Studies on the *Sox9* gene regulation through the testis-specific enhancer in the Amami spiny rat (*Tokudaia osimensis*), an XO/XO mammal

アマミトゲネズミにおける精巣特異的エンハンサーを介した *Sox9* 遺伝子制御の研究

A DISSERTATION

Submitted to the Graduate School of Life Science,

Hokkaido University

In partial fulfilment of the requirements for the degree

DOCTOR OF LIFE SCIENCE

Shoichiro Mitsukawa

September 2025

Contents

Abbreviation

ATAC-seq	assay for transposase-accessible chromatin with high-throughput sequencing
BCA	bicinchoninic acid
CBB	Coomassie brilliant blue
ChIP-seq	chromatin immunoprecipitation sequencing
CNV	copy number variation
DAVID	Database for Annotation, Visualization, and Integrated Discovery
DEPC	diethyl pyrocarbonate
DIG	digoxigenin
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DSD	disorders of sex development
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
ESD	environmental sex determination
FBS	fetal bovine serum
GO	gene ontology
GSD	genetic or genotypic sex determination
GTF	Gene transfer format
IPTG	isopropyl β-D-thiogalactopyranoside
LC-MS/MS	liquid chromatography-tandem mass spectrometry
NBT/BCIP	nitroblue Tetrazolium/5-Bromo-4-chloro-3-indolyl-phosphate
NCBI	National Center for Biotechnology Information
N.S.	no significance
RIPA	radioimmunoprecipitation
RSEM	RNA sequencing by expectation maximization
RT-PCR	reverse transcription polymerase chain reaction

RNA-seq	RNA sequencing
STAR	spliced transcripts alignment to a reference
Tris	tris (hydroxymethyl) aminomethane
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde
PMSF	phenylmethylsulfonyl fluoride
PVDF	polyvinylidene difluoride
qPCR	quantitative polymerase chain reaction
scRNA-seq	single-cell RNA sequencing
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEM	Standard error of the mean
TBS	tris-buffered saline
TPM	transcripts per million
WISH	whole-mount *in situ* hybridization

General Introduction

For all organisms, reproduction is essential for the flourishing of the species. Sexually reproducing organisms must develop testes and ovaries individually to provide sperm and eggs. Therefore, a sex-determining mechanism is an important process that determines the sex of an individual. The sex determination system can be classified into two groups: ESD and GSD[1]. Particularly, temperature-dependent sex determination is well known as ESD in reptiles such as alligators, most turtles, and some lizards. In *Alligator mississippiensis*, eggs hatched at 34°C during a critical temperature-sensitive period produce male offspring, while eggs hatched at temperatures below 30°C produce female offspring[2]. Gene expression patterns associated with male differentiation are activated by transient receptor potential cation channel subfamily V member 4 (TRPV4), which is a thermosensitive TRP channel at male-producing temperature[3].

On the other hand, mammals, birds, amphibians, many other reptiles, and fish follow GSD. GSD mainly depends on the sex-determining gene on a sex chromosome. In birds and some poikilothermal vertebrates such as frogs, have female heterogametic sex chromosomes. Females possess one Z and one W chromosome (ZW), whereas males possess two Z chromosomes (ZZ). In *Xenopus laevis*, doublesex and mab3 related transcription factor W (*DM-W*) on W chromosome is found as a sex determining gene[4]. *DM-W* is a paralogue of doublesex and mab3 related transcription factor 1 (*DMRT1*/*Dmrt1*), which is essential for testicular differentiation in mice[5,6] or is related to testis formation in chicken[7,8].

Conversely, most mammals and some other vertebrates show a male heterogametic XX/XY system[9]. Males possess one X and one Y chromosome (XY), while females possess two X chromosomes (XX). In medaka, a paralogue of *Dmrt1*, *Dmy*, is identified as a sex determining gene[10,11]. In most placental mammals, sex-determining region of the Y chromosome (*SRY*/*Sry*) functions as a sex-determining gene[12-14]. This gene belongs to the SRY-box (*SOX*) family, members of which contain a high mobility group (HMG) domain. *SRY*/*Sry* evolved from SRY-box 3 (*SOX3*/*Sox3*) independently in mammals[15,16]. SRY regulates SRY-box 9 (*SOX9*/*Sox9*) as a transcription factor by binding to its enhancer with an HMG domain, resulting in the trigger of testicular differentiation[17].

SOX9/*Sox9*, located on an autosome, is a transcription factor that belongs to the SRY family. It was first identified in the early 1990s as a gene closely related to the *SRY* gene and was initially characterized for its role in chondrogenesis and skeletal development[18,19]. SOX9 contains an HMG domain that allows it to bind to specific DNA sequences and regulate gene expression primarily involved in various developmental processes[20]. In chondrogenesis, *Sox9* acts as a master regulator, promoting the differentiation of chondrocyte precursor cells and controlling the expression of a cartilage-specific gene, such as collagen type II alpha 1 chain (*Col2a1*)[21]. *Sox9* is also required to maintain the testes in adult, and functions as a pivotal regulator in embryonic testis differentiation. After *SOX9*/*Sox9* up-regulation by SRY in male gonads, SOX9 activates the transcription of various downstream

genes, including anti-Müllerian hormone (*AMH*/*Amh*), which promotes Sertoli cell differentiation and testis cord formation[22].

The testis differentiation mechanism by *SRY*/*Sry* and *SOX9*/*Sox9* is widely conserved in most placental mammals[23,24]. Interestingly, the genus *Tokudaia*, a member of Murinae (Muridae), is the exception that does not follow the system. The genus *Tokudaia* consists of three species: Okinawa spiny rat (*Tokudaia muenninki*; *T. muenninki*), Tokunoshima spiny rat (*Tokudaia tokunoshimensis*; *T. tokunoshimensis*), and Amami spiny rat (*Tokudaia osimensis*; *T. osimensis*)[25]. They inhabit endemic three islands in the southernmost part of Japan. All three species are endangered (The IUCN Red List of Threatened Species, DOI: e.T21972A22409515; e.T136695A22409747; e.T21973A22409638) and have been protected by the Japanese government as a natural monument and national endangered species of wild fauna and flora since 1972 and 2016, respectively.

T. muenninki (2*n*=44) is the only species in the genus with XX/XY sex chromosomes[26]. However, this species possesses unique sex chromosomes fused with a pair of autosomes. The short arms derived from autosomes are called “neo-X” and “neo-Y” regions on the X and Y chromosomes, respectively. These regions are homologues of mouse chromosomes 11 and 16[27]. Due to sex chromosome-autosome fusions, Y chromosome of *T. muenninki* is exceptionally large, containing over 70 copies of *Sry*, of which at least three retain open reading frames[27]. However, all copies share a mutation from alanine to serine at position 21 in the HMG domain[26]. The DNA-binding ability of *T. muenninki* SRY is slightly lower than that of the mouse

homologue, but is retained in *T. muenninki*[28]. The function *in vivo* using transgenic mice remains unclear due to the lack of a polyglutamine-rich (Q-rich) domain, resulting in instability of the SRY protein in mouse cells[29].

T. osimensis (2*n*=25, XO/XO) and *T. tokunoshimensis* (2*n*=45, XO/XO) lack the Y chromosome and *Sry*[30-32]; thus, it has evolved a new sex-determining mechanism independent of *Sry*. Indeed, expression of *Sox9*, a target of *Sry*, and its downstream regulation are conserved in these species as well as in general placental mammals[33,34]. In *T. osimensis*, a single sex-related genomic difference was found in a male-specific duplication of a 17-kb unit located 430 kb upstream of *Sox9* on the short arm of chromosome 3[35]. The duplicated unit contains a 1,262-bp element, termed *T. osimensis* Enh14 (tosEnh14), homologous to mouse Enh14 (mEnh14). The sequence identity between mEnh14 (1,288 bp) and tosEnh14 (1,262 bp) is 87.7%[35]. mEnh14 showed robust testis-specific β-Gal activity in an *in vivo* reporter gene assay, and DNaseI-seq, ATAC-seq, and H3K27ac ChIP-seq data all suggest that this enhancer is active and accessible only in Sertoli cells[36].

The above information indicates that *Sox9* is regulated by tosEnh14 without *Sry*, resulting in sexual dimorphism in *T. osimensis*. This study aimed to elucidate the molecular mechanism of *Sox9* regulation by tosEnh14 on testis development and sex determination. Particularly, transcription factors that regulate tosEnh14 were investigated. In Chapter I, one adult male mouse was used to explore for transcription factors that bind to tosEnh14, aiming to investigate how tosEnh14 regulates *Sox9* expression in the adult testis. The screening was performed in terms of DNA-protein

interactions. A reporter gene assay was subsequently performed to evaluate the transcriptional activity of the transcription factor candidates identified in the screening. In Chapter II, screening of transcription factors was performed using the gonads of embryos just after the sex determination. This screening was focused on the timing and levels of expression.

Chapter I

Exploration of a candidate transcription factor using an adult mouse testis

1.1 Abstract

Expression of *Sox9* is regulated by distal enhancers. According to a previous report, *Sox9* is not significantly activated by well-known transcription factors for *Sox9* in *T. osimensis*. To identify novel transcription factors regulating *Sox9* through the tosEnh14, I performed biochemical analysis here. Southwestern blotting was performed to detect the proteins binding to the tosEnh14 probe, showing testis-specific. Detected samples were analyzed by LC-MS/MS, and 275 genes were identified. Referring to the results of GO analysis and the RNA-seq data in a *Nr5a1*-expressing progenitor cell at an embryonic stage, *Amhr2*, *Fubp1*, *Gtf2f1*, *Rcor3*, *Yap1*, *Bsdc1*, and *Fam118b* were focused on as the candidates. For further screening of the seven candidates, His-tagged recombinant proteins of each candidate were produced bacterially and subjected to the Southwestern blotting. As a result, the tosEnh14 probe showed specific binding to the putative recombinant protein products derived from RCOR3. Finally, the transcription activity of tosEnh14 and its potential regulation by RCOR3, were examined in COS-7 cells. As a result, two copies of tosEnh14 with RCOR3 induced a modest increase in enhancer activity, whereas a single copy of it showed significant inactivation. However, *RCOR3* knockdown did not result in significant changes in tosEnh14 activity, suggesting that RCOR3 alone may not be a major regulator of tosEnh14.

1.2 Introduction

Several *cis*-regulatory regions are spread over a gene desert of >2 Mb upstream of the *Sox9*-coding sequence[37]. Consequently, multiple enhancers of *Sox9* have been identified upstream of *Sox9* in humans and mice[38]. The testis-specific enhancer of *SOX9* (termed TES), a 3.2-kb element located 13 kb upstream of *Sox9* in mice, was first identified as a testis-specific enhancer to upregulate *Sox9*[39]. It has a 1.4-kb core sequence termed TESCO (TES core) containing binding sites for SRY, nuclear receptor subfamily 5, group A, member 1 (NR5A1, also known as SF1), and SOX9. Although the *Sox9* expression level is decreased in TESCO- or TES-knockout mice, sex reversal is not induced[40]. By further screening based on chromatin accessibility screening methods, a 557-bp element (termed Enh13) located 565 kb upstream was identified in the region homologous to human XYSR. Enh13-deletion mice show a remarkable reduction of *Sox9* expression and male-to-female sex reversal[36], providing evidence that Enh13 is an essential enhancer to initiate testicular development.

T. osimensis also has a highly conserved Enh13 sequence[41]. The nucleotide sequence identity between *T. osimensis* and mouse Enh13 (mEnh13) is 90%. All binding sites for NR5A1, SOX9, and SRY are conserved in *T. osimensis* Enh13 (tosEnh13). In a previous study, *in vitro* reporter gene assays demonstrated that tosEnh13 was substantially activated in response to co-transfection of NR5A1 and SOX9; however, activity was not significantly increased after co-transfection of NR5A1 and mouse SRY[41]. This result implies that tosEnh13 lacks SRY enhancer activity in *Sry*-deficient species.

Enh14, described in the general introduction part, is a testis-specific enhancer that is active during embryonic gonad development in mice. However, *Sox9* expression is not altered in E13.5 XY gonads of mEnh14-knockout mice, indicating that mEnh14 controls *Sox9* expression during a different developmental time window and/or functions redundantly with Enh13 and/or TES in mice[36,42]. An *in vivo* reporter assay showed that tosEnh14 drives *Sox9* expression in embryonic gonads of mice, and embryonic gonads of XX mice in which mEnh14 is replaced with duplicated tosEnh14 show increased *Sox9* expression and decreased forkhead box L2 (*Foxl2*) expression, which is a marker gene of ovarian differentiation[35]. An *in vitro* reporter gene assay was performed to examine whether tosEnh14 can be activated by mouse SRY in combination with *T. osimensis* SOX9 or *T. osimensis* NR5A1[41]. The activity of a single copy of tosEnh14 was low for all combinations. Two copies of tosEnh14 showed a 2.5-fold increase in activity compared with a single copy of tosEnh14 with a combination of NR5A1 and SOX9. However, this was significantly lower than the level required for sex determination in mouse embryos, i.e., the level at which mEnh13 is activated in response to SRY/SOX9 and NR5A1. Hence, I hypothesized that an unknown transcription factor(s) regulates *Sox9* through tosEnh14.

In addition to sex determination and testis differentiation, *Sox9* also plays a pivotal role in supporting adult Sertoli cell maintenance and spermatogenesis by maintaining *Dmrt1* expression. Deletion of *Sox9* mice showed transdifferentiation from Sertoli cells to granulosa cells, resulting male-to-female sex reversal[43]. However, no studies have investigated the regulatory function of tosEnh14 in adult testes. In this

chapter, I attempted to identify candidate proteins that bind to tosEnh14 to reveal the mechanism of *Sox9* regulation via tosEnh14. I performed Southwestern blotting using adult mouse testis extracts to screen for binding proteins to the tosEnh14. After the mass spectrometry, seven candidate genes were picked up following molecular functions. The binding assay was carried out using recombinant proteins of these genes. Lastly, transcription activity of tosEnh14 and its potential regulation by one of the identified candidate transcription factors for *Sox9* were measured in COS-7 cells.

1.3 Materials and Methods

Animal

With permission from the Agency for Cultural Affairs and the Ministry of the Environment in Japan, all animals were released at their capture sites after a small piece was cut from the tip of their tail. Genomic DNA was extracted from the tail tissues as reported previously[26,44]. C57BL/6N strain mice were used for the extraction of proteins and total RNAs. All the animal experiments were approved by the Institutional Animal Care and Use Committee of the National University Corporation, Hokkaido University, and performed following the Guidelines for the Care and Use of Laboratory Animals, Hokkaido University.

Protein extraction

Protein was extracted from a testis and a spleen of an identical adult mouse. Each tissue was homogenized in RIPA buffer (Santa Cruz Biotechnology, TX) including 1 mmol/L sodium orthovanadate in ultrapure water, 2 mmol/L PMSF in DMSO, and protease inhibitor cocktail in DMSO (NACALAI TESQUE, Kyoto, Japan), and rotated for 2 hrs at room temperature. Debris was removed by centrifugation at 10,000 × *g* for 5 mins at 4°C. The supernatants were used as protein extracts and preserved at -80°C until use. The protein concentration was measured by the BCA method using a BCA protein assay kit (FUJIFILM Wako Pure Chemical, Osaka, Japan).

Cloning of tosEnh14 and DNA probe synthesis

A fragment of the tosEnh14 amplified by PCR from genomic DNA was cloned into the pGEM-T Easy vector (Promega, Madison, WI). The nucleotide sequence was confirmed using BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, MA) with M13 primer sets by ABI Prism 3100 genetic analyzer (Thermo Fisher Scientific). The primer sequences are shown in Table 1. As the probes, the tosEnh14 and green fluorescent protein (*GFP*) were labeled DIG each with specific primer sets using PCR DIG Labelling Mix (Roche, Basel, Switzerland) according to the manufacturer's instructions. pTagGFP2-N vector (Evrogen, Moscow, Russia) was used for a PCR template to synthesize *GFP* probe. After ethanol purification of DIG-labelled PCR products, their concentration was measured by dsDNA assay using Qubit 4 Fluorometer (Thermo Fisher Scientific).

