX target genes solidifying sumoylation as an important regulator of NOBOX function.

Discussion

Reversable posttranslational modifications expand the functional diversity of proteins with sumoylation being a key regulatory input for a wide-ranging number of transcription factors[20,37,38]. Surprisingly little is known about the posttranslational modifications, including sumoylation that could regulate two critical oocyte transcription factors, SOHLH1 and NOBOX during ovarian follicle development. *Sohlh1* has a more restricted expression pattern in mouse ovaries than *Nobox*. *Sohlh1* is expressed in oocytes beginning at embryonic day (E) 15.5 until the it is downregulated at the secondary follicle stage, and SOHLH1 is necessary for the expression of other key oocyte transcription factors such as *Nobox* and *Lhx8*[1,12]. *Nobox* begins to be expressed in oocytes similar to *Sohlh1* (E15.5)[12], but unlike *Sohlh1*, *Nobox* continues to be expressed in oocytes until they are fully grown[11]. NOBOX directly regulates critical oocyte genes including *Gdf9* and *Pou5f1*[2,13]. Our *in silico* analysis suggests multiple PTMS may regulated NOBOX and potentially SOHLH1. Understanding how SOHLH1 and NOBOX functions are regulated by PTMs is likely essential to fully recreate the oocyte developmental program and leverage it for therapeutic and technological advances, including those that derive oocyte-like cells from stem cells or primordial germ cell-like cells[10].

We utilized overexpression assays with non-deconjugatable SUMO constructs similar to previous studies[18,26] to determine if SOHLH1 was sumoylated and which lysine was required. Co-expression of non-deconjugatable SUMO constructs is technically useful to eliminate the conjugation-deconjugation dynamics that make SUMO detection in native

proteins challenging[39]. Our *in vitro* assays indicate that mouse SOHLH1 is modified by both by SUMO1 and SUMO2/3, but molecular weight bands that correspond to mono-sumoylated SOHLH1 are detected for only SUMO2/3. Mutation of K345 in SOHLH1 was sufficient to eliminate this band, but did not affect higher molecular weight bands indicating additional lysines may be sumoylated. It is unclear if the higher molecular weight bands are artifacts due to use of the non-deconjugatable SUMOs, or if they represent biologically important chains. SUMO1 and SUMO2/3 are known to form chains which can be used for the recruitment of ubiquitin ligases[40-42]. Yet, it remains unclear whether these chains have a distinct function from mono-SUMO modification[40-44]. Additionally, immunoblotting with anti-SOHLH1 revealed a mono-modified SOHLH1 protein in the Myc-SUMO1 IP lysate not present when blotting for Myc which could indicate additional PTMs are present on SOHLH1 leading to multiple higher molecular weight bands. Despite the unknown functions of the higher weighted bands, SOHLH1^{K345R}, which lacks mono-SUMO2/3, increases the percentage of SOHLH1 localized to the nucleus. If loss of sumoylation was a critical event for SOHLH localization *in vivo*, we expected this to be recapitulated in our *Ube2i* cKO mouse model. Our analysis of *Ube2i* cKO oocytes demonstrates an increased nuclear localization of SOHLH1 compared to control littermates. These data support a hypothesis that sumoylation by SUMO2/3 is the mechanism for SOHLH1 shuttling from the nucleus to the cytoplasm in primordial and primary follicles.

Regulated subcellular localization of SOHLH1 by sumoylation is a novel discovery as previous studies on wild-type oocytes had described SOHLH1 in the cytoplasm and nucleus, but lacked a quantification of the localization and a mechanism for its shift[1,12]. SOHLH1 has been noted as a necessary factor responsible for SOHLH2 cytoplasmic to nuclear shuttling, but the trigger for SOHLH1 translocation was not identified[12]. Our data indicate sumoylation is a critical regulator for the localization of both SOHLH1 by direct regulation of SOHLH1, and potentially, as an indirect regulator of SOHLH2 localization. Furthermore, our data using the *Nobox-Luc* reporter assay indicate that SOHLH1 directly upregulates *Nobox* expression. Interestingly, the higher concentration of SOHLH1^{K345R} in the nucleus was not sufficient to increase its transcriptional activity, suggesting that there are additional regulatory inputs for SOHLH1 transcriptional activation. For instance, as a helix-loop-helix transcription factor, SOHLH1 can homodimerize, but also interact with several cofactors including SOHLH2, FIGLA, and LHX8 [12,45]. SOHLH1 clearly activates the *Nobox* promoter, but the lack of appropriate cofactors in our model system may be a rate limiting step. While localization of SOHLH1 is altered in oocytes of primordial and primary follicles, its downregulation in secondary follicles is unchanged by loss of sumoylation, as indicated in analysis of *Ube2i* cKO ovaries[1]. Though sumoylation can regulate protein stability this does not appear to be the case for SOHLH1[46,47]. Other PTMs, such as ubiquitin, may also drive this process, though this remains to be determined.

Our previous study showed that sumoylation directly conjugates to NOBOX, but the consequence of this modification remained unexplored. As with SOHLH1, there is a single consensus sumoylation site in NOBOX at K125. However, generating a K125R mouse knockin did not alter the fecundity of females, and furthermore, in HEK-293T cells, mutation of K125 had no effect on NOBOX sumoylation. This supports the current understanding that only around 60% of sumoylation falls within the consensus motif[23].

Using predictive software for non-consensus sumoylation sites followed by *in vitro* analysis, we positively identified K97 in NOBOX as a key site for conjugation of SUMO2/3. Mutation of K97 modulates the activity of NOBOX to increase expression of the key oocyte factor, *Gdf9*, demonstrating that NOBOX transcriptional modulation by post-translational modifications may be another mechanism by which the oocyte controls GDF9 levels during follicle development.

We detected both mono-SUMO1 and mono-SUMO2/3 modified NOBOX. While we determined that K97 is a key residue for SUMO2/3 modification, none of our mutants disrupted mono-SUMO1 modifications. Therefore, additional lysines must be modified by sumoylation on NOBOX that fall outside of current predictive algorithms. In addition, as with SOHLH1, laddering of higher molecular weight bands is seen when NOBOX is co-transfected with the non-deconjugatable forms of SUMO1 and SUMO2/3. Mutation of NOBOX K97 led to loss of the fastest migrating band, which most likely is mono-SUMO2/3 modification as this migrates at the predicted molecular weight of 85kDa. The double mutation of NOBOX K97/125 also led to the loss of fastest migrating band, and of the next higher band. This could indicate some cooperativity between K97 and K125 on NOBOX as multiple lysines can work together to enhance regulation by multi-modification of the target[48-50]. However, as with SOHLH1, it is possible these higher bands are an artifact of the deconjugation resistant SUMOs. Further investigation into the functionality of the higher molecular weight bands will be needed to determine if they are biologically relevant to NOBOX regulation.