Southwestern blotting

Fifty micrograms proteins of testis and spleen extracts were subjected to SDS-PAGE with a 12%(w/v) polyacrylamide gel at a constant current of 20 mA for approximately 90 mins. These resolved proteins were transferred onto a PVDF membrane (pore size: 0.45 μm, Merck KGaA, Darmstadt, Germany). The membrane was hydrophilized in methanol for 1 min and then was treated in a blotting buffer (20% (v/v) methanol, 0.04 mol/L aminocaproic acid, 0.025 mol/L Tris). Two sheets of filter paper, which were respectively immersed in three blotting buffers (1: 10% (w/v) methanol, 0.3

mol/L Tris, 2: 10% (v/v) methanol, 0.025 mol/L Tris, 3: 20% (v/v) methanol, 0.04 mol/L aminocaproic acid, 0.025 mol/L Tris), sandwich the membrane in order of concentration of aminocaproic acid and Tris. Semi-dry transfer was performed at constant current of 65 mA for 65 mins. The membrane was blocked in N101 (NOF, Tokyo, Japan) diluted 5-fold with ultrapure water for 1 hr at room temperature. After blocking, the probe was added at a final concentration of 50 ng/3 mL to hybridize with proteins overnight at room temperature. The probe and protein were crosslinked by UV for 3 mins. The membrane was rinsed in TBS-T (137 mmol/L NaCl, 2.7 mmol/L KCl, 25 mmol/L Tris-HCl (pH7.4), 0.1% (w/v) polyoxyethylene sorbitan monolaurate (Tween-20)) and blocked for 1 hr in 5% (w/v) nonfat milk in TBS-T. Anti-Digoxigenin-POD Fab fragments (Roche) were used at dilutions of 1:3000 and incubated for 1 hr at room temperature. After the reaction, the membrane was washed in TBS-T for 5 min three times. Immobilon (Merck KGaA) was used to gain chemiluminescence and the signals were detected by LAS-3000 (FUJIFILM, Tokyo, Japan).

Mass spectrometry

SDS-PAGE was carried out with a 12% (w/v) polyacrylamide gel at a constant current of 20 mA for approximately 90 mins. After CBB staining, the protein bands were excised and purified using *In-gel* tryptic digestion kit (Thermo Fisher Scientific) following the manufacturer's instruction. These samples were analyzed by LC-

MS/MS at the Open Facility, Global Facility Center, Creative Research Institution, Hokkaido University.

GO analysis

GO analysis was performed using the DAVID 6.8 (https://david.ncifcrf.gov/). GO terms on molecular functions were obtained by functional annotation clustering. A typical annotation name was adopted in each category.

RNA extraction and cDNA synthesis

Total RNA was extracted from a testis of an adult mouse using TRIzol (Thermo Fisher Scientific) according to the manufacturer's instruction. Total RNA was reverse transcribed using PrimeScript™ Reverse Transcriptase (Takara Bio, Shiga, Japan) and an Oligo (dT) 12-18 primer (Thermo Fisher Scientific).

RT-PCR and cloning

Based on the sequence of mouse anti-Mullerian hormone type 2 receptor (*Amhr2*, NM_001356575), far upstream element binding protein 1 (*Fubp1*, NM_057172), general transcription factor IIF, polypeptide 1 (*Gtf2f1*, NM_133801), REST corepressor 3 (*Rcor3*, NM_001346740), yes-associated protein 1 (*Yap1*, NM_001171147), BSD domain containing 1 (*Bsdc1*, NM_133889), and family with

sequence similarity 118 member B (*Fam118b*, NM_001286604), specific primer sets were designed (Table 1). The amplicons were confirmed by electrophoresis in 0.8% (w/v) agarose gel containing Midori Green (NIPPON Genetics, Tokyo, Japan) to visualize DNA in TAE buffer (40 mmol/L Tris, 2.2 mmol/L EDTA・2Na, 1.2% (v/v) acetic acid). After the treatment of 10 × A-attachment (TOYOBO, Osaka, Japan) to adenylate 3' end of the amplicons, each fragment was cloned into the pGEM-T Easy vector. The nucleotide sequence of each gene was confirmed by Sanger sequence method.

Recombinant protein expression and extraction

To construct plasmids that express His-tagged recombinant proteins bacterially, full length of the coding sequence of each gene was subcloned from the pGEM-T Easy vector to the pET30a (+) vector (Merck KGaA). The fragments were digested by *Kpn*I and *Sal*I and ligated to the pET30a (+) vector. The construct was transformed into competent *Escherichia coli*, strain BL21, and a kanamycin-resistant clone was obtained. Recombinant proteins were expressed with 0.4 mmol/L-1.0 mmol/L IPTG at 37˚C for 3-5 hrs and extracted using xTractor buffer (Takara Bio) including 20 μg deoxyribonuclease I (Merck KGaA) and 1 mg lysozyme (FUJIFILM Wako Pure Chemical) following the manufacturer's instruction. After centrifugation at 10,000 × *g* for 5 mins at 4°C, supernatant was collected. Pellet was resuspended by 1%(v/v)

Triton-X in PBS (137 mmol/L NaCl, 2.7 mmol/L KCl, 8.1 mmol/L NaH_2PO_4, 1.5 mmol/L KH_2PO_4 (pH7.4)).

Reporter and expression plasmid construction

tosEnh14 fragment and two copies of tosEnh14 fragment were previously subcloned into the *Kpn*I and *Mlu*I or *Xho*I restriction site of pGL3 basic vector (Promega) containing the human beta-globin promoter[41,45,46]. Full length of coding sequences of *Rcor3* was amplified with gene-specific primers added *Kpn*I or *Xho*I restriction site. Primer sequences are shown in Table 1. The PCR products were cloned into corresponding sites of pcDNA3.1(+) (Thermo Fisher Scientific).

Reporter gene assays

COS-7 cells were plated in a 24-well microplate at a density of 0.8×10^5 cells/well and cultured in DMEM supplemented with 10% heat-inactivated FBS at 37°C in an atmosphere of 5% CO_2. The cells in each well were co-transfected with 260 ng of the enhancer reporter plasmid, several combinations of 200 ng of each expression vector, and 40 ng of firefly luciferase control reporter vectors, pRL-CMV (Promega), using Lipofectamine 3000 (Thermo Fisher Scientific). Forty-eight hours after the transfections, reporter activity was measured using the Dual-Luciferase Reporter Assay System (Promega) in accordance with the manufacturer's instructions. The reporter activity was normalized to *Renilla* luciferase activity as an internal control. Each experiment was performed in triplicate at least. The data were presented as the

mean ± standard deviation (SD). Statistical significance of the assay data was analyzed using Tukey's range tests.

shRNA design

siDirect (https://sidirect2.rnai.jp/) was used for shRNA design. Two output sequences with the fewest off-targets were selected as shRNA sequences (Table 1). These sequences were cloned into *Bam*HI and *Hin*dIII restriction sites of pBAsi-hU6 Pur DNA vector (Takara). The cloned sequences, including the target and hairpin loop sequences, were obtained by annealing sense oligo DNA and antisense sequence oligo DNA sequences (Fasmac, Kanagawa, Japan).

Cell culture and establishment of knocked down cell lines

COS-7 cells were plated in a 24-well microplate at a density of 0.8×10^5 cells/well and cultured in DMEM supplemented with 10% FBS at 37°C in an atmosphere of 5% CO_2. The cells were transfected with 500 ng of pBAsi-hU6 Pur DNA vector containing a shRNA cassette. Two days later, the transfected cells were collected and subcultured in DMEM supplemented with 10% FBS contained 1 μg/mL puromycin. Single colonies were obtained by puromycin selection.

qPCR

Total RNA was extracted from cultured cells using TRIzol (Thermo Fisher Scientific) according to the manufacturer's instruction. Total RNA was reverse transcribed using

PrimeScript™ RT reagent Kit with gDNA Eraser (Perfect Real Time) (Takara) following the manufacturer's instruction. qPCR was performed using SYBR Green qPCR Master Mix (Thermo Fisher Scientific) with the ABI 7300 Fast Real-Time PCR System (Thermo Fisher Scientific). The reaction was subjected in triplicate. The expression level was calculated by calibration curve method and then normalized to the value of *ACTB* using gene-specific primers (Table 1).

1.4 Results

Identification of 275 genes as initial candidates for factors that regulate *Sox9* through tosEnh14

Testis extracts of an adult mouse underwent Southwestern blotting to detect proteins that bind to tosEnh14 (Fig. 1A). In order to identify the specific or highly expressing components in testis, spleen extracts were used as a negative protein control from the reasons why same developmental origin from mesoderm and easier isolation of tissue. A *GFP*-encoding sequence was used as a negative DNA control to discriminate non-specific DNA-binding proteins such as deoxyribonuclease I and II[47,48]. Two bands were observed only in testis using the tosEnh14 probe. LC-MS/MS, *de novo* protein sequencing, and analysis using protein identification software (Proteome Discoverer 1.4.0.288) identified 228 and 81 genes from the higher and lower molecular weight bands, respectively (Fig. 1B). Genes counted in both the higher and lower molecular weight bands were counted as one gene, and identified 275 genes in were selected as initial candidates (Table 2).

Identification of seven genes by GO analysis

For the next step of screening, I investigated the functional characteristics of the 275 candidate genes by GO analysis. Approximately 40 out of the 275 genes were keratin related gene. The top ten GO terms in each category, namely, cellular component (CC), molecular function (MF), and biological process (BP), are shown in Fig. 2A, B,

and C, respectively. First, I selected genes associated with the GO terms “nucleus” in CC. Next, “DNA binding” in MF, and “regulation of transcription by RNA polymerase II”, “positive regulation of transcription by RNA polymerase II” was focused on but they were not ranked in 10th. Consequently, *Fubp1*, *Gtf2f1*, *Rcor3*, and *Yap1* were selected as the candidate genes for *Sox9* regulator. In addition to that, *Amhr2*, encodes a receptor of AMH, was included in the candidates as it is involved in sex determination and differentiation in several species of fish. Especially, it has been reported as a sex-determining gene in *Takifugu rubripes*[49]. *Bsdc1* was unknown function by GO analysis, therefore they were not excluded as the candidates. Although *Fam118b* was not annotated DNA-binding, it is reported to have an ability to bind with YAP1[50] which is one of the candidates in this study. In summary, a total of seven genes, *Amhr2*, *Fubp1*, *Gtf2f1*, *Rcor3*, *Yap1*, *Bsdc1*, and *Fam118b*, were selected (Table 3).

RCOR3 specifically bound to tosEnh14

To confirm that the proteins of 7 candidate genes bind to the tosEnh14, recombinant proteins of each candidate were produced by the bacterial expression system and subjected to the Southwestern blotting. FUBP1 and YAP1 recombinant proteins failed to express in the *E. coli*, unfortunately. The results of AMHR2, GTF2F1, RCOR3, BSDC1, and FAM118B were shown in Fig. 3A. The tosEnh14 probe showed the binding to the putative recombinant protein products derived from GTF2F1 and

RCOR3 (a band near 70 kDa indicated by an arrowhead in Fig. 3A). GTF2F1 showed binding to the *GFP*-encoding probe indicating non-specific binding. On the other hand, the band derived from RCOR3 was not observed using the *GFP*-encoding sequence probe. In addition, I confirmed that 6 × Histidine antibody (Bioss, MA) recognized a 70 kDa band by Western blotting analysis (Fig. 3B). Some bands at small molecular sizes seemed to appear by protein degradation. These results indicate that RCOR3 is a final candidate which binds to tosEnh14.

Binding sites for RCOR3 are present in tosEnh14

RCOR3 contains a switching-defective protein 3, adaptor 2, nuclear receptor co-repressor, and transcription factor IIIB (SANT) /Myb domains which can be classified into three groups: DNA-binding domain, protein-protein interaction module, and the domain that can be involved in either of these functions[51]. The Myb DNA-binding domain specifically recognizes the sequence YAAC(G/T)G[52,53], where Y represents a pyrimidine base. Although tosEnh14 does not have the same sequence, partially similar sequences are present in tosEnh14 (Fig. 4). This suggests that RCOR3 directly binds to tosEnh14 and regulate *Sox9*.

RCOR3 induced difference activity depend on copy number of tosEnh14

To investigate the transcriptional activity of tosEnh14 and its potential regulation by RCOR3, a luciferase reporter assay was performed in COS-7 cells. Co-transfection

with RCOR3 significantly reduced enhancer activity when a single copy of tosEnh14 was introduced, compared to the condition without RCOR3 (Fig. 5A). However, when two copies of tosEnh14 were co-transfected, luciferase activity was significantly promoted by RCOR3 compared to condition without it and the single copy of tosEnh14. These results suggest that the effect of RCOR3 on tosEnh14 activity may depend on enhancer dosage, showing a repressive effect at low levels and a potential activating role when enhancer copies are increased.

tosEnh14 was not activate by RCOR3 in the knocked down cell line

To further analyze the role of RCOR3 in regulating tosEnh14 activity, stable *RCOR3* knockdown cell lines were generated using shRNA constructs. Considering that transfection efficiency may not reach 100%, shRNA-expressing stable clones were established to ensure consistent and long-term suppression of *Rcor3* expression. More than ten cell colonies were obtained. After preliminary screening by PCR, four clones were evaluated to assess knocked down efficiency by qRT-PCR (Fig. 5B). Among the established clones, 4C and 3A showed decreases in *RCOR3* mRNA levels compared to the empty vector control. Clone 4C demonstrated a reduction in *RCOR3* expression of approximately 46%, indicating effective knockdown. In contrast, clones 5D and 1B exhibited expression levels similar to the control, suggesting that knockdown was insufficient. Despite being transfected with the same shRNA, the knockdown

efficiency of Clone 4C and 5D was drastically different. After all, clone 4C was chosen as an *RCOR3* knocked down cell line.

A luciferase assay was performed in the *RCOR3* knockdown cell line to evaluate the effect of suppressing endogenous *RCOR3* on transcriptional activity of tosEnh14. The shRNAs were designed to target only endogenous *RCOR3* derived from green monkey and not affect exogenously introduced mouse *Rcor3* mRNA. Interestingly, *RCOR3* knockdown did not lead to any significant changes in any of the tested conditions, regardless of the copy number of tosEnh14 (Fig. 5C). This result suggests that endogenous *RCOR3* alone is not sufficient to activate tosEnh14, indicating that additional co-factors may be required for its regulatory function.

1.5 Discussion

Several methods are available to detect DNA–protein interactions, including ChIP and EMSA[54]. However, these techniques often require information of a target protein and are mainly used to identify the DNA regions to which the protein binds. In contrast, methods that allow screening for unknown proteins that bind to a specific DNA sequence are relatively limited. In this study, Southwestern blotting approaches were employed for screening a large number of potential transcription factors. Here, I showed a possibility that RCOR3 is a transcription factor that binds to the tosEnh14 by Southwestern blotting analyses. RCOR3 was identified as a factor that specifically binds to tosEnh14, and its binding was experimentally confirmed, supporting the validity of this approach. On the other hand, despite initially identifying approximately 275 candidate genes, only one remained as a final candidate. This significant reduction may reflect a limitation that proteins are hybridized to DNA under denaturing conditions, which may hinder the detection of proteins that require intact tertiary structures for DNA binding.

The identification of transcription factors involved in gene regulation is a crucial step in understanding the molecular mechanisms underlying specific biological processes. Given its identification as a potential transcriptional regulator for tosEnh14, the regulatory role of RCOR3 was investigated in this chapter. Interestingly, the reporter assay indicated that the transcriptional activity of tosEnh14 with RCOR3 varied depending on the copy number (Fig. 5A). These results may reproduce the difference of *Sox9* expression levels between males and females that

have duplicated tosEnh14 or single copy of it in *T. osimensis*. These findings raise the possibility that tosEnh14 may exhibit copy-number-dependent dual functionality, functioning as a silencer or inactive element at low copy number, and as an active enhancer only when present in higher numbers. This phenomenon is consistent with previous reports of dual-function *cis*-regulatory elements, whose transcriptional roles shift depending on genomic context, chromatin state, and dosage. For example, insulator elements can block enhancer-promoter communication at low copy numbers but act as activators at higher copy numbers in the *Drosophila*[55,56]. In the case of *T. osimensis*, male-specific duplications of tosEnh14 could release from repression of *Sox9*. However, the results did not provide compelling evidence to support a direct regulatory effect of RCOR3 on tosEnh14. Although a modest increase in enhancer activity was observed with two copies of tosEnh14 and RCOR3 present, this effect was not replicated when *RCOR3* was knocked down.