Loss of SUMO2/3 conjugation at NOBOX K97 and K97/125 leads to an increase in *Gdf9* reporter activity, but no change with the *Pou5f1-luc* reporter. Differential regulation suggests that sumoylation is one mechanism to fine-tuning NOBOX activity. It has been previously hypothesized that NOBOX may directly interact with YY1 to regulate gene expression in the oocyte[36]. YY1 acts as both a transcriptional activator or repressor[51]. Human YY1, which a BLAST analysis showed shares 92% protein sequence homology with mouse YY1, contains two predicted SUMO-interacting motifs, while mouse YY1 contains three. We were unable to detect any direct interaction between YY1 and NOBOX nor any interaction with sumoylation-deficient NOBOX. However, SUMO is still known to be responsible for the recruitment of chromatin modifiers and other cofactors responsible for transcriptional activation or repression[52,53]. The possibility remains that SUMO is regulating NOBOX function through an unidentified transcriptional complex. *In vivo* analysis of a NOBOX^{K97R} mouse model would be highly useful in future studies to more clearly define the role of sumoylation in regulating NOBOX during early folliculogenesis.

Our finding that NOBOX^{K97R} and NOBOX$^{K97R/125R}$ expression increased activity with *Gdf9-luc* is opposite to the gene expression changes detected in *Ube2i* cKO ovaries, which show a significant decrease in *Gdf9* mRNA[18]. This likely indicates additional regulatory inputs on *Gdf9* expression, as oocyte-specific deletion of all sumoylation will disrupt multiple pathways causing both direct and indirect effects on transcription. Our current study provides a more direct analysis of the relationship between SUMO and NOBOX, uncovering that sumoylation has a repressive effect on NOBOX transcriptional activation of *Gdf9*. This may be useful for technological advancements to modify expression patterns

in oocytes, such as adding exogenous protein or DNA to oocytes for *in vitro* or *in vivo* maturation[54-56]. Being able to utilize sumo-deficient forms of NOBOX could allow for an increase in transcripts such as *Gdf9* without disrupting earlier developmental pathways that likely cause *Gdf9* downregulation in the *Ube2i* cKO model. This is especially relevant as many primary ovarian insufficiency (POI) *NOBOX* mutations have been shown to have defective activation of the *Gdf9* promoter[4,5].

The regulation of transcription factors involved in early folliculogenesis has important translational relevance. *NOBOX,* as previously mentioned, is one of the most highly mutated proteins in women diagnosed with POI, and SOHLH1 mutations have also been identified in POI patients[4,5,7]. Currently, known POI mutations include one lysine in NOBOX, K371, which could be a direct site of post-translational regulation[5]. Additionally, other mutations to non-lysine amino acids hold the potential to disrupt docking sites for necessary PTM enzymes such as UBE2I or E3 ligases. Sumoylation targets provide an intriguing therapeutic route as inhibitors for the SUMO E1 complex and the SUMO E2, UBE2I, are readily available[57-59]. Additionally, there is promising drug development for some of the SENPs, proteins necessary for desumoylation, providing another possible route of therapeutic targeting[60]. While there are challenges to using pharmaceuticals because of non-tissue specificity, the prospect remains for their potential utilization as an infertility intervention by being able to increase or decrease the activity of critical oocyte transcription factors.

As technology moves towards the generation of germ cells from pluripotent stem cells, NOBOX and SOHLH1 have been defined as essential transcription factors for this process[10]. PTM modification of these transcription factors may be an untapped resource for protocol optimization by controlling their activity. For instance, because SUMO2/3 acts in a repressive manner for both SOHLH1 and NOBOX, albeit through different mechanisms, SUMO inhibitors could be utilized to upregulate the activity of these proteins, increasing GDF9 production to crosstalk to somatic cells, which is often utilized as a marker for successfully creation of oocyte-like cells[61,62]. Conversely, SENP inhibitors could be used to manipulate the activity of these proteins during early follicle development. Post-translational modification in oocyte biology remains a highly underexplored area of female reproduction, but our studies show that is a critically important area that could hold the key to many therapeutic and technological advances.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors wish to thank Drs. Deborah L. Johnson (Baylor College of Medicine) for the *Myc-Sumo1ρ* and *Myc-Sumo2ρ* expression plasmids and Aleksandar Rajkovic (University of California, San Francisco) for the *pFLAG-Nobox*, rabbit anti-SOHLH1 antibody, and *Gdf9* and *Pou5f1* luciferase reporter plasmids. We thank Amanda Rodriguez for initial characterization of the $Nobox^{K125R}$ line. We additionally thank Ernesto Salas and Hannia Torralba-Salazar for technical assistance. We thank the Directors and staff at the following BCM advanced technology cores for assistance with these studies: Genetically Engineered Rat and Mouse Core, Human Tissue Acquisition and Pathology Core, Optical Imaging & Vital Microscopy Core, and Integrated Microscopy Core. Core services at Baylor College of Medicine were supported from the NIH (NCI grant P30 CA125123).

Grant Support:

These studies were supported by NIH grants R01 HD085994 and T32 HD098068 (to S.A.P.). Advanced Technology Core services at Baylor College of Medicine are supported from the NIH (NCI grant P30 CA125123). BKP was supported by a predoctoral fellowship from T32 HD098068.

Data Availability:

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

1. Pangas SA, Choi Y, Ballow DJ, et al. Oogenesis requires germ cell-specific transcriptional regulators Sohlh1 and Lhx8. Proc Natl Acad Sci U S A. May 23 2006;103(21):8090–5. doi:10.1073/pnas.0601083103 [PubMed: 16690745]
2. Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. Science. Aug 20 2004;305(5687):1157–9. doi:10.1126/science.1099755 [PubMed: 15326356]
3. Pan H, O'Brien MJ, Wigglesworth K, Eppig JJ, Schultz RM. Transcript profiling during mouse oocyte development and the effect of gonadotropin priming and development in vitro. Dev Biol. Oct 15 2005;286(2):493–506. doi:10.1016/j.ydbio.2005.08.023 [PubMed: 16168984]
4. Bouilly J, Bachelot A, Broutin I, Touraine P, Binart N. Novel NOBOX loss-of-function mutations account for 6.2% of cases in a large primary ovarian insufficiency cohort. Hum Mutat. Oct 2011;32(10):1108–13. doi:10.1002/humu.21543 [PubMed: 21837770]
5. Bouilly J, Roucher-Boulez F, Gompel A, et al. New NOBOX mutations identified in a large cohort of women with primary ovarian insufficiency decrease KIT-L expression. J Clin Endocrinol Metab. Mar 2015;100(3):994–1001. doi:10.1210/jc.2014-2761 [PubMed: 25514101]
6. Li L, Wang B, Zhang W, et al. A homozygous NOBOX truncating variant causes defective transcriptional activation and leads to primary ovarian insufficiency. Human Reproduction. 2016;32(1):248–255. doi:10.1093/humrep/dew271 [PubMed: 27836978]
7. Zhao S, Li G, Dalgleish R, et al. Transcription factor SOHLH1 potentially associated with primary ovarian insufficiency. Fertil Steril. Feb 2015;103(2):548–53.e5. doi:10.1016/j.fertnstert.2014.11.011 [PubMed: 25527234]
8. Bayram Y, Gulsuner S, Guran T, et al. Homozygous loss-of-function mutations in SOHLH1 in patients with nonsyndromic hypergonadotropic hypogonadism. J Clin Endocrinol Metab. May 2015;100(5):E808–14. doi:10.1210/jc.2015-1150 [PubMed: 25774885]
9. Jolly A, Bayram Y, Turan S, et al. Exome Sequencing of a Primary Ovarian Insufficiency Cohort Reveals Common Molecular Etiologies for a Spectrum of Disease. J Clin Endocrinol Metab. Aug 1 2019;104(8):3049–3067. doi:10.1210/jc.2019-00248 [PubMed: 31042289]
10. Hamazaki N, Kyogoku H, Araki H, et al. Reconstitution of the oocyte transcriptional network with transcription factors. Nature. 2021;589(7841):264–269. doi:10.1038/s41586-020-3027-9 [PubMed: 33328630]
11. Belli M, Cimadomo D, Merico V, Redi CA, Garagna S, Zuccotti M. The NOBOX protein becomes undetectable in developmentally competent antral and ovulated oocytes. Int J Dev Biol. 2013;57(1):35–9. doi:10.1387/ijdb.120125mz [PubMed: 23585350]
12. Shin YH, Ren Y, Suzuki H, et al. Transcription factors SOHLH1 and SOHLH2 coordinate oocyte differentiation without affecting meiosis I. J Clin Invest. Jun 1 2017;127(6):2106–2117. doi:10.1172/jci90281 [PubMed: 28504655]
13. Choi Y, Rajkovic A. Characterization of NOBOX DNA Binding Specificity and Its Regulation of Gdf9 and Pou5f1 Promoters. Journal of Biological Chemistry. 2006;281(47):35747–35756. doi:10.1074/jbc.m604008200 [PubMed: 16997917]
14. Capili AD, Lima CD. Taking it step by step: mechanistic insights from structural studies of ubiquitin/ubiquitin-like protein modification pathways. Curr Opin Struct Biol. Dec 2007;17(6):726–35. doi:10.1016/j.sbi.2007.08.018 [PubMed: 17919899]