There is a possibility that RCOR3 requires additional co-factors in order to effectively exert its regulatory function. Actually, tosEnh14 activity is more reinforced by both NR5A1 and SOX9 than by either of them alone[41]. Previous studies have reported that RCOR3 acts as a transcriptional co-repressor alongside other transcription factors and chromatin modifiers[57]. RCOR3 consists of an Egl-27 and an MTA1 homology 2 (ELM2) domain, a SANT/Myb domain, and repressor element-1-silencing transcription factor (REST) corepressor helical domain. The function of the ELM2 domain has been unknown. The metastasis associated 1 (MTA1) protein is a part of the complex that has ATP-dependent nucleosome remodeling and histone

deacetylation activities[58] and is usually found at the N terminus of a myb-like DNA-binding domain. These reports suggest that the ELM2 domain is involved in protein-protein interaction or DNA-binding. The SANT domain is found in nuclear receptor co-repressor and the subunits of many chromatin remodeling complexes[59]. Nuclear receptor co-repressor 1 (NCOR1), which is also known to have the SANT domain, was reported to have histone deacetylase regulator, nuclear receptor binding, and transcription corepressor ability[60]. The structure of the SANT is very similar to the Myb DNA-binding domain[61]. Both domains consist of tandem repeats of three alpha-helices, which are arranged in a helix-turn-helix motif, each alpha helix containing a bulky aromatic residue[59,62,63].

RCOR3 belongs to REST corepressor (CoREST) family. REST is an essential vertebrate zinc finger transcription factor known to suppress gene expression in multiple cell types[64,65]. The transcription silencing effect of REST depends on CoRESTs as its transcriptional corepressors[66]. However, RCOR3 has also been demonstrated to act as a transcriptional activator[57]. RCOR3 exerts the positive effect by competitively inhibiting its paralogs, REST corepressor 1 (RCOR1) and REST corepressor 2 (RCOR2), which form repressive complexes with lysine-specific demethylase I (LSD1), also known as *Kdm1a*. Under normal conditions, LSD1–RCOR1/2 complexes catalyze the removal of mono- and dimethyl groups from histone H3 lysine 4 (H3K4me1/2)[67]. This modification is enriched at active enhancers and promoters and is generally associated with active transcription[68]. By binding to LSD1 and displacing RCOR1/2, RCOR3 effectively blocks this demethylase activity,

thereby preserving H3K4 methylation marks at target *cis*-regulatory elements. Although it is unknown that these modifications occurred in testis as well, LSD1 and the RCOR family proteins are expressed in testis[69,70].

Taken together with these functions and a result of binding to DNA (Fig. 3), RCOR3 might be relevant at *Sox9* expression by tosEnh14. If RCOR3 can be recruited and directly bind to tosEnh14, H3K4me1/2 marks at tosEnh14 remain, thereby promoting an open chromatin environment conducive to *Sox9* activation. Alongside other co-activators such as NR5A1 or SOX9 itself, the preservation of H3K4 methylation mediated by RCOR3 may therefore contribute to the balance of *Sox9* expression in retention of adult testis. Moreover, a possibility remains that other transcription factor candidates identified in mass spectrometry may work as a primary regulator of *Sox9*, rather than RCOR3. In conclusion, while the current results do not definitively confirm RCOR3 as a direct enhancer activator, they suggest its potential involvement in the regulation of chromatin states. Therefore, further investigation is warranted to identify other primary regulators of tosEnh14 and to clarify the functional partners or indirect roles of RCOR3.

1.6 Tables

Table 1 Primers or oligo DNAs used in this chapter.

Target Sequence/gene	Experiment	Sequence 5' to 3' (Sense/ Antisense)
pUC/M13	Cycle sequencing	AGGGTTTTCCCAGTCACGAC
		TCACACAGGAAACAGCTATGAC
tosEnh14	Southwestern blotting	GGATAGAGAAGCTTGAAATGC
		TTGCATGCTTTCAGGTGCTT
GFP	Southwestern blotting	ATGGTGAGCAAGGGCGAGGA
		TTACTTGTACAGCTCGTCCA
Amhr2	RT-PCR Southwestern blotting	GGTACCATGCTGGGGACACTGGGGC
		GTCGACCTCATTTACATACACCTGAACAG
Fubp1	RT-PCR	GGTACCATGGCCGACTACTCCACAGTG
		GTCGACTTGGCCCTGAGGTGCTGGAG
Gtf2f1	RT-PCR Southwestern blotting	GGTACCATGGCGGCTCTAGGCTCCAG
		GTCGACCTCTTTGAGGGAGAAGTGCATC
Rcor3	RT-PCR Southwestern blotting	GGTACCATGCCCGGCATGATGGAGAAAG
		GTCGACGCAGGGCGAATGAGAGGTGG
	Reporter gene assay	GGTACCATGCCCGGCATGATGGAGAAAG
		CTCGAGTTAGTGCAGTGAAGGCTGTGAATC
RCOR3	shRNA #1	GATCCA**CGCCTTTCTAGAGTATATGTTCA**CTGTGAAGCCA
		CAGATGGG**TGAACATATACTCTAGAAAGGCG**TTTTTTTA
		AGCTTAAAAAAA**CGCCTTTCTAGAGTATATGTTCA**CCCAT
		CTGTGGCTTCACAG**TGAACATATACTCTAGAAAGGCG**TG
	shRNA #2	GATCCG**CACTTAAATTGTCTTACATATGT**CTGTGAAGCCA
		CAGATGGG**ACATATGTAAGACAATTTAAGTG**CTTTTTTA
		AGCTTAAAAAAG**CACTTAAATTGTCTTACATATGT**CCCAT
		CTGTGGCTTC**ACAGACATATGTAAGACAATTTAAGTG**CG
	qRT-PCR	CGGAGCCGAATACCAAGCTC
		TGGAGACCATACAAGCATCCC
Yap1	RT-PCR	GGTACCATGGAGCCCGCGCAACAGC
		GTCGACTAACCACGTGAGAAAGCTTTC
Bsdc1	RT-PCR Southwestern blotting	GGTACCATGGCGGAAGGGGAGGACGTG
		GTCGACTTCCCAGTCCTCCCACTCTACATC
Fam118b	RT-PCR Southwestern blotting	GGTACCATGGCTTCTACAGGGAGCCAG
		GTCGACAGTACTACAGCCTCTTATTTCACC
ACTB	qRT-PCR	GATCAGCAAGCAGGAGTATGACG
		AAGGGTGTAACGCAACTAAGTCATAG

Bold sequences indicate the shRNA sequences.

Table 2-1 Genes identified from an adult testis by LC-MS/MS.

Upper band	
Accession	Gene symbol
Q9D454	*4933412E24Rik*
Q8BMP6	*Acbd3*
Q3V140	*Acrbp*
P63268	*Actg2*
Q5SUE7	*Adad1*
Q3UMC0	*Afg2a*
Q60662	*Akap4*
P07724	*Alb*
P09242	*Alpl*
Q8K592	*Amhr2*
P50516	*Atp6v1a*
Q8C3R1	*Brat1*
Q80Y55	*Bsdc1*
P06683	*C9*
Q8C633	*Cabs1*
P14211	*Calr*
P35564	*Canx*
O08529	*Capn2*
Q80ZU5	*Ccdc181*
Q99LM2	*Cdk5rap3*
Q8BI72	*Cdkn2aip*
Q497N6	*Cep295nl*
P23953	*Ces1c*
Q8BHB9	*Clic6*
Q6NVF9	*Cpsf6*
P63154	*Crnkl1*
Q8BIQ5	*Cstf2*
Q8C7E9	*Cstf2t*
Q99LI7	*Cstf3*
Q9CWL8	*Ctnnbl1*
P70698	*Ctps1*
Q60598	*Cttn*
P16381	*D1Pas1*
Q501J6	*Ddx17*
Q62167	*Ddx3x*
Q62095	*Ddx3y*
Q61496	*Ddx4*
Q61656	*Ddx5*
Q9JMC3	*Dnaja4*
Q8K1M6	*Dnm1l*
Q9QXP4	*Donson*
Q9DA79	*Dpep3*
Q99KK7	*Dpp3*
Q8BGD9	*Eif4b*
Q03173	*Enah*
O55026	*Entpd2*

Lower band	
Accession	Gene symbol
Q3V140	*Acrbp*
P63268	*Actg2*
P45376	*Akr1b1*
P21300	*Akr1b7*
P45377	*Akr1b8*
P70694	*Akr1c6*
P07724	*Alb*
P07356	*Anxa2*
Q9WV54	*Asah1*
Q9DB16	*Cab39l*
Q06890	*Clu*
P06797	*Ctsl*
Q9QZL9	*Dkkl1*
Q99LN9	*Dohh*
E9Q557	*Dsp*
P57776	*Eef1d*
Q3UGC7	*Eif3j1*
Q66JS6	*Eif3j2*
P17182	*Eno1*
Q8C569	*Fam118b*
P16858	*Gapdh*
O70252	*Hmox2*
O88569	*Hnrnpa2b1*
P51660	*Hsd17b4*
P08113	*Hsp90b1*
Q61696	*Hspa1a*
P17879	*Hspa1b*
P16627	*Hspa1l*
P20029	*Hspa5*
P63017	*Hspa8*
Q02257	*Jup*
P04104	*Krt1*
P02535	*Krt10*
Q61781	*Krt14*
Q61414	*Krt15*
Q9Z2K1	*Krt16*
Q9QWL7	*Krt17*
P05784	*Krt18*
P19001	*Krt19*
Q3TTY5	*Krt2*
Q8VCW2	*Krt25*
Q9Z320	*Krt27*
A6BLY7	*Krt28*
P07744	*Krt4*
Q6IFX2	*Krt42*
Q922U2	*Krt5*

Table 2-2 Genes identified from an adult testis by LC-MS/MS.

Upper band	
Accession	Gene symbol
Q8CHU3	*Epn2*
Q9D9V2	*Eqtn*
Q61545	*Ewsr1*
Q66LM6	*Fam170a*
Q9WUA2	*Farsb*
Q8VDH1	*Fbxo21*
Q9D824	*Fip1l1*
P50608	*Fmod*
Q91WJ8	*Fubp1*
P56959	*Fus*
Q61584	*Fxr1*
P97855	*G3bp1*
P97379	*G3bp2*
P16858	*Gapdh*
Q64467	*Gapdhs*
Q9CZD3	*Gars1*
P97494	*Gclc*
P03995	*Gfap*
Q3THK7	*Gmps*
P36916	*Gnl1*
Q3THK3	*Gtf2f1*
Q8BMS1	*Hadha*
Q8R081	*Hnrnpl*
Q8BIL5	*Hook1*
Q91X72	*Hpx*
P07901	*Hsp90aa1*
P11499	*Hsp90ab1*
P08113	*Hsp90b1*
Q8K0U4	*Hspa12a*
Q61696	*Hspa1a*
P17879	*Hspa1b*
P16627	*Hspa1l*
P48722	*Hspa4l*
P20029	*Hspa5*
P63017	*Hspa8*
P38647	*Hspa9*
Q8BKE9	*Ift74*
Q8BKC5	*Ipo5*
Q99MN1	*Kars1*
Q9D3R6	*Katnal2*
Q60749	*Khdrbs1*
E9Q0C6	*Kiaa1210*
O88448	*Klc2*
Q9DBS5	*Klc4*
P04104	*Krt1*
P02535	*Krt10*

Lower band	
Accession	Gene symbol
P50446	*Krt6a*
Q9Z331	*Krt6b*
Q9R0H5	*Krt71*
Q6IME9	*Krt72*
Q6NXH9	*Krt73*
Q6IFZ9	*Krt74*
Q8BGZ7	*Krt75*
Q3UV17	*Krt76*
Q6IFZ6	*Krt77*
Q8VED5	*Krt79*
P11679	*Krt8*
Q0VBK2	*Krt80*
Q99M74	*Krt82*
Q99M73	*Krt84*
P06151	*Ldha*
P16125	*Ldhb*
P00342	*Ldhc*
Q9DBH5	*Lman2*
Q9DAP0	*Lrrc46*
P70670	*Naca*
Q60817	*Naca*
Q61937	*Npm1*
Q9CRD0	*Ociad1*
Q9D051	*Pdhb*
Q9D819	*Ppa1*
Q9DCL8	*Ppp1r2*
Q91WE2	*Psme3ip1*
Q91YR9	*Ptgr1*
P14869	*Rplp0*
Q3URD3	*Slmap*
Q9ER00	*Stx12*
Q80VL1	*Tdrkh*
Q9JMI7	*Tex101*
Q8CDN6	*Txnl1*
Q9WUP7	*Uchl5*

Table 2-3 Genes identified from an adult testis by LC-MS/MS.

Upper band	
Accession	Gene symbol
P08730	*Krt13*
Q61781	*Krt14*
Q61414	*Krt15*
Q9Z2K1	*Krt16*
Q9QWL7	*Krt17*
P05784	*Krt18*
P19001	*Krt19*
Q3TTY5	*Krt2*
A1L317	*Krt24*
Q8VCW2	*Krt25*
Q9Z320	*Krt27*
A6BLY7	*Krt28*
Q61765	*Krt31*
Q62168	*Krt32*
Q61897	*Krt33b*
Q497I4	*Krt35*
B1AQ75	*Krt36*
Q6IFX3	*Krt40*
Q6IFX2	*Krt42*
Q922U2	*Krt5*
P50446	*Krt6a*
Q9Z331	*Krt6b*
Q9R0H5	*Krt71*
Q6NXH9	*Krt73*
Q6IFZ9	*Krt74*
Q8BGZ7	*Krt75*
Q3UV17	*Krt76*
Q6IFZ6	*Krt77*
Q8VED5	*Krt79*
Q9Z2T6	*Krt85*
P11438	*Lamp1*
Q05CL8	*Larp7*
P48678	*Lmna*
P14733	*Lmnb1*
Q505F5	*Lrrc47*
Q99KG5	*Lsr*
P51885	*Lum*
A2A9R3	*Mageb4*
P26645	*Marcks*
Q8R3C0	*Mcmbp*
C0HKD8	*Mfap1a*
P16332	*Mmut*
P26041	*Msn*
Q9R190	*Mta2*
Q80WJ7	*Mtdh*
Q99MD9	*Nasp*

Table 2-4 Genes identified from an adult testis by LC-MS/MS.

Upper band	
Accession	Gene symbol
P09405	*Ncl*
Q91VD9	*Ndufs1*
P32507	*Nectin2*
Q8BG30	*Nelfa*
Q91YP2	*Nln*
P60670	*Nploc4*
P46460	*Nsf*
Q5NCR9	*Nsrp1*
Q61503	*Nt5e*
A3KGV1	*Odf2*
Q8K3K8	*Optn*
P29341	*Pabpc1*
Q91ZA3	*Pcca*
Q64338	*Pde1c*
P27773	*Pdia3*
P08003	*Pdia4*
Q8K124	*Plekho2*
Q61074	*Ppm1g*
Q9QUR6	*Prep*
O08795	*Prkcsh*
Q8R326	*Pspc1*
Q8BSI6	*R3hcc1*
Q91WG2	*Rabep2*
Q9CT10	*Ranbp3*
Q8C2Q3	*Rbm14*
Q6PGA0	*Rcor3*
P26043	*Rdx*
P62983	*Rps27a*
Q9D394	*Rufy3*
Q8CBY1	*Samd4a*
Q60710	*Samhd1*
Q80UG5	*Septin9*
P29621	*Serpina3c*
Q80X76	*Serpina3f*
Q5I2A0	*Serpina3g*
P07759	*Serpina3k*
Q03734	*Serpina3m*
Q91WP6	*Serpina3n*
Q61247	*Serpinf2*
Q64213	*Sf1*
Q62203	*Sf3a2*
Q7TSG5	*Sh3d21*
P10852	*Slc3a2*
Q62376	*Snrnp70*
Q9WV80	*Snx1*
Q91ZR2	*Snx18*

Table 2-5 Genes identified from an adult testis by LC-MS/MS.

Upper band	
Accession	Gene symbol
Q9CWK8	*Snx2*
Q3V0Q6	*Spag8*
Q80YT5	*Spata20*
Q8K1N4	*Spats2*
Q9D5A0	*Spesp1*
Q8VE97	*Srsf4*
Q8VC66	*Ssx2ip*
Q7TMK9	*Syncrip*
Q9D3R9	*Taf7l*
Q9CXF4	*Tbc1d15*
Q80VL1	*Tdrkh*
P97499	*Tep1*
Q921I1	*Tf*
Q62313	*Tgoln1*
Q62314	*Tgoln2*
Q8BH55	*Thnsl1*
Q8C1A5	*Thop1*
Q9CZW5	*Tomm70*
A2AI08	*Tprn*
Q9CQN1	*Trap1*
O54887	*Tsks*
Q3UDE2	*Ttll12*
P68369	*Tuba1a*
Q9JJZ2	*Tuba8*
Q9ERD7	*Tubb3*
P46686	*Tulp2*
Q6PAM1	*Txlna*
Q6P902	*Txndc2*
P0CG49	*Ubb*
P0CG50	*Ubc*
Q8R317	*Ubqln1*
Q9QZM0	*Ubqln2*
Q8C5U9	*Ubqln3*
Q99NB8	*Ubqln4*
Q6P5G6	*Ubxn7*
Q3TLD5	*Uri1*
Q9D2J4	*Vsig1*
P29788	*Vtn*
Q80ZK9	*Wdtc1*
Q6P1B1	*Xpnpep1*
P23475	*Xrcc6*
P46938	*Yap1*
Q9Z2C8	*Ybx2*
Q8BJ05	*Zc3h14*

Table 3 Final candidate genes and their GO terms.

			GO term in BP		
Gene symbol	Gene name	Gene ID	regulation of transcription by RNA polymerase II	positive regulation of transcription by RNA polymerase II	Note
Amhr2	anti-Mullerian hormone type 2 receptor	110542			Sex determination gene in some fish
Fubp1	far upstream element (FUSE) binding protein 1	51886	✓		
Gtf2f1	general transcription factor IIF, polypeptide 1	98053		✓	
Rcor3	REST corepressor 3	214742	✓		
Yap1	yes-associated protein 1	22601		✓	
Bsdc1	BSD domain containing 1	100383			Unknown function
Fam118b	family with sequence similarity 118, member B	109229			Interaction with YAP1

1.7 Figures

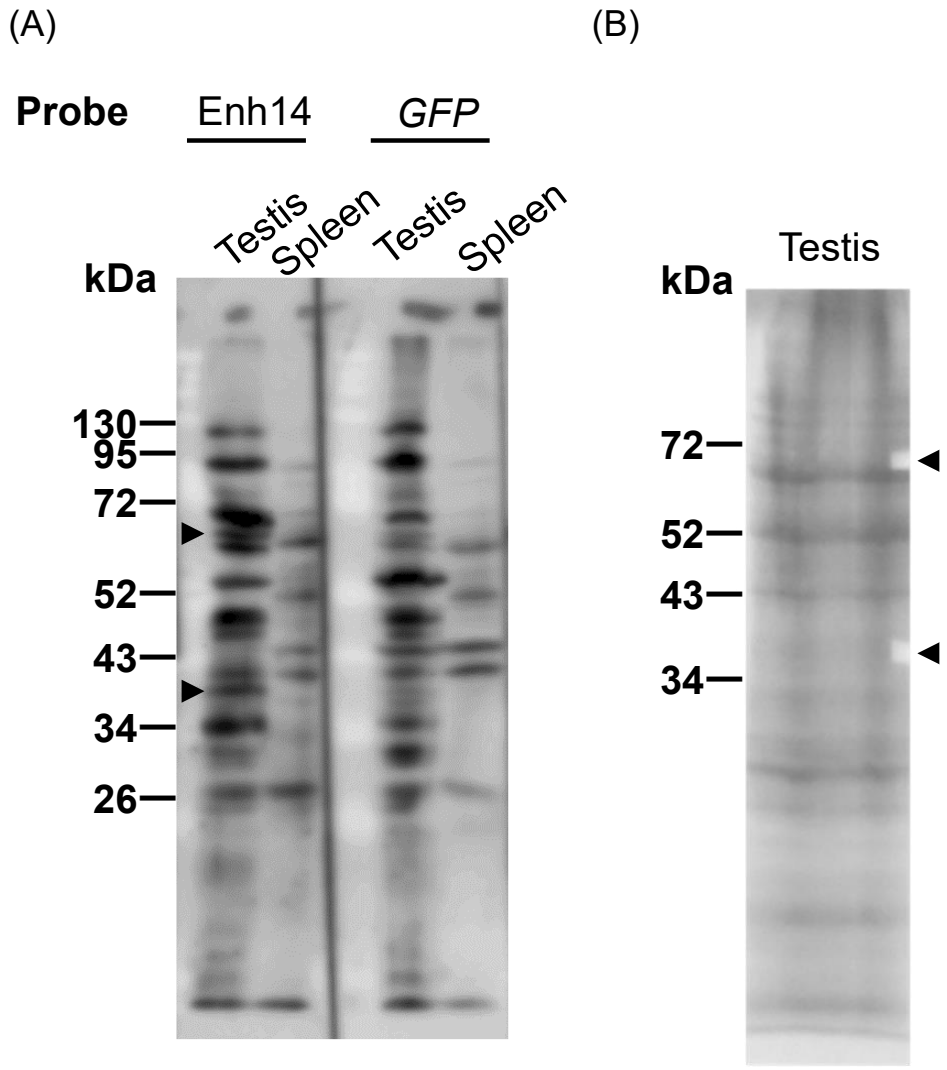

Fig. 1 Southwestern blotting analysis for screening proteins which bind to the tosEnh14

(A) Testis and spleen extracts were subjected to SDS-PAGE and blotted onto the PVDF membrane. DIG-labeled tosEnh14 and *GFP*-encoding sequence were used as the probe. Signals were detected using an anti-DIG antibody. Two arrowheads indicate testis-specific protein binding to the tosEnh14 probe. (B) CBB staining for LC-MS/MS. Each band marked by an arrowhead was excised and subjected to LC-MS/MS.

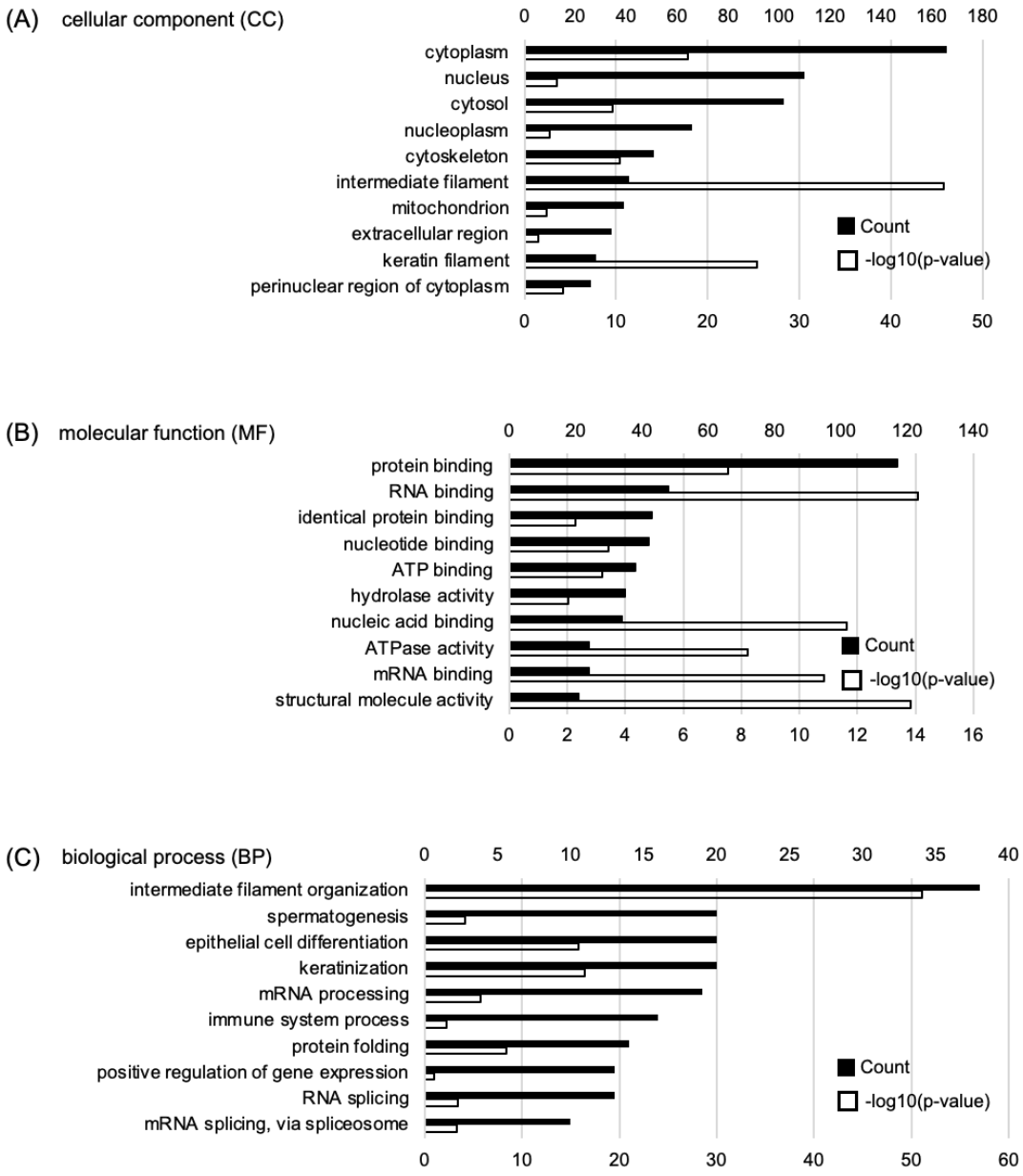

Fig. 2 GO analysis of genes identified by mass spectrometry.

(A) BP, (B) CC, and (C) MF. The top ten GO terms are shown.

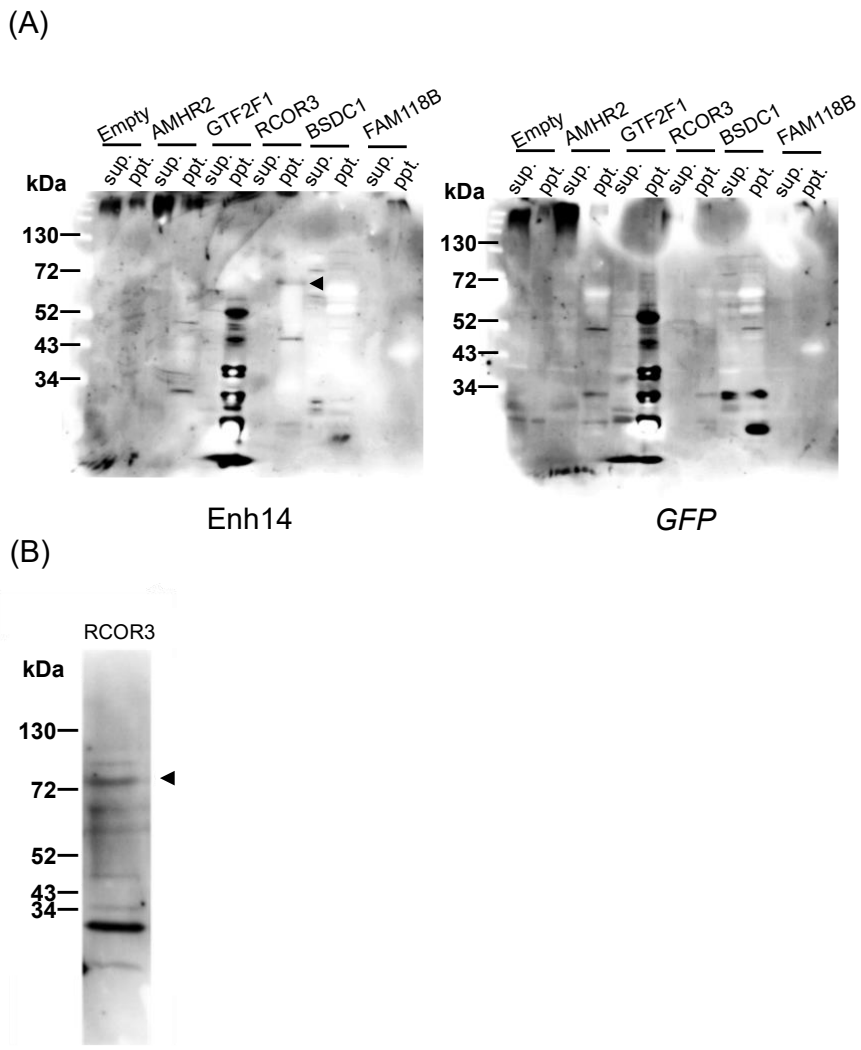

Fig. 3 Southwestern blotting analysis for the binding check to the tosEnh14.

(A) The extracts from *E. coli* after IPTG treatment for 5 hrs were subjected to SDS-PAGE and blotted onto the PVDF membrane. The left and right panel shows the results of using DIG-labeled tosEnh14 and *GFP*-encoding probes. Signals were detected using an anti-DIG antibody. An arrowhead indicates the protein band interacted with the tosEnh14 probe specifically. (B) Western blotting analysis using a 6 × Histidine antibody detected the recombinant RCOR3 proteins. An arrowhead indicates the protein band of full-length RCOR3.

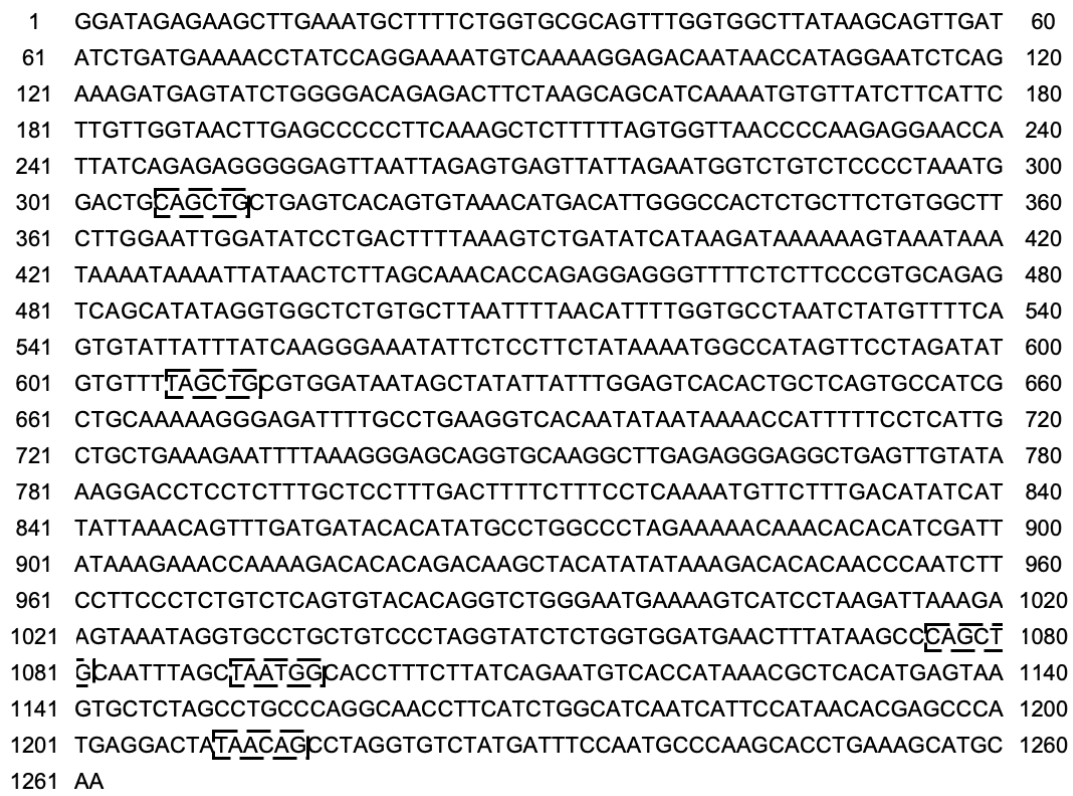

Fig. 4 Putative binding sites for RCOR3 in tosEnh14.