15. Johnson ES. Protein Modification by SUMO. Annual Review of Biochemistry. 2004;73(1):355–382. doi:10.1146/annurev.biochem.73.011303.074118

16. Wang ZB, Ou XH, Tong JS, et al. The SUMO pathway functions in mouse oocyte maturation. Cell Cycle. Jul 1 2010;9(13):2640–6. doi:10.4161/cc.9.13.12120 [PubMed: 20543581]

17. Hickey CM, Wilson NR, Hochstrasser M. Function and regulation of SUMO proteases. Nature Reviews Molecular Cell Biology. 2012/12/01 2012;13(12):755–766. doi:10.1038/nrm3478 [PubMed: 23175280]

18. Rodriguez A, Briley SM, Patton BK, et al. Loss of the E2 SUMO-conjugating enzyme Ube2i in oocytes during ovarian folliculogenesis causes infertility in mice. Development. Dec 2 2019;146(23)doi:10.1242/dev.176701

19. Gill G Something about SUMO inhibits transcription. Curr Opin Genet Dev. Oct 2005;15(5):536–41. doi:10.1016/j.gde.2005.07.004 [PubMed: 16095902]

20. Lyst MJ, Stancheva I. A role for SUMO modification in transcriptional repression and activation. Biochem Soc Trans. Dec 2007;35(Pt 6):1389–92. doi:10.1042/bst0351389 [PubMed: 18031228]

21. Shiio Y, Eisenman RN. Histone sumoylation is associated with transcriptional repression. Proc Natl Acad Sci U S A. Nov 11 2003;100(23):13225–30. doi:10.1073/pnas.1735528100 [PubMed: 14578449]

22. Verger A, Perdomo J, Crossley M. Modification with SUMO. A role in transcriptional regulation. EMBO Rep. Feb 2003;4(2):137–42. doi:10.1038/sj.embor.embor738 [PubMed: 12612601]

23. Zhao Q, Xie Y, Zheng Y, et al. GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Research. 2014;42(W1):W325–W330. doi:10.1093/nar/gku383 [PubMed: 24880689]

24. Ren J, Gao X, Jin C, et al. Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0. PROTEOMICS. 2009;9(12):3409–3412. doi:10.1002/pmic.200800646 [PubMed: 29658196]

25. Suzuki H, Ahn HW, Chu T, et al. SOHLH1 and SOHLH2 coordinate spermatogonial differentiation. Dev Biol. Jan 15 2012;361(2):301–12. doi:10.1016/j.ydbio.2011.10.027 [PubMed: 22056784]

26. Rohira AD, Chen CY, Allen JR, Johnson DL. Covalent small ubiquitin-like modifier (SUMO) modification of Maf1 protein controls RNA polymerase III-dependent transcription repression. J Biol Chem. Jun 28 2013;288(26):19288–95. doi:10.1074/jbc.M113.473744 [PubMed: 23673667]

27. Suzumori N, Yan C, Matzuk MM, Rajkovic A. Nobox is a homeobox-encoding gene preferentially expressed in primordial and growing oocytes. Mechanisms of Development. 2002;111(1-2):137–141. doi:10.1016/s0925-4773(01)00620-7 [PubMed: 11804785]

28. Kaneko T, Li L, Li SS. The SH3 domain--a family of versatile peptide- and protein-recognition module. Front Biosci. May 1 2008;13:4938–52. doi:10.2741/3053 [PubMed: 18508559]

29. Wang D, Liang Y, Xu D. Capsule network for protein post-translational modification site prediction. Bioinformatics. Jul 15 2019;35(14):2386–2394. doi:10.1093/bioinformatics/bty977 [PubMed: 30520972]

30. Wang D, Liu D, Yuchi J, et al. MusiteDeep: a deep-learning based webserver for protein post-translational modification site prediction and visualization. Nucleic Acids Res. Jul 2 2020;48(W1):W140–w146. doi:10.1093/nar/gkaa275 [PubMed: 32324217]

31. Wang D, Zeng S, Xu C, et al. MusiteDeep: a deep-learning framework for general and kinase-specific phosphorylation site prediction. Bioinformatics. Dec 15 2017;33(24):3909–3916. doi:10.1093/bioinformatics/btx496 [PubMed: 29036382]

32. Chen L, Ma Y, Qian L, Wang J. Sumoylation regulates nuclear localization and function of zinc finger transcription factor ZIC3. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2013/12/01/ 2013;1833(12):2725–2733. doi:10.1016/j.bbamcr.2013.07.009 [PubMed: 23872418]

33. Du JX, Bialkowska AB, McConnell BB, Yang VW. SUMOylation regulates nuclear localization of Krüppel-like factor 5. J Biol Chem. Nov 14 2008;283(46):31991–2002. doi:10.1074/jbc.M803612200 [PubMed: 18782761]

34. Wen D, Wu J, Wang L, Fu Z. SUMOylation Promotes Nuclear Import and Stabilization of Polo-like Kinase 1 to Support Its Mitotic Function. Cell Rep. Nov 21 2017;21(8):2147–2159. doi:10.1016/j.celrep.2017.10.085 [PubMed: 29166606]