Putative binding sites for Myb DNA-binding like domains of RCOR3 are indicated by dashed frames.

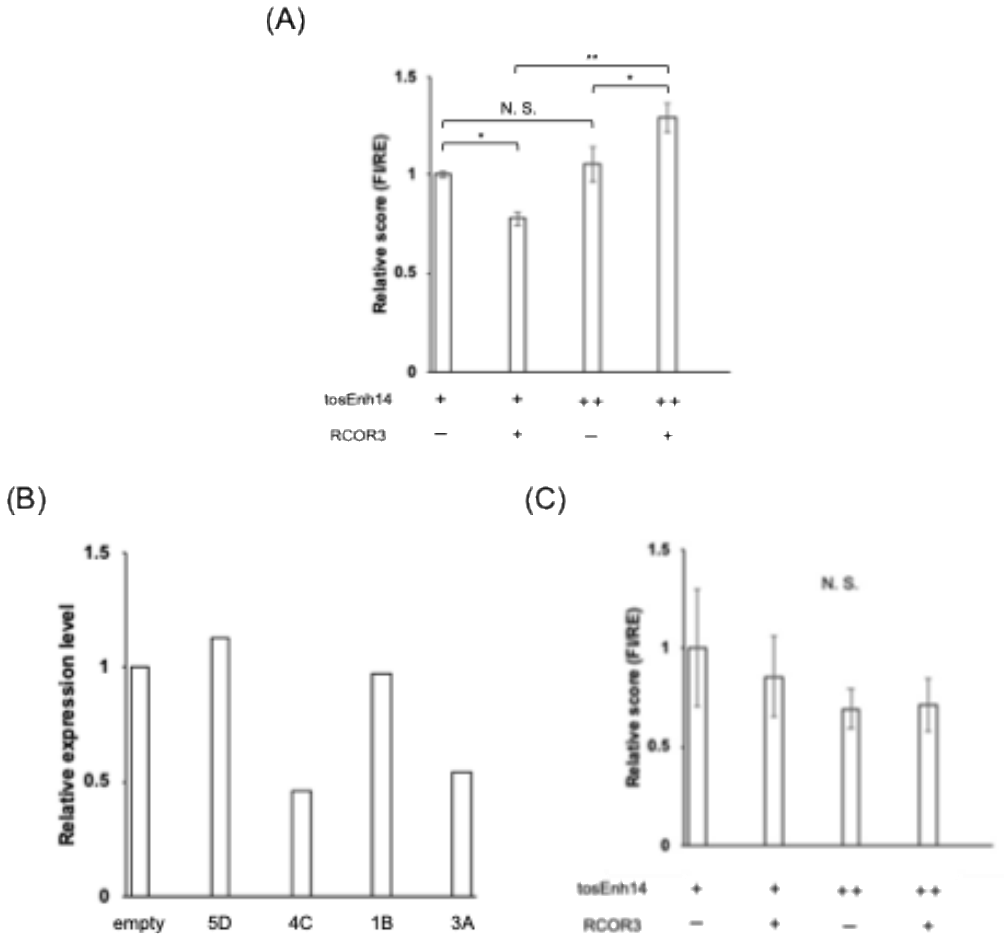

Fig. 5 Relative luciferase activity of Enh14 with or without RCOR3 in COS-7 cells.

(A) A luciferase assay was performed using the single copy of tosEnh14 (+) or two copies of tosEnh14 (++) reporter vectors. Data are shown as mean ± SEM (n = 3). * $p < 0.05$. ** $p < 0.01$. (B) Relative RCOR3 mRNA expression levels in COS-7 cells transfected with shRNA constructs or empty vector control, as determined by qRT-PCR. 5D, 4C, 1B, and 3A mean each clone name. (C) A luciferase assay was performed using the single copy of tosEnh14 (+) or two copies of tosEnh14 (++) reporter vectors in knocked down cells 4D. Data are shown as mean ± SEM (n≥3).

Chapter II

Identification of candidate transcription factors expressing in embryonic gonads

2.1 Abstract

I explored novel transcription factors that regulate *Sox9* via tosEnh14 using mouse embryonic samples. To detect proteins that bind to tosEnh14 DNA, Southwestern blotting analysis was performed using mouse embryonic gonad extracts. Bands of a similar molecular weight but prominent in males and faint in females were subjected to mass spectrometry analysis. Peptides derived from 174 genes were identified, and eight genes associated with GO terms such as "DNA binding" and "regulation of transcription by RNA polymerase II" were selected. For further screening, the expression level of each gene was examined using scRNA-seq data of mouse progenitor cells, which differentiate into Sertoli cells in mouse embryonic testes and granulosa cells in embryonic ovaries. Finally, five genes (*Elf2*, *Etv6*, *Fiz1*, *Gtf2f1*, and *Trim27*) whose expression was confirmed in seminiferous tubules of embryonic day 13.5 (E13.5) XY embryos by whole-mount *in situ* hybridization were selected as candidates. Hence, the five genes may regulate *Sox9* through tosEnh14 in embryonic gonads.

2.2 Introduction

SRY binds to the enhancers of *Sox9* and synergistically upregulates *Sox9* expression with NR5A1[39]. An additional factor, GATA-binding protein 4 (GATA4), interacts with NR5A1 and zinc finger protein FOG family member 2 (FOG2) to regulate expression of *Sry* and *Sox9*[71,72]. After SRY expression disappears, SOX9 binds to its enhancer in order to maintain its expression via self-regulation[39].

In humans, the *SOX9* enhancer has been found in patients with DSD patients who have a complete *SOX9*-coding sequence but mutations in the upstream region. Referring to genetic information of a large number of patients, overlapping regions have been defined as potentially regulatory elements. To functionally validate these candidate regulatory regions, cell-based reporter gene assays have been employed to demonstrate enhancer activity. This integrative approach combining human genetic data and experimental enhancer assays has provided strong evidence for the existence of long-range enhancers that regulate *SOX9* expression and contribute to human sex development. Actually, CNV mapping in the genomes of DSD patients with a *SRY*-positive 46,XY karyotype identified a 32.5-kb region located 600 kb upstream of *SOX9*, called XYSR[73]. Subsequent analysis of DSD patients carrying CNVs refined this to a small 5.2-kb region[46]. Additionally, a 68-kb region ~600 kb upstream of *SOX9* called XXSR (also known as RevSex) was identified by screening of CNVs in DSD patients with a *SRY*-negative 46,XX karyotype[73-75] and further refined it to 24 kb[46]. By screening of additional patient genomic and bioinformatic data, three putative enhancers were identified: eSR-A in XYSR, eSR-B in RevSex, and eALDI[46].

By cell-based reporter assays, all three enhancers showed synergistic activity with different combinations of testis-specific regulators including SRY, NR5A1, and SOX9.

In mice, the process of sex determination occurs during a narrow window of development in a fetal gonad. This begins around embryonic day (E) 10.5 and becomes morphologically evident by E12.5. Bipotential gonads, which is derived from the genital ridge, has a capacity to differentiate into testes or ovaries depending on genetic sex[76,77]. In XY embryos, *Sry* is transiently activated at E10.5 and peaks around E11.5 in the pre-Sertoli cells of the developing testis, and then disappeared at E12.0[78,79]. Only *Sry*-positive cells in XY gonads express high levels of *Sox9*, subsequently differentiating into Sertoli cells[80,81]. Following the loss of *Sry* expression, SOX9 maintains its own expression through a positive feedback loop[39]. SOX9 activates the downstream targets required for testis differentiation such as *Amh*[22], fibroblast growth factor 9 (*Fgf9*)[82], and desert hedgehog (*Dhh*)[83]. By contrast, in XX embryos where *Sry* is absent, *Sox9* is not activated. Consequently, the supporting cells differentiate into granulosa cells, leading to ovarian development. This process is promoted by the expression of ovarian development factors, such as wingless-type MMTV integration site family, member 4 (*Wnt4*), R-spondin 1 (*Rspo1*) and *Foxl2*, which repress testis development pathways[84,85].

In this chapter, mouse embryonic gonads extracts were used to explore unknown transcription factors that bind to the tosEnh14 DNA sequence at sex determination period. I also performed Southwestern blotting and subsequently mass spectrometry

analysis, which identified 174 candidate genes. Further screening involved GO analysis and expression analysis using sets of RNA-seq data of progenitor cells, which differentiate into Sertoli cells in mouse embryonic testes. To further validate expression localization of candidate genes, I confirmed expression of each gene in E13.5 XY mouse gonads by WISH. Finally, five genes with expression patterns resembling that of *Sox9* were identified as candidates for transcription factors that regulate *Sox9* through tosEnh14.

2.3 Materials and methods

Animals and embryo staging

B6D2F1/Jcl mice were used for protein and total RNA extraction and WISH. All mice were purchased from a local vendor (Hokudo, Sapporo, Japan). Timed mating was used for all embryonic experiments. E0.5 was defined as noon on the day of vaginal plug detection.

Cloning of tosEnh14 and DNA probe synthesis

I followed the method in Chapter I.

Southwestern blotting

Proteins were extracted from four gonads and mesonephros of E13.5 mice with RIPA buffer (Merck KGaA) containing 1 mmol/L sodium orthovanadate prepared in ultrapure water, 2 mmol/L PMSF prepared in DMSO, and a protease inhibitor cocktail prepared in DMSO (NACALAI TESQUE). Twenty micrograms of protein was subjected to SDS-PAGE with a 10% (w/v) polyacrylamide gel at a constant current of 20 mA for approximately 90 min. Subsequent processes were described in Chapter I.

Mass spectrometry analysis

I followed the method in Chapter I.

GO analysis

GO analysis was performed using DAVID (v2024q2) (https://davidbioinformatics.nih.gov/). For other procedures, I followed the method in Chapter I.

scRNA-seq analysis

Data from the NCBI BioProject (accession no. GSE97519) were used for the analysis in XY[86]. Each dataset obtained by scRNA-seq (GSE97519) was first trimmed in TrimGalore v0.6.7. Next, using whole mouse genome data (GCF_000001635) in the database as a reference, trimmed reads were mapped using STAR v2.7.9a with the GTF file. The TPM values were calculated by RSEM v1.3.3. Data for *Sox9*-expressing cells was extracted from XY. Likewise, GSE119766 was used for the analysis in XX[87]. Data for *Foxl2*-expressing cells were extracted from XX.

RNA probe design and synthesis

Total RNA was isolated from a gonad of a E13.5 mouse using TRIzol (Thermo Fisher Scientific) according to the manufacturer's instructions. cDNA was synthesized using

a PrimeScript™ RT Reagent Kit with gDNA Eraser (Perfect Real Time) (Takara Bio) following the manufacturer's instructions. Partial sequences of candidate genes were amplified and cloned into the pGEM-T Easy Vector. Primer sequences are shown in Table 4. The nucleotide sequence was confirmed by Sanger sequence method. After linearization of the plasmids by digestion with *Apa*I, *Spe*I, or *Sph*I, RNA probes were synthesized using a MaxiScript SP6/T7 Transcription Kit (Thermo Fisher Scientific) and labeled with digoxigenin-11-deoxyuridine 5-triphosphate using DIG RNA Labeling Mix (Roche).

WISH

Standard protocols for WISH were used with minor modifications[88]. Briefly, gonad/mesonephros complexes were dissected and fixed in 4% (w/v) PFA prepared in PBS-T (DEPC treated PBS containing 0.1% (v/v) Triton X-100) overnight at 4°C. Samples were gradually dehydrated using ethanol (25%, 50%, and 75% (v/v)) diluted in PBS-T and stored in 100% ethanol at -30°C. After rehydration, samples were treated with 5 μg/mL proteinase K in PBS-T at 37°C for 10 min and then fixed in 4% PFA prepared in PBS-T at room temperature for 15 min. Samples were pre-hybridized in hybridization buffer (750 mmol/L NaCl and 75 mmol/L trisodium citrate pH 7.0, 50% (v/v) formamide, 2% (w/v) Blocking Reagent (Roche), 0.1% (w/v) CHAPS, 0.1% (v/v) Triton X-100, 1 μg/mL yeast tRNA (Merck KGaA), 50 μg/mL heparin sodium, and 5 mmol/L EDTA) at 65°C for 1 hr. Subsequently, samples were

hybridized with DIG-labeled RNA probes overnight at 65°C. The next day, samples were incubated in blocking solution (1.5% (w/v) Blocking Reagent, 187.5 mmol/L NaCl, 50 mmol/L Tris, 10 mmol/L KCl, 10 mmol/L maleic acid, and 0.1% (v/v) Triton X-100) and then with Anti-Digoxigenin-AP, Fab fragments (Roche) diluted 1:2000 in the solution at 37°C for 1 hr each. Alkaline phosphatase was detected using NBT/BCIP Stock Solution (Roche). At least three urogenital complexes were used for expression analysis of each gene and sex. Antisense and sense probes were tested in parallel, and only signals specific to antisense probes were reported as gene-specific expression.

2.4 Results

Identification of 174 genes as initial candidates for factors that regulate *Sox9* through tosEnh14

Gonad extracts of E13.5 mice underwent Southwestern blotting to detect proteins that bind to tosEnh14 (Fig. 6A). An *in vivo* reporter gene assay was previously performed using transgenic mice carrying two fragments of tosEnh14 with the *Sox9* promoter and the reporter gene *lacZ*[35]. XY transgenic embryos showed robust β-Gal activity in the testis compared with the weak activity in the ovary of XX transgenic embryos. According to this result, among the bands detected by Southwestern blotting, bands of a similar molecular weight but prominent in males and faint in females (indicated by arrows in Fig. 6A) were subjected to LC-MS/MS (Fig. 6B). In total, 50, 15, and 47 genes were identified from the higher molecular weight band of males, females, and both sexes, respectively (Fig. 6C left). In total, 63, 10, and 32 genes were identified from the lower molecular weight band of males, females, and both sexes, respectively (Fig. 6C right). Genes counted in both the higher and lower molecular weight bands were counted as one gene, and 174 genes identified in males and both sexes were selected as initial candidates (Table 5). Base on the suggestion that *Sox9* regulators are also expressed in female gonads[35], genes identified in both sexes were also considered as candidates.

Identification of seven candidate genes by GO and RNA-seq analyses

For the next step of screening, I investigated the functional characteristics of the 174 candidate genes by GO analysis. The top ten GO terms in each category, namely, cellular component (CC), molecular function (MF), and biological process (BP), are shown in Fig. 7A, B, and C, respectively. I selected genes associated with the GO terms "nucleus" in CC, "DNA binding" in MF, and "regulation of transcription by RNA polymerase II", "positive regulation of transcription by RNA polymerase II", and "positive regulation of DNA-templated transcription" in BP because these are characteristics of the transcription factors of interest (Fig. 7C). Additionally, genes with GO term "positive regulation of DNA-binding transcription factor activity" which was not ranked 10th, were selected. This left eight genes as candidates: E74-like factor 2 (*Elf2*), ets variant 6 (*Etv6*), Flt3 interacting zinc finger protein 1 (*Fiz1*), *Gtf2f1*, v-rel reticuloendotheliosis viral oncogene homolog A (avian) (*Rela*), tripartite motif-containing 27 (*Trim27*), tripartite motif-containing 28 (*Trim28*), and zinc finger protein 58 (*Zfp58*) (Table 6).

To determine whether these candidate genes are expressed in embryonic gonad cells at the developmental stage during which *Sox9* expression is regulated, the expression levels of the candidate genes were confirmed using scRNA-seq data reported previously[86]. The scRNA-seq data was used from the NCBI bioproject which was obtained from NR5A1-GFP-positive cells in XY gonads of transgenic male mice at various embryonic stages, E10.5, E11.5, E12.5, E13.5, and E16.5. A portion of NR5A1-positive cells in embryonic gonads at sex determination differentiate into

Sertoli cells by transiently expressing *Sry* and upregulating *Sox9* expression via SRY. NR5A1-positive cells lacking *Sry* and *Sox9* expression differentiate into other cell types such as Leydig cells. Therefore, I extracted the scRNA-seq data of cells expressing *Sox9* and calculated TPM value of each candidate gene, as well as *Sry* and *Sox9* as controls (Fig. 8A). In mice, expression of *Sry* starts at E10.5, peaks at E11.5, and disappears after E12.5[89]. I confirmed the same transient pattern in the TPM calculation of *Sry* (Fig. 8A). *Sox9* is expressed in male and female gonads at E10.5. However, it is upregulated only in male gonads at E11.5 because *Sry* regulates its expression[80,90]. A similar pattern i.e., upregulation at E11.5, was confirmed in the TPM calculation of *Sox9* (Fig. 8A). These results indicate that this TPM calculation was reliable. In most mammalian species, *Sry* is essential for initial male sex differentiation. Insufficient *Sry* expression causes *Sox9* reduction and gonadal feminization[91]. However, the peak expression level is extremely low and the timing of expression is limited[92]. Therefore, I thought that the candidate transcription factors would need to be expressed at levels equal to or higher than *Sry* at between E10.5 and E12.5. The expression levels of *Zfp58* were lower than that of *Sry*, therefore, *Zfp58* was excluded from candidates.

To elucidate expression level of the candidate genes in female supporting cells, I also calculated TPM value using scRNA-seq data of XX embryonic gonads[87]. I extracted *Foxl2*-expressing cells data as an antagonist of *Sox9*. Values for TPM calculations in XX were obtained from E12.5, E13.5, and E16.5 because *Foxl2* expression is initiated around E12.5[93]. Expression level of *Elf2*, *Trim27*, and *Trim28*

were higher in XY than in XX in all stages (Fig. 8B). For other genes, the difference in expression levels between XY and XX varied depending on the stage. To further examine the results obtained from these *in silico* analyses, I performed *in vivo* analysis using mouse embryonic gonads.

Five candidate genes are clearly expressed in seminiferous tubules

To examine the expression patterns of the candidate genes in male and female embryonic gonads at E13.5, I performed WISH using urogenital tissues (gonad on the mesonephros). The expression pattern of *Sox9* is shown in Fig. 9 as a positive control. *Sox*9 was highly expressed in seminiferous tubules of male gonads, but was weakly expressed in female gonads. These patterns are identical to those reported previously[94]. *Rela* was not clearly expressed in seminiferous tubules of male gonads. *Trim28* was highly expressed not only in male but in female. Therefore, *Rela* and *Trim28* were excluded as candidates.

All candidate genes except *Rela* were clearly expressed in seminiferous tubules, but their expression levels differed. The male expression levels of *Etv6*, *Fiz1*, *Gtf2f1*, *Trim27*, and *Trim28* were identical to the results of the TPM calculation (Fig. 8A). However, while *Elf2* was relatively highly expressed according to WISH (Fig. 9), it was weakly expressed according to the TPM calculation (Fig. 8A). Therefore, I may have observed expression in cells other than Sertoli cells in WISH. All genes were also expressed in female gonads. In addition, *Fiz1*, *Gtf2f1*, and *Trim27* were expressed

at lower levels in females than in males, but *Elf2*, *Etv6*, and *Trim28* were also strongly expressed in females.

Binding sites for three candidates are present in tosEnh14

After these screening processes, five genes were identified as final candidates: *Elf2*, *Etv6*, *Fiz1*, *Gtf2f1*, and *Trim27*. Of these candidates, the DNA-binding sites for ELF2 and ETV6 have been reported, and their consensus binding motifs are GGAA and ccGGAAgt, respectively[95-98]. According to JASPAR 2024[99], the putative binding sites for human ELF2 and ETV6, which share 100% identity with mouse ELF2 and ETV6 in the DNA-binding domain ETS, are present in tosEnh14 (Fig. 10). These results support the possibility that ELF2 and ETV6 are transcription factors that bind to tosEnh14. The binding site for FIZ1 is unknown, but the consensus binding sequence for Sp1, which contains the same DNA-binding motif as FIZ1, is GGGCGG or GAG repeats[100], indicating that FIZ1 binds to similar sequences. Of note, tosEnh14 contains a GAGGAG sequence (Fig. 10). *Gtf2f1* encodes RAP74 protein, whose binding site is also unknown. RAP74 is a subunit of the general transcription initiation factor TFIIF[101]. TFIIF binds to TATA box[102,103], suggesting that RAP74 binds to the TATA box. Although it has not been reported that RAP74 binds to an enhancer region, promoter-enhancer interaction often observed. Indeed, tosEnh14 contains a TATA-like sequence (Fig. 10). As for the remaining gene, *Trim27*, no binding sites have been reported and therefore I could not confirm them.

2.5 Discussion

I identified five candidate genes as regulators of *Sox9*. Two of them, *Elf2* and *Etv6*, are members of the ELF and TEL subfamilies of the ETS transcription factor family, respectively, which contain a DNA-binding domain termed the ETS domain[104]. Members of the ETS family play important roles in various basic biological phenomena such as development, differentiation, cell proliferation, immune responses, and apoptosis[105]. *Elf2*, also known as *Nerf*, plays an important role in hemopoietic development[106]. ELF2 contains a transactivation domain at the N-terminus[104]. It has two major isoforms, ELF2A and ELF2B, which activate and repress expression of their target genes, respectively[107]. Mouse *Elf2*a is more highly expressed in the testis than *Elf2b*, whereas *Elf2b* is generally expressed at a higher level in the thymus and spleen[108]. ELF2 transcriptionally regulates target genes involved in lymphocyte development and function, such as B and T cell co-receptor proteins and tyrosine kinases, by binding to promoter or enhancer regions[106,109]. It also regulates valosin-containing protein (*VCP*) expression by binding to the 5'-flanking region of *VCP* in MCF-7 cells[110]. However, the function of *Elf2* in gonads or testes has not been reported. These findings support that *Elf2* is involved in embryonic testicular development.

Etv6, also known as *Tel*, was first identified by its rearrangement in human chronic myelomonocytic leukemia associated with a chromosomal translocation[111]. Expression of *Etv6* starts from E7.0 at the latest and spreads widely in embryo and adult mouse tissues, including the testis[112]. *Etv6* homozygous knockout is embryonic

lethal in mice and no embryos survive after E13.5, suggesting that *Etv6* is essential for normal development, particularly hematopoiesis and vascular development[112]. In ETV6-overexpressing cells, *p53* and its downstream target genes are upregulated, inducing apoptosis[113]. Target genes of ETV6, as well as of ELF2, in embryonic gonads are unknown; hence, it is necessary to validate the function of these candidates at the sex determination stage.

One of the remaining three candidates, *Trim27*, also known as *Rfp*, belongs to the tripartite motif (TRIM) family and is classified into the IV subfamily based on the structure of the C-terminal region. TRIM proteins in class IV contain a PRY domain and a SPRY domain in addition to the RING, b-box, and coiled-coil domains, similar to other subfamily proteins[114]. Although RFP is well-known as a DNA-binding protein and a transcriptional repressor[115-117], it also functions as a positive regulator[118,119]. *Rfp* initiates expression at least as early as E8.5, and is expressed in testis and some other tissues in adult mice[120,121]. In testis, RFP is localized in the nuclei of Sertoli cells and in the cytoplasm of round spermatids[122], indicating that *Rfp* is related to embryonic development and spermatogenesis. This report is the first finding that RFP may be associated with *Sox9* in fetal Sertoli cells. FIZ1 was originally isolated as an *Flt3* (receptor tyrosine kinase)-interacting zinc finger protein[123]. *Fiz1* is a widely expressed gene of unknown function. In the retina, FIZ1 acts as a transcriptional regulatory protein complex at the active rod- and cone-specific gene promoter region in mice[124]. FIZ1 contains 11 C(2)H(2)-type zinc fingers related to DNA binding. The C(2)H(2) zinc finger domain is also found in Sp1, which

binds to the promoter region of Wilms' tumor 1 (*Wt1*) and *Sry* and regulates their expression[125,126]. Interestingly, tosEnh14 contains an Sp1 binding site, suggesting that FIZ1 has the potential to regulate *Sox9* in embryonic gonads.

The final candidate gene, *Gtf2f1*, encodes RAP74, which heterodimerizes with RAP30 to form the general transcription initiation factor TFIIF[101]. TFIIF binds to RNA polymerase II and helps to recruit it to the initiation complex in collaboration with TFIIB[127]. Subunits of both TFIIF and Pol II crosslink to nucleotides located upstream and downstream of the TATA box[102,103]. The *Sox9* promoter contains a TATA-like sequence that is conserved in mouse and human[128]. It is possible that RAP74 regulates *Sox9* by associating with the promoter and tosEnh14.

In testicular development, ets variant 2 (*Etv2*) is upregulated by SRY and then ETV2 regulates *Sox9* by binding to its promoter region in mouse[129]. This regulation is conserved in *T. osimensis*[34], but its involvement with tosEnh14 has not been demonstrated. In this study, *Etv2* was undetected by mass spectrometry analysis, suggesting that it is unlikely to function through tosEnh14. In contrast, ELF2 and ETV6, which belong to the same ETS transcription factor family as ETV2, were identified by mass spectrometry and have predicted binding sites in tosEnh14, suggesting that they are especially strong candidates for regulating *Sox9* through tosEnh14. To verify their function, it is necessary to conduct reporter gene assays to assess their transcriptional activity and confirm their roles as tosEnh14 regulators.

2.6 Tables

Table 4 Primers used in the chapter.

Target Sequence/gene	Experiment	Sequence 5' to 3' (Sense/ Antisense)
tosEnh14	Southwestern blotting	GGATAGAGAAGCTTGAAATGC
		TTGCATGCTTTCAGGTGCTT
Elf2	WISH	AGCCAGAGAAGGAAAAGGAAACAC
		AGGAACCGTCGTGAGTAGGG
Etv6	WISH	CATTGGGAGAATAGCAGACTGTAG
		TGCTCGGTGGCTTGTCTAAG
Fiz1	WISH	AACGACCCTACCGATGCTCTG
		CATGAGTGAGCTTGTGGCGTTC
Gtf2f1	WISH	AGCAAGCTGGAGCAGGGGAAG
		GGCATCACTCTTTGAGGGAGAAGTG
Rela	WISH	TACTTGCCAGACACAGATGATCGC
		CTGAGAGACCATTGGGAAGCC
Trim28	WISH	GTGAAATACACCAAGGACCACAC
		CTCAGCCACAATCTTGCCAAAGG

Table 5-1 Genes identified from embryonic gonads by LC-MS/MS.

Upper band

Male	
Accession	Gene symbol
Q80US4	*Actr5*
Q91Z58	*AI661453*
O88845	*Akap10*
P50516	*Atp6v1a*
Q0VF22	*Ccdc138*
Q9DC48	*Cdc40*
A3KGW5	*Cercam*
Q9CXV9	*Dcun1d5*
Q6P9R1	*Ddx51*
Q9Z247	*Fkbp9*
P97379	*G3bp2*
P26049	*Gabra3*
Q8CHY6	*Gatad2a*
Q3THK3	*Gtf2f1*
Q61696	*Hspa1a*
P01869	*Ighg1*
P01837	*Igkc*
E9Q1P8	*Irf2bp2*
Q3TCX3	*Khdc4*
Q3U0V1	*Khsrp*
Q9QWT9	*Kifc1*
Q9DBS5	*Klc4*
Q61084	*Map3k3*
Q8R3C0	*Mcmbp*
Q9CQG2	*Mettl16*
P54729	*Nub1*
Q99JX7	*Nxf1*
Q7TSV4	*Pgm2*
P33611	*Pola2*
P49443	*Ppm1a*
Q8CEC6	*Ppwd1*
Q8CIG8	*Prmt5*
Q8VCT3	*Rnpep*
D3YXK1	*Samd1*
Q60710	*Samhd1*
Q9DBC0	*Selenoo*
Q9EP97	*Senp3*
Q64213	*Sf1*
P10852	*Slc3a2*
Q9CWK8	*Snx2*
Q8VE97	*Srsf4*
Q9CQN1	*Trap1*
Q61510	*Trim25*
Q62158	*Trim27*
Q62318	*Trim28*

Female	
Accession	Gene symbol
P30658	*Cbx2*
Q8CJ40	*Crocc*
P70303	*Ctps2*
Q8CIB9	*Esco2*
Q9CPN8	*Igf2bp3*
A2A9R3	*Mageb4*
P09405	*Ncl*
Q04207	*Rela*
Q9DBU3	*Riok3*
P14576	*Srp54*
Q8R317	*Ubqln1*
Q8VHI6	*Wasf3*
P59326	*Ythdf1*
Q91YT7	*Ythdf2*
Q8BYK6	*Ythdf3*

Both	
Accession	Gene symbol
P02772	*Afp*
P07724	*Alb*
Q3TMW1	*Ccdc102a*
Q6NVF9	*Cpsf6*
Q8BIQ5	*Cstf2*
Q501J6	*Ddx17*
P58252	*Eef2*
Q8BGD9	*Eif4b*
Q9WUA2	*Farsb*
Q61576	*Fkbp10*
P35922	*Fmr1*
P56959	*Fus*
Q61584	*Fxr1*
Q60760	*Grb10*
Q9D0E1	*Hnrnpm*
P07901	*Hsp90aa1*
P20029	*Hspa5*
Q8R3Y8	*Irf2bp1*
Q99MN1	*Kars1*
Q99JN2	*Klhl22*
Q05CL8	*Larp7*
P14733	*Lmnb1*
P21619	*Lmnb2*
Q505F5	*Lrrc47*
Q6NSQ7	*Ltv1*
Q8VCF0	*Mavs*
Q5NCR9	*Nsrp1*
P29341	*Pabpc1*
Q9EQ61	*Pes1*
Q0VGB7	*Ppp4r2*
O08795	*Prkcsh*
Q922H1	*Prmt3*
Q9CT10	*Ranbp3*
Q8VH51	*Rbm39*
Q8C796	*Rcor2*
Q8VEE4	*Rpa1*
Q80UG5	*Septin9*
Q8BKJ9	*Sirt7*
Q91ZR2	*Snx18*
Q9DBG7	*Srpra*
Q6A028	*Swap70*
P40142	*Tkt*
Q9CZW5	*Tomm70*
Q02053	*Uba1*
Q6P5G6	*Ubxn7*

Table 5-2 Genes identified from embryonic gonads by LC-MS/MS.

Upper band					
Male			**Both**		
Accession	Gene symbol		Accession	Gene symbol	
Q99KW3	*Triobp*		P23475	*Xrcc6*	
Q3UDE2	*Ttll12*		Q8R060	*Zwilch*	
Q8BGG7	*Ubash3b*		Q3TMW1	*Ccdc102a*	
Q99NB8	*Ubqln4*		Q6NVF9	*Cpsf6*	
Q62523	*Zyx*		Q8BIQ5	*Cstf2*	

Table 5-3 Genes identified from embryonic gonads by LC-MS/MS.