35. Liu B, Shuai K. Regulation of the sumoylation system in gene expression. Curr Opin Cell Biol. Jun 2008;20(3):288–93. doi:10.1016/j.ceb.2008.03.014 [PubMed: 18468876]
36. Griffith GJ, Trask MC, Hiller J, et al. Yin-Yang1 Is Required in the Mammalian Oocyte for Follicle Expansion1. Biology of Reproduction. 2010;84(4):654–663. doi:10.1095/biolreprod.110.087213 [PubMed: 21123818]
37. Rosonina E A conserved role for transcription factor sumoylation in binding-site selection. Current Genetics. 2019/12/01 2019;65(6):1307–1312. doi:10.1007/s00294-019-00992-w [PubMed: 31093693]
38. Rosonina E, Akhter A, Dou Y, Babu J, Sri Theivakadadcham VS. Regulation of transcription factors by sumoylation. Transcription. Aug 8 2017;8(4):220–231. doi:10.1080/21541264.2017.1311829 [PubMed: 28379052]
39. Da Silva-Ferrada E, Lopitz-Otsoa F, Lang V, Rodríguez MS, Matthiesen R. Strategies to Identify Recognition Signals and Targets of SUMOylation. Biochem Res Int. 2012;2012:875148. doi:10.1155/2012/875148 [PubMed: 22811915]
40. Tatham MH, Jaffray E, Vaughan OA, et al. Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. J Biol Chem. Sep 21 2001;276(38):35368–74. doi:10.1074/jbc.M104214200 [PubMed: 11451954]
41. Yang M, Hsu CT, Ting CY, Liu LF, Hwang J. Assembly of a polymeric chain of SUMO1 on human topoisomerase I in vitro. J Biol Chem. Mar 24 2006;281(12):8264–74. doi:10.1074/jbc.M510364200 [PubMed: 16428803]
42. Jansen NS, Vertegaal ACO. A Chain of Events: Regulating Target Proteins by SUMO Polymers. Trends Biochem Sci. Feb 2021;46(2):113–123. doi:10.1016/j.tibs.2020.09.002 [PubMed: 33008689]
43. Matic I, van Hagen M, Schimmel J, et al. In vivo identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. Mol Cell Proteomics. Jan 2008;7(1):132–44. doi:10.1074/mcp.M700173-MCP200 [PubMed: 17938407]
44. Salas-Lloret D, González-Prieto R. Insights in Post-Translational Modifications: Ubiquitin and SUMO. Int J Mol Sci. Mar 18 2022;23(6)doi:10.3390/ijms23063281
45. Wang Z, Liu CY, Zhao Y, Dean J. FIGLA, LHX8 and SOHLH1 transcription factor networks regulate mouse oocyte growth and differentiation. Nucleic Acids Res. Apr 17 2020;48(7):3525–3541. doi:10.1093/nar/gkaa101 [PubMed: 32086523]
46. Cheng J, Kang X, Zhang S, Yeh ET. SUMO-specific protease 1 is essential for stabilization of HIF1alpha during hypoxia. Cell. Nov 2 2007;131(3):584–95. doi:10.1016/j.cell.2007.08.045 [PubMed: 17981124]
47. Li S, Wang M, Qu X, et al. SUMOylation of PES1 upregulates its stability and function via inhibiting its ubiquitination. Oncotarget. Aug 2 2016;7(31):50522–50534. doi:10.18632/oncotarget.10494 [PubMed: 27409667]
48. Desterro JM, Rodriguez MS, Hay RT. SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. Mol Cell. Aug 1998;2(2):233–9. doi:10.1016/s1097-2765(00)80133-1 [PubMed: 9734360]
49. Shalizi A, Gaudillière B, Yuan Z, et al. A calcium-regulated MEF2 sumoylation switch controls postsynaptic differentiation. Science. Feb 17 2006;311(5763):1012–7. doi:10.1126/science.1122513 [PubMed: 16484498]
50. Escobar-Ramirez A, Vercoutter-Edouart AS, Mortuaire M, et al. Modification by SUMOylation Controls Both the Transcriptional Activity and the Stability of Delta-Lactoferrin. PLoS One. 2015;10(6):e0129965. doi:10.1371/journal.pone.0129965 [PubMed: 26090800]
51. Agarwal N, Theodorescu D. The Role of Transcription Factor YY1 in the Biology of Cancer. Crit Rev Oncog. 2017;22(1-2):13–21. doi:10.1615/CritRevOncog.2017021071 [PubMed: 29604933]
52. Ouyang J, Gill G. SUMO engages multiple corepressors to regulate chromatin structure and transcription. Epigenetics. Oct 1 2009;4(7):440–4. doi:10.4161/epi.4.7.9807 [PubMed: 19829068]
53. Ouyang J, Shi Y, Valin A, Xuan Y, Gill G. Direct binding of CoREST1 to SUMO-2/3 contributes to gene-specific repression by the LSD1/CoREST1/HDAC complex. Mol Cell. Apr 24 2009;34(2):145–54. doi:10.1016/j.molcel.2009.03.013 [PubMed: 19394292]

54. Mobarak H, Heidarpour M, Tsai P-SJ, et al. Autologous mitochondrial microinjection; a strategy to improve the oocyte quality and subsequent reproductive outcome during aging. Cell & Bioscience. 2019/11/29 2019;9(1):95. doi:10.1186/s13578-019-0360-5 [PubMed: 31798829]

55. Wu X, Shen W, Zhang B, Meng A. The genetic program of oocytes can be modified in vivo in the zebrafish ovary. Journal of Molecular Cell Biology. 2018;10(6):479–493. doi:10.1093/jmcb/mjy044 [PubMed: 30060229]

56. Stein P, Schindler K. Mouse oocyte microinjection, maturation and ploidy assessment. J Vis Exp. Jul 23 2011;(53)doi:10.3791/2851

57. Fukuda I, Ito A, Hirai G, et al. Ginkgolic acid inhibits protein SUMOylation by blocking formation of the E1-SUMO intermediate. Chem Biol. Feb 27 2009;16(2):133–40. doi:10.1016/j.chembiol.2009.01.009 [PubMed: 19246003]

58. Fukuda I, Ito A, Uramoto M, et al. Kerriamycin B inhibits protein SUMOylation. J Antibiot (Tokyo). Apr 2009;62(4):221–4. doi:10.1038/ja.2009.10 [PubMed: 19265871]

59. Hirohama M, Kumar A, Fukuda I, et al. Spectomycin B1 as a novel SUMOylation inhibitor that directly binds to SUMO E2. ACS Chem Biol. Dec 20 2013;8(12):2635–42. doi:10.1021/cb400630z [PubMed: 24143955]

60. Tokarz P, Woźniak K. SENP Proteases as Potential Targets for Cancer Therapy. Cancers (Basel). Apr 24 2021;13(9)doi:10.3390/cancers13092059

61. Salvador LM, Silva CP, Kostetskii I, Radice GL, Strauss JF, 3rd. The promoter of the oocyte-specific gene, Gdf9, is active in population of cultured mouse embryonic stem cells with an oocyte-like phenotype. Methods. Jun 2008;45(2):172–81. doi:10.1016/j.ymeth.2008.03.004 [PubMed: 18593614]

62. Qing T, Shi Y, Qin H, et al. Induction of oocyte-like cells from mouse embryonic stem cells by co-culture with ovarian granulosa cells. Differentiation. 2007/12/01/ 2007;75(10):902–911. doi:10.1111/j.1432-0436.2007.00181.x [PubMed: 17490416]

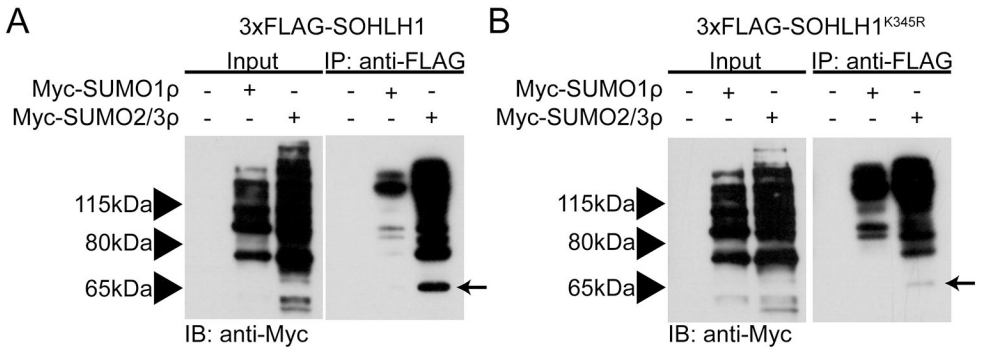

Figure 1: Modification of SOHLH1 by sumoylation.

(A) Immunoprecipitation of HEK-293T cells co-transfected with *3xFLAG-Sohlh1* and *Myc-Sumo1ρ* or *Myc-Sumo2/3ρ*. Whole cell protein lysates were immunoprecipitated (IP) with an anti-FLAG antibody and immunoblotted (IB) for Myc. Molecular weight (MW) is shown in kilodaltons (kDa) to the left of each blot, with the input lysate shown in the left panel and the FLAG-IP shown in the right. Arrow indicates the predicted MW for mono-SUMO2/3 modification of SOHLH1. (B) Immunoprecipitation of 3xFLAG-SOHLH1^{K345R} mutant. Arrow indicates significantly diminished band of mono-SUMO2/3 modified SOHLH1^{K345R}.

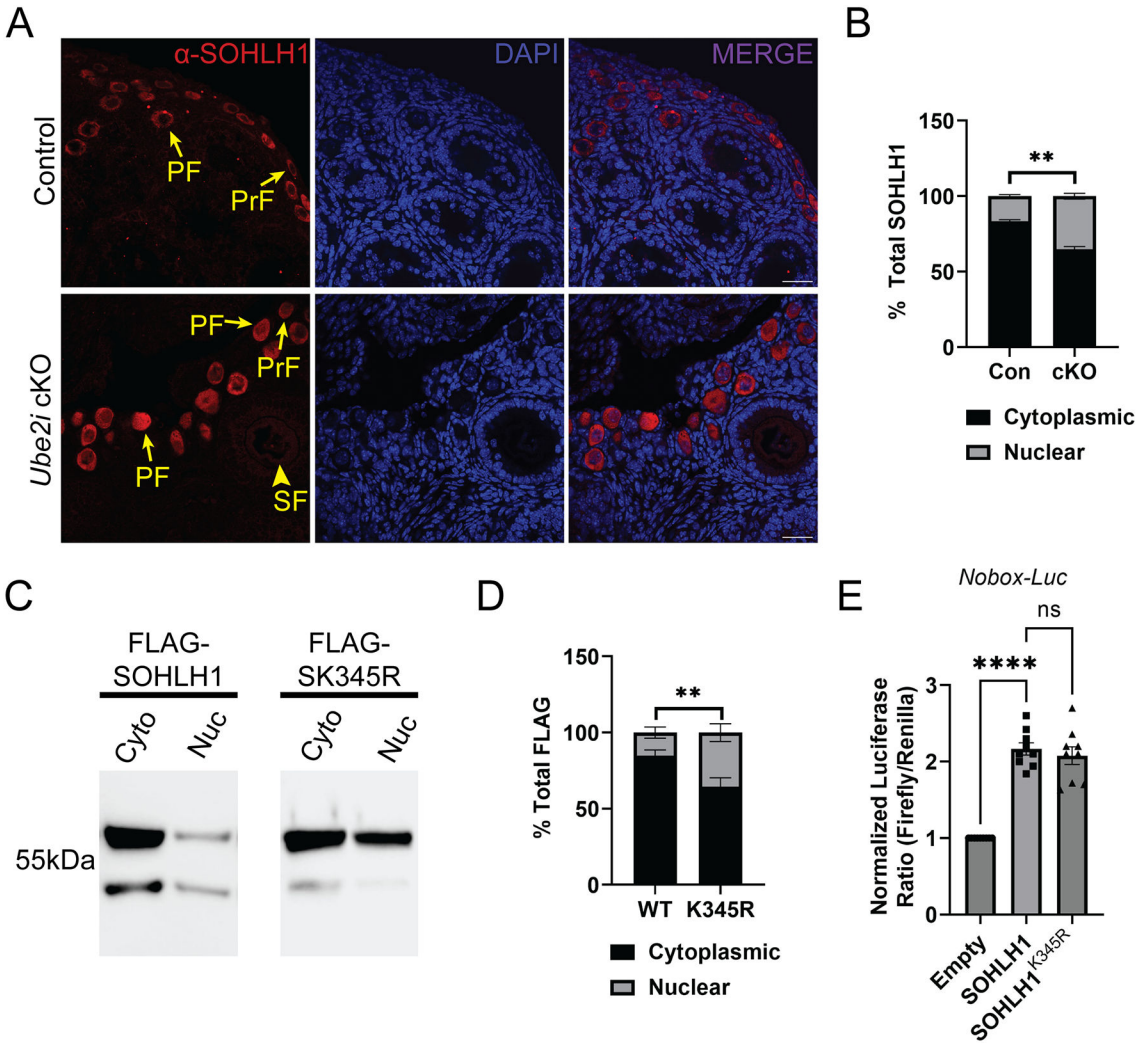

Figure 2: SOHLH1^{K345R} enhances the nuclear localization of SOHLH1.