Lower band

Male	
Accession	Gene symbol
P50544	*Acadvl*
Q5XG73	*Acbd5*
P02772	*Afp*
Q8C0I1	*Agps*
P07724	*Alb*
P27106	*Amh*
Q9DBR3	*Armc8*
Q80X53	*Ccdc116*
Q99LM2	*Cdk5rap3*
Q9WUM3	*Coro1b*
P52825	*Cpt2*
Q9R194	*Cry2*
Q8BIQ5	*Cstf2*
A2ATU0	*Dhtkd1*
Q8BL99	*Dop1a*
P58252	*Eef2*
Q9JHW4	*Eefsec*
Q9JHC9	*Elf2*
Q8K2I9	*Fbh1*
Q9WTJ4	*Fiz1*
P56959	*Fus*
P97855	*G3bp1*
Q9EST1	*Gsdma*
P20029	*Hspa5*
Q8BK58	*Hspbap1*
Q9CPN8	*Igf2bp3*
Q02257	*Jup*
Q3U0V1	*Khsrp*
O08677	*Kng1*
O54785	*Limk2*
P14733	*Lmnb1*
P24527	*Lta4h*
P17897	*Lyz1*
Q91YJ5	*Mtif2*
Q99MD9	*Nasp*
Q8R480	*Nup85*
Q8BH04	*Pck2*
Q8BSN5	*Phf23*
Q80UG2	*Plxna4*
P36993	*Ppm1b*
Q8VI36	*Pxn*
Q9D0I9	*Rars1*
Q8VH51	*Rbm39*
Q9JJF3	*Riox1*
Q8BLK9	*Rps6kc1*

Female	
Accession	Gene symbol
P61222	*Abce1*
P80315	*Cct4*
Q9D1L0	*Chchd2*
Q8K1M6	*Dnm1l*
Q9ERG0	*Lima1*
Q8BTW9	*Pak4*
Q7M6Y3	*Picalm*
Q8BP71	*Rbfox2*
P14576	*Srp54*
Q60770	*Stxbp3*

Both	
Accession	Gene symbol
Q9Z0X1	*Aifm1*
P14824	*Anxa6*
P49919	*Cdkn1c*
Q8BI72	*Cdkn2aip*
Q501J6	*Ddx17*
Q61656	*Ddx5*
Q8BMF4	*Dlat*
Q8BJW6	*Eif2a*
Q8JZQ9	*Eif3b*
Q8BGD9	*Eif4b*
P97360	*Etv6*
P97379	*G3bp2*
Q9D0E1	*Hnrnpm*
O88477	*Igf2bp1*
Q5SF07	*Igf2bp2*
Q60787	*Lcp2*
Q9CRC8	*Lrrc40*
P09405	*Ncl*
Q9D6Z1	*Nop56*
Q02819	*Nucb1*
Q63850	*Nup62*
Q9EQ28	*Pold3*
Q9D787	*Ppil2*
Q04207	*Rela*
Q91YQ5	*Rpn1*
Q8BKJ9	*Sirt7*
Q9CSN1	*Snw1*
Q60864	*Stip1*
Q9CZB3	*Thumpd2*
P40142	*Tkt*
Q62318	*Trim28*
Q02053	*Uba1*

Table 5-4 Genes identified from embryonic gonads by LC-MS/MS.

Lower band	
Male	
Q5I2A0	*Serpina3g*
Q61247	*Serpinf2*
Q91WC0	*Setd3*
Q64213	*Sf1*
Q9DAG4	*Spmip9*
Q80ZI1	*Trnp1*
Q6P7W5	*Tsen2*
Q8R317	*Ubqln1*
Q3TIX9	*Usp39*
Q9Z0K8	*Vnn1*
Q91YD9	*Wasl*
O88342	*Wdr1*
Q9D994	*Wdr38*
Q3TUF7	*Yeats2*
P59326	*Ythdf1*
Q91YT7	*Ythdf2*
Q8BYK6	*Ythdf3*
P16372	*Zfp58*

Table 6 Candidate genes and their GO terms.

			GO term in BP			
Gene symbol	Gene name	Gene ID	regulation of transcription by RNA polymerase II	positive regulation of transcription by RNA polymerase II	positive regulation of DNA-templated transcription	positive regulation of DNA-binding transcription factor activity
Elf2	E74-like factor 2	69257	✓		✓	
Etv6	ets variant 6	14011	✓	✓		
Fiz1	Flt3 interacting zinc finger protein 1	23877	✓	✓		
Gtf2f1	general transcription factor IIF, polypeptide 1	98053		✓		
Rela	v-rel reticuloendotheliosis viral oncogene homolog A (avian)	19697	✓	✓	✓	
Trim27	tripartite motif-containing 27	19720				✓
Trim28	tripartite motif-containing 28	21849			✓	
Zfp58	zinc finger protein 58	238693	✓			

2.7 Figures

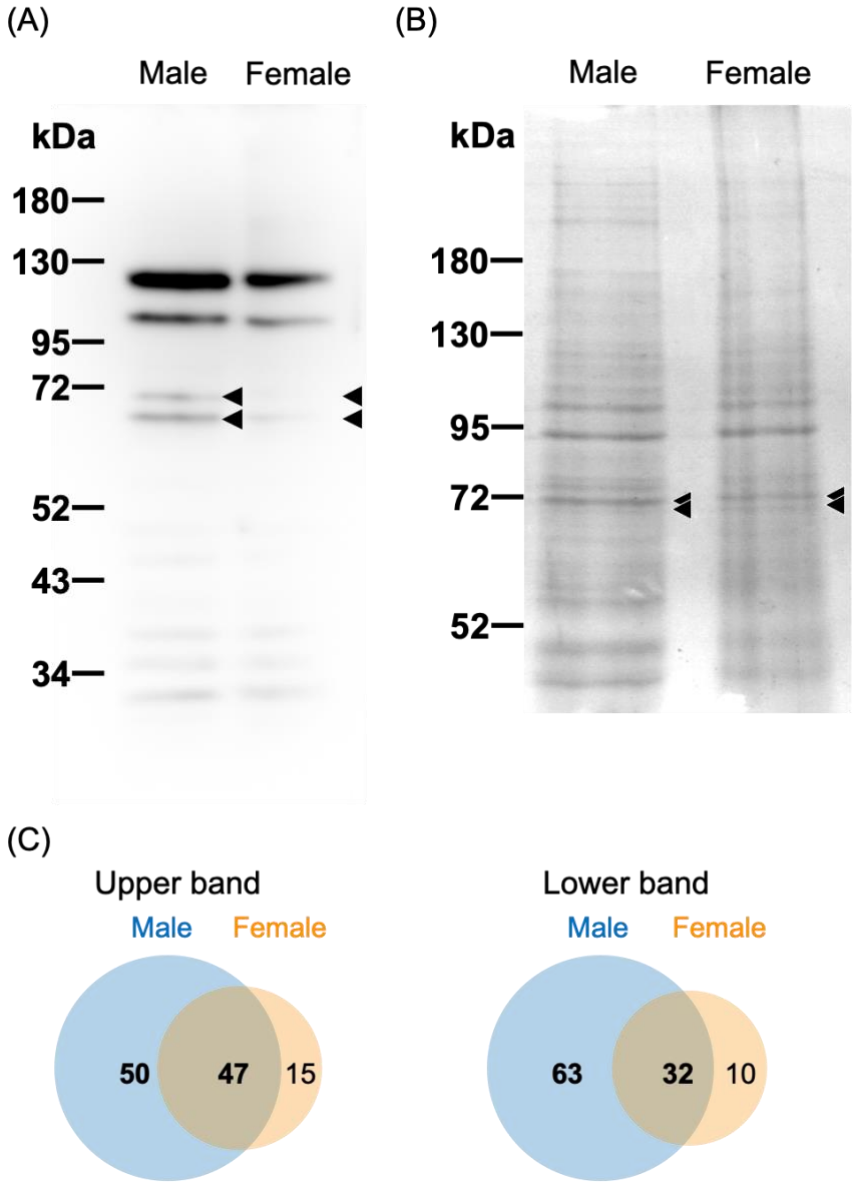

Fig. 6 Screening of proteins that bind to tosEnh14 DNA.

(A) Southwestern blotting analysis to detect proteins that bind to tosEnh14. Arrowheads indicate bands of a similar molecular weight that are prominent in males but faint in females. (B) CBB staining for LC-MS/MS. Each band marked by an arrowhead was excised and subjected to LC-MS/MS. (C) Venn diagram of genes identified in males and/or females.

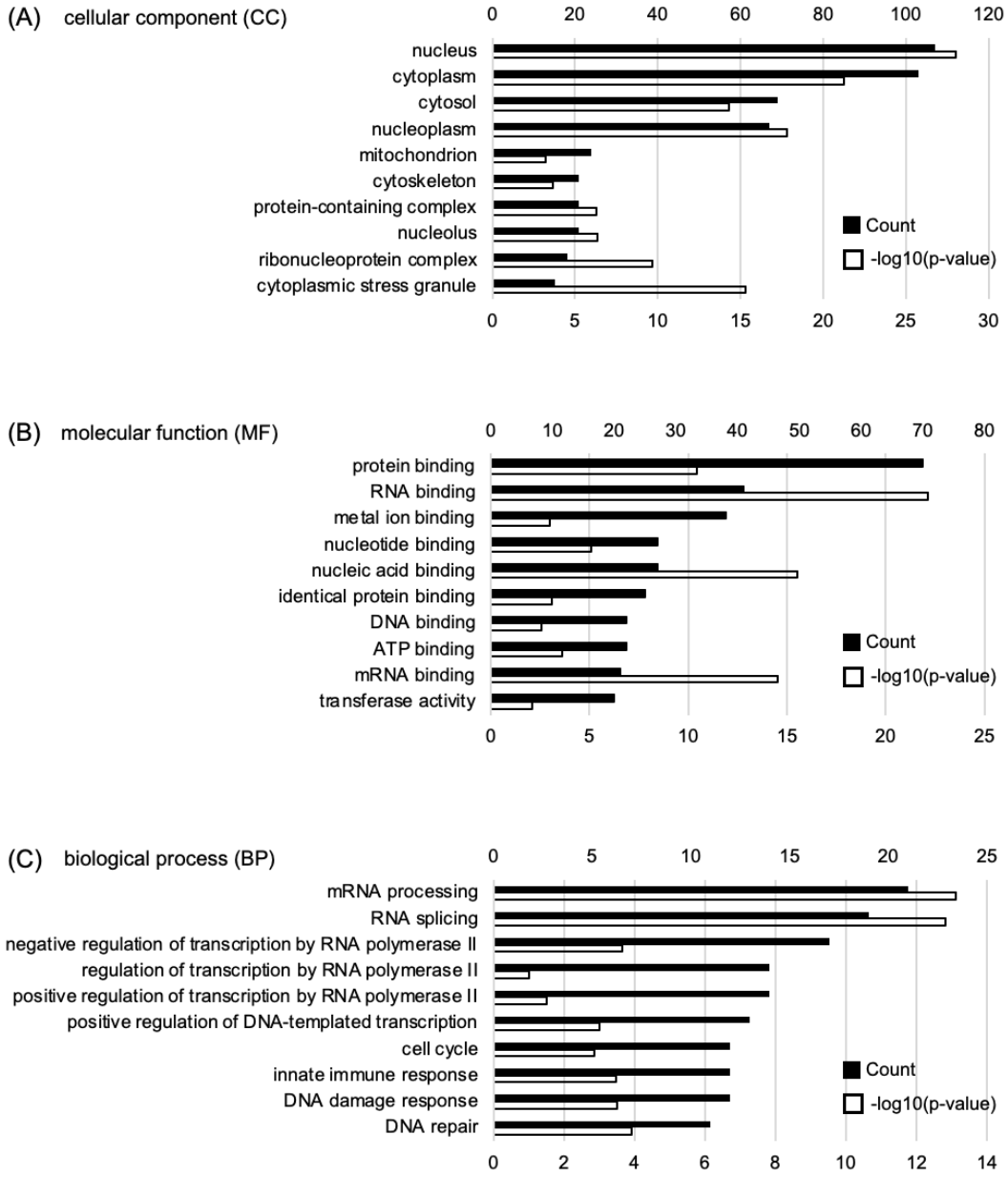

Fig. 7 GO analysis of genes identified by mass spectrometry.

(A) BP, (B) CC, and (C) MF. The top ten GO terms are shown.

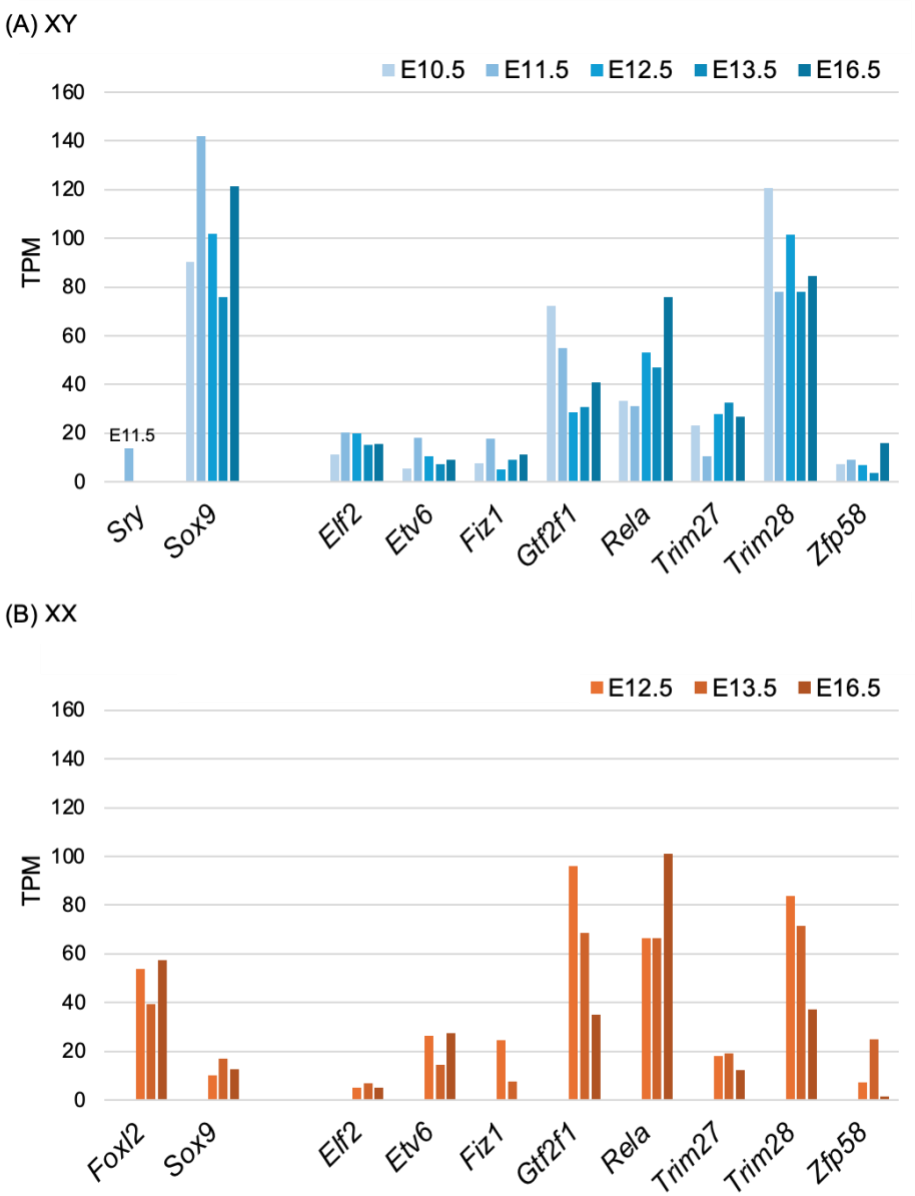

Fig. 8 Expression levels of *Sry* or *Foxl2*, *Sox9*, and eight candidate genes.

The TPM value of each gene was calculated using scRNA-seq data of progenitor cells, which differentiate into (pre-)Sertoli cells (A) or (pre-)granulosa cells (B) in mouse embryonic gonads.

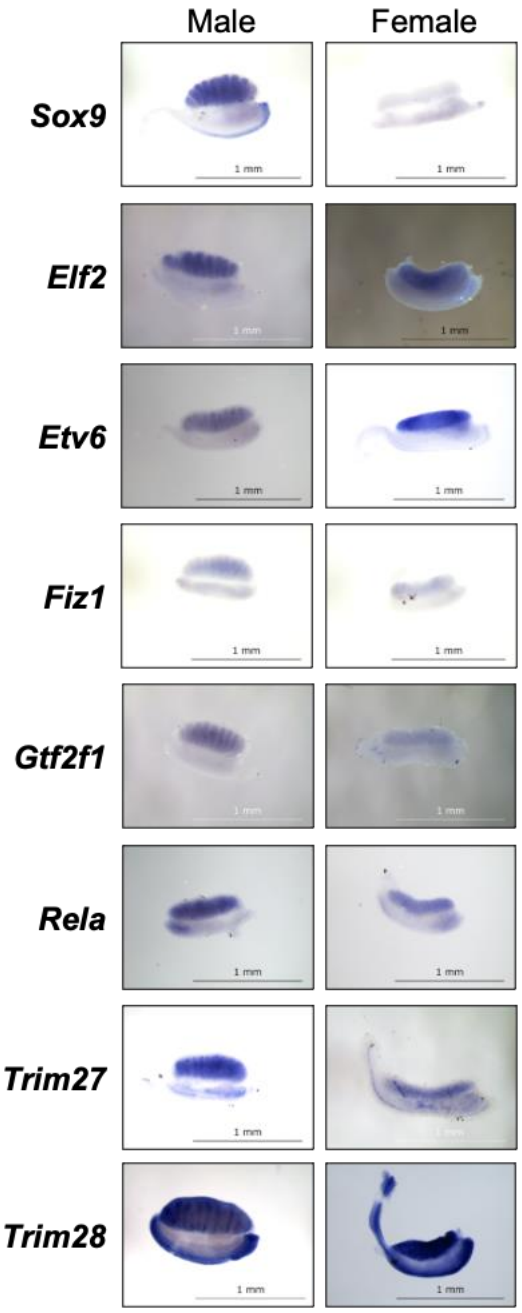

Fig. 9 Expression of *Sox9* and candidate genes in E13.5 mouse gonads examined by WISH.