(A) Indirect immunofluorescence (IF) of SOHLH1 in control or *Ube2i* cKO 14-day old ovary sections by confocal microscopy. Red indicates SOHLH1 immunoreactivity; blue is DAPI nuclear stain. Arrows indicate representative follicles, PrF: Primordial Follicle, PF: Primary Follicle, SF: Secondary Follicle. Scale bar indicates 20 μm. (B) Quantification of anti-SOHLH1 fluorescent intensity in Panel A by confocal microscopy. "Con" denotes control ovaries, cKO denotes "*Ube2i* cKO". n=55 oocytes across three independent ovaries per genotype. (C) Immunostaining for α-FLAG in cell lysates transfected with either 3xFLAG-SOHLH1 or 3xFLAG-SOHLH1^{K345R} and then fractioned into cytoplasmic (cyto) and nuclear (nuc) fractions. Representative sample shown. n=3 per group (D) Quantification of the amount of wild-type 3xFLAG-SOHLH1 (WT) or 3xFLAG-SOHLH1^{K345R} (K345R) protein found in each respective portion in lysates of transfected HEK-293T cells. n=3 separate cell lysates per group. Statistical analysis in panels B, C by Fisher's Exact test. (E) Normalized luciferase expression for the *Nobox-luc* co-transfected in HEK-293T cells with empty parent vector pcDNA3.1, or expression vectors for 3xFLAG-SOHLH1 or 3x-FLAG-SOHLH1^{K345R}. Three replicates were averaged for each experiment and the experiment was

repeated independently nine times (n=9). Each bar represents the mean +/− s.e.m. Statistical analysis by one way ANOVA with multiple comparisons against WT SOHLH1.

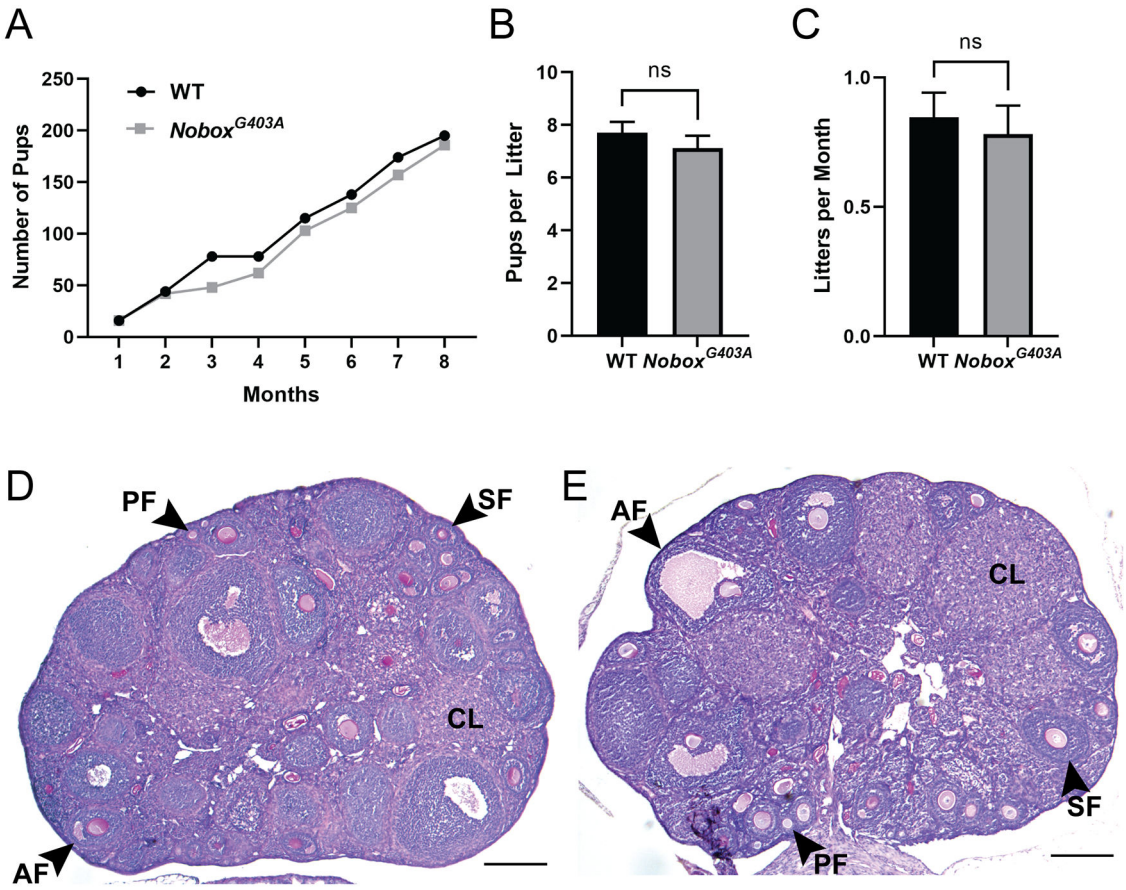

Figure 3: *Nobox*$^{G403A/G403A}$ females have normal fecundity.

(A) Cumulative number of pups for four wild-type (WT) and four homozygous *Nobox*$^{G403A/G403A}$ female mice bred in continuous paired matings. “Months” indicates months of breeding in pair-mated cages after mice reached sexual maturity (6-weeks of age for females). (B) Average pups per litter and (C) average litters per month for wild type (WT) (n=5) or *Nobox*$^{G403A/G403A}$ (n=5) females bred continuously for eight months. Bars indicate mean +/− s.e.m. No significant difference was found. Statistical analysis by Student’s *t*-test between genotypes. (D,E) PAS histological analysis of ovaries collected from six week old WT (D) or homozygous *Nobox*$^{G403A/G403A}$ (E) mice. Arrows indicate relevant follicular stages: PF: primary follicle, SF: secondary follicle, AF: antral follicle, CL: corpus luteum. Scale bar is 100 μm.

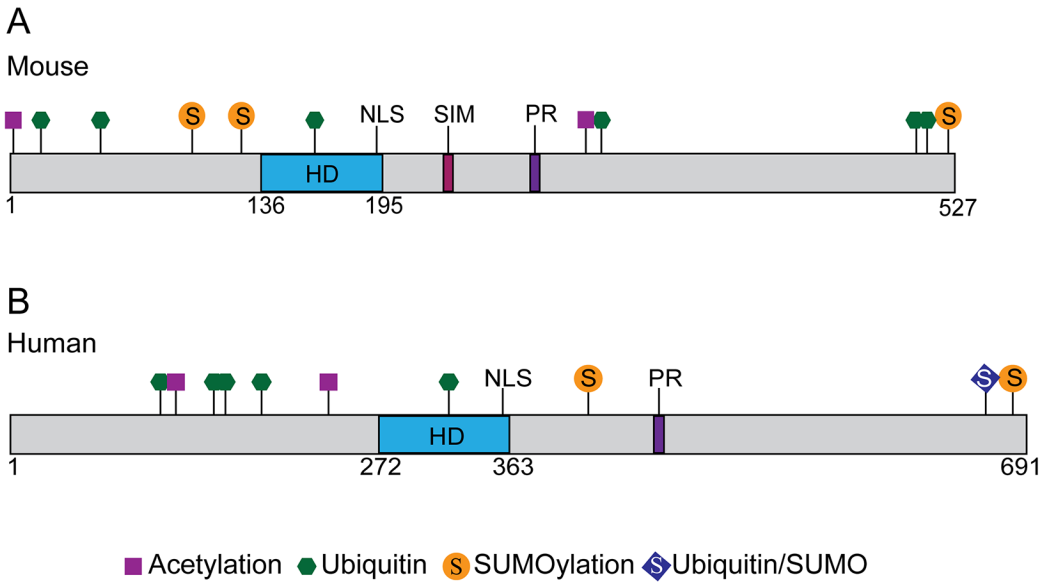

Figure 4:

In silico analysis of human and mouse NOBOX. (A) Mouse NOBOX and (B) human NOBOX depicted with all predicted lysines modifications. Additional important protein domains present in each protein include the homeodomain (HD), nuclear localization signal (NLS), SUMO-interacting motif (SIM), and proline-rich region (PR).

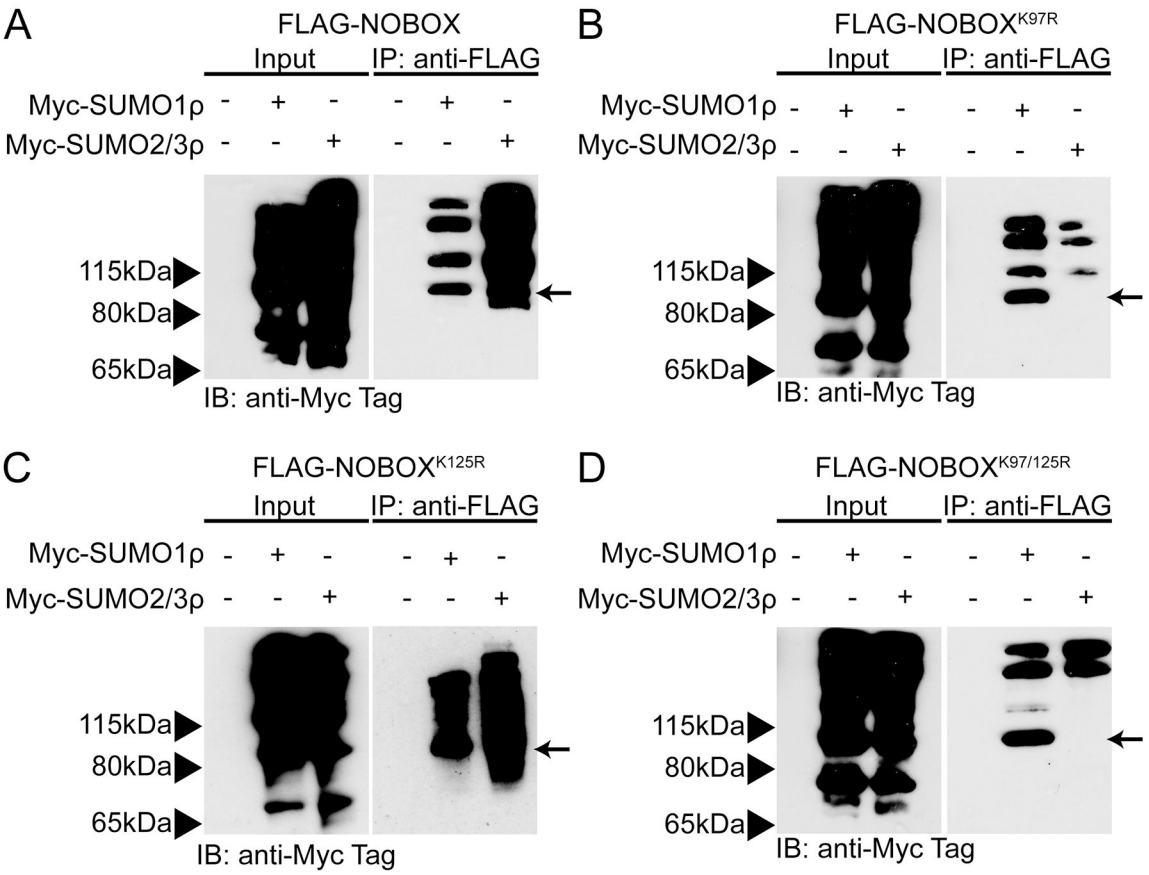

Figure 5: Modification of NOBOX by sumoylation.

(A) Immunoprecipitation (IP) of cell lysates with anti-FLAG from HEK-293T cells co-transfected with *FLAG-Nobox* and *Myc-Sumo1ρ* or *Myc-Sumo2/3ρ* and immunoblotting (IB) against Myc. Molecular weight (MW) is shown in kilodaltons (kDa) to the left of each blot, with the input lysate shown in the left panel and the FLAG-IP shown in the right. Arrow indicates the predicted MW for mono-SUMO modification of NOBOX. (B-D) Experiment performed as in panel A, except *FLAG-Nobox*^{K97R} (B), *FLAG-Nobox*^{K125R} (C), or *FLAG-Nobox*$^{K97/125R}$ (D) were co-transfected with *Myc-Sumo1ρ* or *Myc-Sumo2/3ρ*.

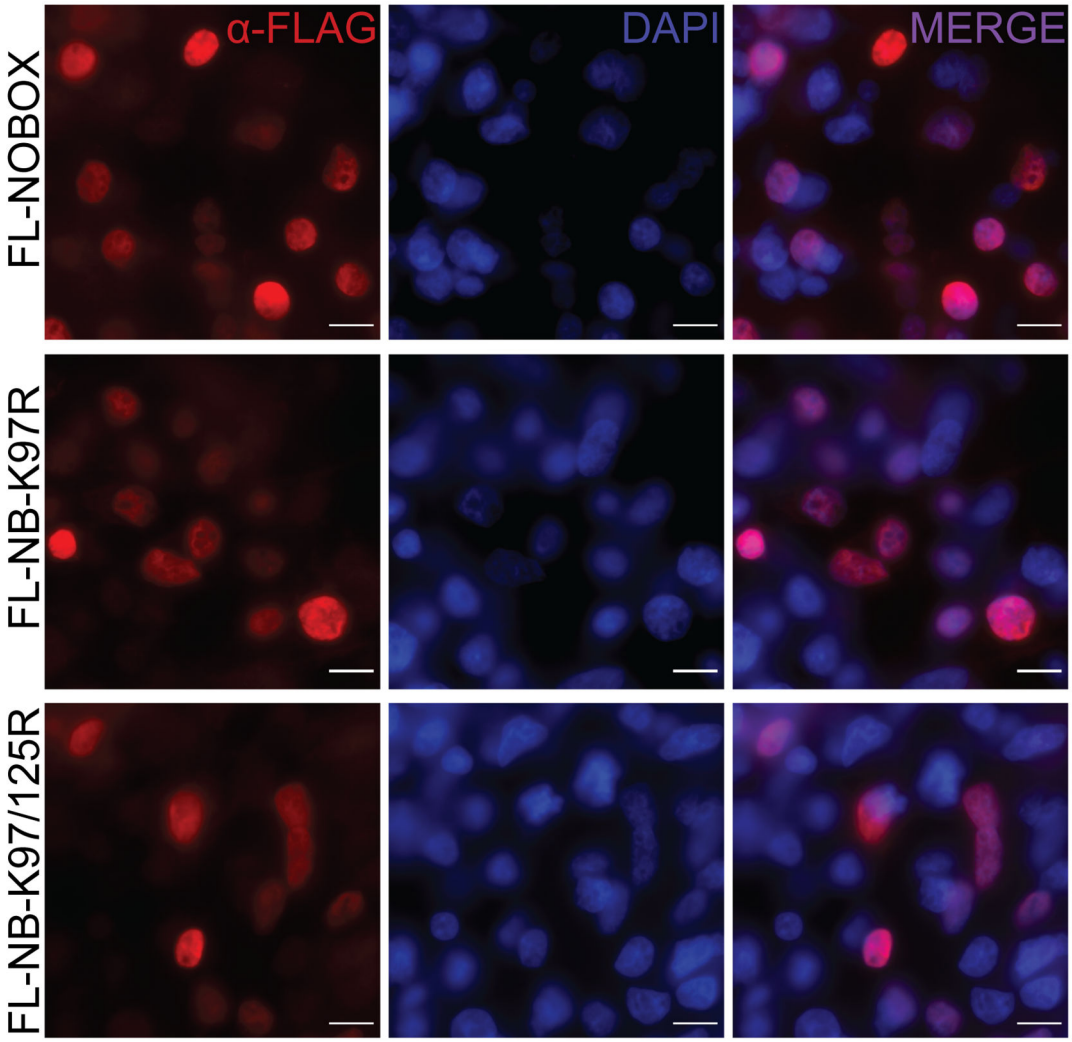

Figure 6: SUMO mutant NOBOX does not alter subcellular localization.

Indirect immunofluorescent imaging by fluorescent microscopy against FLAG in HEK-293T cells transfected with expression plasmids for FLAG-tagged (FL) NOBOX, NOBOX (NB)-K97R or NOBOX (NB)-K97/125R. Red fluorescence indicates FLAG immunoreactivity to the FLAG-tagged NOBOX protein (left column); blue indicates DAPI nuclear stain (middle column), and merged images (right column). All panels taken at the same exposure and magnification. Scale bar indicates 20 μm.

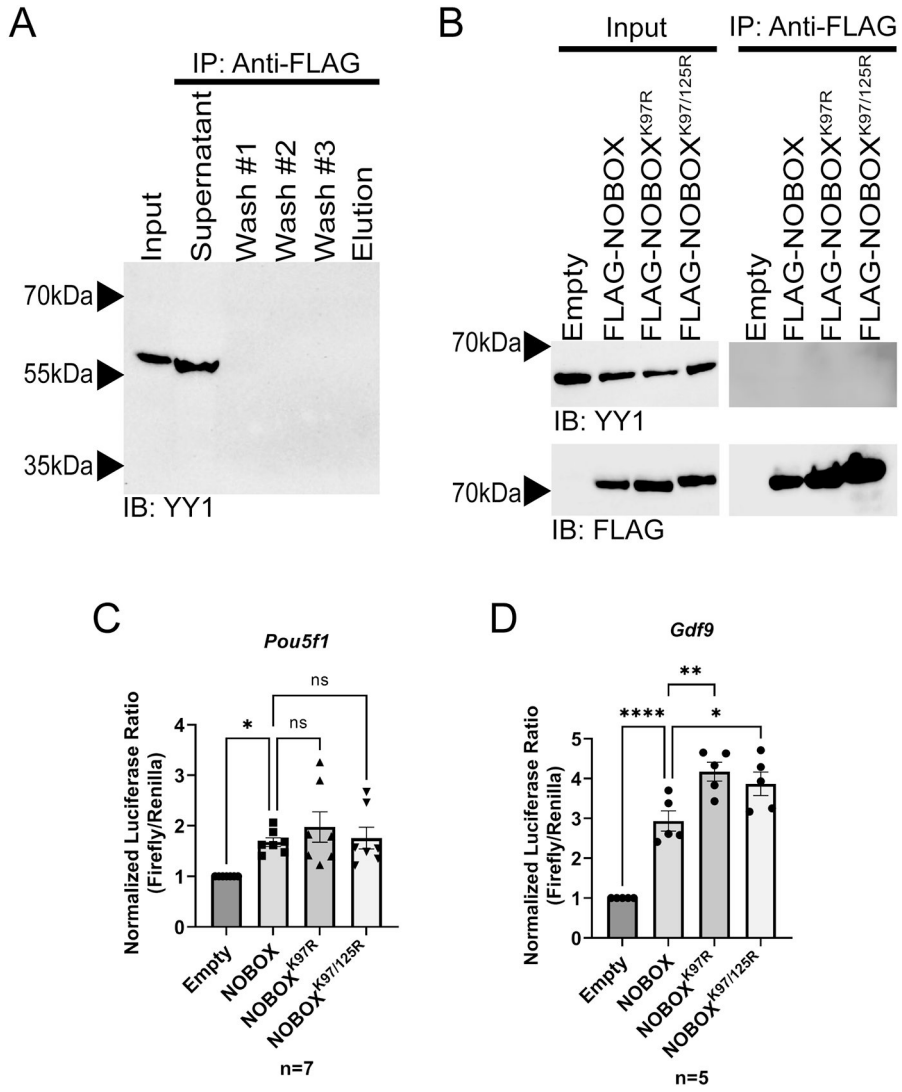

Figure 7: NOBOX^{K97R} or NOBOX$^{K97/125R}$ alters NOBOX transcriptional activity but do not drive an interaction with YY1.

(A) Protein lysates from HEK-293T cells transfected with FLAG-NOBOX were immunoprecipitated for anti-FLAG, washed three times, and immunoblotted. Endogenous YY1 is seen at the appropriate molecular weight of 60kDa in the input and supernatant lanes. (B) HEK-293T cells were transfected with expression constructs for FLAG-NOBOX or with the NOBOX-SUMO mutants and then protein lysates immunoprecipitated for anti-FLAG. Immunoblotting was first done with YY1, but no bands could be detected in the IP lanes. Immunoblotting with FLAG shows appropriate expression of transfected proteins (indicated in IB:FLAG immunoblots). (C,D) Normalized luciferase expression for the *Pou5f1-luc* (C) or *Gdf9-luc* (D) co-transfected in HEK-293T cells with empty parent vector pcDNA3.1, or expression vectors for untagged NOBOX, NOBOX^{K97R}, or NOBOX$^{K97/125R}$. Three replicates were averaged for each experiment (shown by triangles in C, or circles in D), and the experiment was repeated independently seven times (C) or five times (D). Each

bar represents the mean +/− s.e.m. Statistical analysis by one way ANOVA with multiple comparisons against WT NOBOX.

Table 1:

List of *NOBOX* mutations found in POI patients that have experimental or predicted validation.

Basepair	Nucleotide Change	Protein Change	Validation	Reference
271	G>T	G91W	Reduced *Gdf9-luc* transcription *in vitro*	[4]
349	C>T	R117W		
907	C>T	R303X		
1025	G>C	S342T		
1048	G>T	V350L		
113	G>T	R44L		[5]
331	G>A	G111R		
1112	A>C	K371T		
1856	C>T	P619L		
68	G>A	G23D	Predicted to be benign	[6]
350	G>A	R117Q		