WISH was performed using partial sequences of each target gene. Scale bars indicate 1 mm.

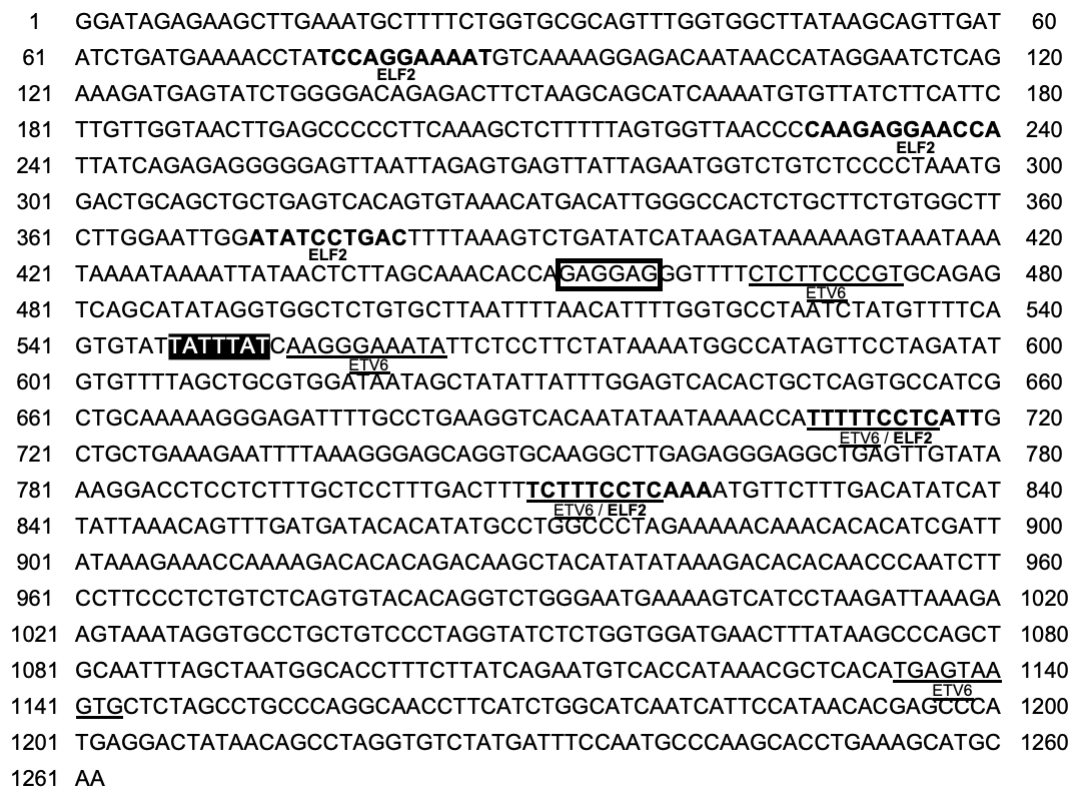

Fig. 10 Putative binding sites for ELF2 and ETV6 in tosEnh14.

Binding sites for ELF2 and ETV6 are bold and underlined, respectively. GAG repeat is surrounded by a black frame and TATA-like sequence is painted by black.

General discussion

The aim of this study was to investigate the regulatory mechanisms of *Sox9* in *T. osimensis*, a rodent species lacking the Y chromosome and *Sry*, with a particular focus on the testis-specific enhancer tosEnh14. Chapter I demonstrated that biochemical screening using Southwestern blotting and mass spectrometry identified RCOR3 as a promising transcription factor candidate that binds specifically to tosEnh14. On the contrary, a similar approach using embryonic mouse gonads revealed five different candidate transcription factors and RCOR3 did not detect in Chapter II. The identification of distinct factors in each chapter suggests that tosEnh14 may interact with different sets of transcription factors depending on developmental stage and be able to serve multiple roles in testis. In hematopoietic differentiation, GATA-binding enhancers where GATA-binding protein 2 (GATA2) is replaced by GATA-binding protein 1 (GATA1) as cells transition from progenitor to more differentiated states, thereby altering gene regulatory outcomes without changing enhancer[130]. Similarly, the distinct transcription factors that engage with the tosEnh14 enhancer in the fetal and adult testes may reflect a mechanism by which the enhancer contributes to different regulatory programs during testis development. During the window of sex determination, ELF2, ETV6 and/or potentially other candidate factors may trigger *Sox9* transcription through tosEnh14, thereby initiating testis differentiation. In testis development, RCOR3 may gradually replace initiation factors and sustain *Sox9* expression by maintaining an active chromatin state, thereby supporting Sertoli cell identity and long-term testicular function.

tosEnh14 with RCOR3 exhibited enhancer activity from the modest increase in transcriptional activity observed in reporter gene assays. Furthermore, the copy number of tosEnh14 causes differences in enhancer activity. These results not only advance understanding of *Sox9* regulation in *T. osimensis* but also provide valuable insights into the function of the homologous mEnh14, which indicates compensation of testis-specific enhancers for *Sox9* such as Enh13[36]. In fact, the sequence identity between mEnh14 and tosEnh14 is 87.7%[35], and almost putative binding sites are also conserved in mEnh14. By identifying transcription factors that interact with tosEnh14 and demonstrating its transcriptional activity *in vitro*, this study offers candidate mechanisms and regulatory contexts that may be conserved in mouse. These findings suggest that multiple conserved transcriptional regulators may act on tosEnh14, even across species.

On the other hand, the difference is not so much and may not generate the *Sox9* expression level required for testis development[35,40]. As mentioned in Chapter I, one possibility is the presence of other factors to intensify the activity of tosEnh14. The regulation of *Sox9* by tosEnh14 is complex and possibly combinatorial, suggesting that RCOR3 may contribute to its regulation but is unlikely to act alone. Another reason is that they imply a possibility that other *cis*-elements besides tosEnh14 in male-specific duplication also contribute to *Sox9* regulation. According to high-quality long-read-based chromosome-level genome assemblies, male-specific sequences that repressed recombination are contained in the 17-kb sequence[131].

Therefore, the identification of other *cis*-elements is required to understand *Sox9* regulatory mechanism in *T. osimensis* further.

Mammals comprise over 6,000 extant species and are classified into three major groups: monotremes, marsupials, and eutherians[132]. The XY sex chromosome system is thought to have originated in the common ancestor of marsupials and eutherians approximately 160 to 180 million years ago[133,134]. Although both marsupials and eutherians possess *SRY*/*Sry*, monotremes such as platypuses and echidnas lack the *SRY* altogether and instead retain *SOX3*, the presumptive evolutionary precursor of *SRY*, on their autosomes[135,136]. Monotremes possess a complex sex chromosome system consisting of five pairs of X and Y chromosomes, which are more similar to the avian ZW system than to the XY chromosomes found in therian mammals[135,137,138]. These observations suggest that *SRY* emerged in the common ancestor of marsupials and eutherians, probably at around the same time as the Y chromosome originated. Despite the evolutionary conservation of *SRY*/*Sry* among most eutherians, the genus *Ellobius* and *Tokudaia* have been discovered in which *Sry* is completely absent. *E. talpinus* and *E. tancrei* lost the Y chromosome, and males also possess two X chromosomes[139-142]. *E. lutescens* has 2*n*=17 and is XO/XO, lacking *Sry*[143]. *Sox9* remains a pivotal gene in testis differentiation and is likely regulated by conserved evolutionarily conserved region (ECR) in TESCO, which may function as an enhancer in *E. lutescens*[142]. However, the specific genomic sequences or regulatory elements in males that activate *Sox9* in this species remain unknown. By contrast, a male-specific duplication including tosEnh14 suggests regulation of *Sox9*

in *T. osimensis*[35]. Despite their geographic and genetic isolation, *E. lutescens* and *T. osimensis* appear to have converged on a common solution, that is enhancement of *Sox9* expression by *cis*-elements as an alternative to *Sry* function. This suggests that enhancer-mediated regulation of *Sox9* could be a flexible and conserved mechanism for sex determination in placental mammals. The emergence of *Sry*-independent regulatory strategies centered on *Sox9* could reflect an underlying robustness in the mammalian sex-determining network, whereby enhancer elements can compensate for the loss of the ancestral regulation by *Sry*. These findings suggest that such enhancer-based mechanisms may constitute a universal backup system intrinsic to mammalian genomes.

Y chromosome carries the sex-determining gene *SRY*/*Sry* in mammals. However, the Y chromosome has progressively degenerated throughout evolution, and its potential disappearance in the future has been suggested[144]. On the other hand, the Y chromosome has lost only a single gene since the divergence of humans and rhesus monkeys (*Macaca mulatta*) 25 million years ago, suggesting that it is unlikely to disappear[145,146]. This study aims to elucidate the sex determination mechanism in *T. osimensis*, a model animal that has already lost its Y chromosome and produces male individuals. In this study, I demonstrated that a regulatory factor for *Sox9* is not a unique gene in *T. osimensis*, but rather a widely conserved gene that has acquired a new role as a sex-determining gene. This implies that the loss of the Y chromosome does not necessarily lead to the extinction of males. Instead, mechanisms of alternative gene activation for male development may already exist or have been

evolutionarily acquired to maintain sex determination stability in mammals that have lost their Y chromosome. Thus, the significance of this study extends beyond the specific context of *T. osimensis*. This study provides new insights into the evolutionary plasticity of sex determination mechanisms.

Acknowledgements

First of all, I would like to express my deepest gratitude to my major advisor, Professor Asato Kuroiwa, Biosystems Science Course, Graduate School of Life Science, Hokkaido University, for her unwavering support, insightful guidance, and invaluable expertise throughout my Ph.D. journey. Her commitment to scientific rigor and passion for research have been a constant source of inspiration and have significantly shaped my growth as a researcher.

I am also extremely grateful to the members of my sub-supervisors, Professor Atsushi Kimura, Biosystems Science Course, Graduate School of Life Science, Hokkaido University, and Associate Professor Shusei Mizushima, for their thoughtful feedback, constructive criticism, and encouraging words. Their diverse perspectives and expertise have greatly enriched my work and challenged me to think critically about my research.

I would like to thank Assistant Professor Ikuya Yoshida, Biosystems Science Course, Graduate School of Life Science, Hokkaido University, for his comments and warm encouragement. He also provided me with helpful comments on the development of my research.

Special thanks to my all-laboratory members who have not only provided insightful discussions but have also been a source of friendship and camaraderie. The countless hours spent together in the lab, troubleshooting experiments, and celebrating small victories have made this journey both rewarding and memorable.

I am deeply indebted to the funding agencies that supported this research, including the Hokkaido University EXEX Doctoral Fellowship [JST SPRING (Grant Number JPMJSP2119)]. Their financial support has been instrumental in enabling me to pursue my studies and complete this dissertation.

A heartfelt thank you to my family for their unconditional love, patience, and unwavering belief in me. To my parents, who have always been my greatest supporters and have taught me the value of perseverance and hard work. To my friends and mates, whose encouragement and understanding have given me with refreshment besides my studies.

Finally, I would like to express my appreciation to all those who have contributed in one way or another way to the completion of this dissertation. Your support, encouragement, and belief in me have made this achievement possible.

Thank you all.

References

1. Pokorná, M., and Kratochvíl, L. (2016). What was the ancestral sex-determining mechanism in amniote vertebrates? Biological Reviews *91*, 1-12.

2. Ferguson, M., and Joanen, T. (1982). Temperature of egg incubation determines sex in *Alligator Mississippiensis*. Nature *296*, 850-853.

3. Yatsu, R., Miyagawa, S., Kohno, S., Saito, S., Lowers, R., Ogino, Y., Fukuta, N., Katsu, Y., Ohta, Y., Tominaga, M., et al. (2015). TRPV4 associates environmental temperature and sex determination in the American alligator. Scientific Reports *5*, 18581.

4. Yoshimoto, S., Ikeda, N., Izutsu, Y., Shiba, T., Takamatsu, N., and Ito, M. (2010). Opposite roles of *DMRT1* and its W-linked paralogue, *DM-W*, in sexual dimorphism of *Xenopus laevis*: implications of a ZZ/ZW-type sex-determining system. Development *137*, 2519-2526.

5. Raymond, C., Kettlewell, J., Hirsch, B., Bardwell, V., and Zarkower, D. (1999). Expression of *Dmrt1* in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. Developmental Biology *215*, 208-220.

6. Raymond, C., Murphy, M., O'Sullivan, M., Bardwell, V., and Zarkower, D. (2000). *Dmrt1*, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. Genes and Development *14*, 2587-2595.

7. Nanda, I., Shan, Z., Schartl, M., Burt, D., Koehler, M., Nothwang, H., Grützner, F., Paton, I., Windsor, D., Dunn, I., et al. (1999). 300 million years of conserved

synteny between chicken Z and human chromosome 9. Nature Genetics *21*, 258-259.

8. Nanda, I., Zend-Ajusch, E., Shan, Z., Grützner, F., Schartl, M., Burt, D., Koehler, M., Fowler, V., Goodwin, G., Schneider, W., et al. (2000). Conserved synteny between the chicken Z sex chromosome and human chromosome 9 includes the male regulatory gene *DMRT1*: a comparative (re)view on avian sex determination. Cytogenetics and Cell Genetics *89*, 67-78.

9. Moore, E., and Roberts, R. (2013). Polygenic sex determination. Current Biology *23*, R510-R512.

10. Matsuda, M., Nagahama, Y., Shinomiya, A., Sato, T., Matsuda, C., Kobayashi, T., Morrey, C., Shibata, N., Asakawa, S., Shimizu, N., et al. (2002). *DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. Nature *417*, 559-563.

11. Matsuda, M., Shinomiya, A., Kinoshita, M., Suzuki, A., Kobayashi, T., Paul-Prasanth, B., Lau, E., Hamaguchi, S., Sakaizumi, M., and Nagahama, Y. (2007). *DMY* gene induces male development in genetically female (XX) medaka fish. Proceedings of the National Academy of Sciences of the United States of America *104*, 3865-3870.

12. Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P., and Lovellbadge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature *346*, 245-250.

13. Sinclair, A., Berta, P., Palmer, M., Hawkins, J., Griffiths, B., Smith, M., Foster, J., Frischauf, A., Lovell-Badge, R., and Goodfellow, P. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature *346*, 240-244.

14. Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., and Lovell-badge, R. (1991). Male development of chromosomally female mice transgenic for *Sry*. Nature *351*, 117-121.

15. Stevanovic, M., Lovellbadge, R., Collignon, J., and Goodfellow, P. (1993). *SOX3* is an X-linked gene related to *SRY*. Human Molecular Genetics *2*, 2013-2018.

16. Sutton, E., Hughes, J., White, S., Sekido, R., Tan, J., Arboleda, V., Rogers, N., Knower, K., Rowley, L., Eyre, H., et al. (2011). Identification of *SOX3* as an XX male sex reversal gene in mice and humans. Journal of Clinical Investigation *121*, 328-341.

17. Kashimada, K., and Koopman, P. (2010). *Sry*: the master switch in mammalian sex determination. Development *137*, 3921-3930.

18. Foster, J., Dominguezsteglich, M., Guioli, S., Kwok, C., Weller, P., Stevanovic, M., Weissenbach, J., Mansour, S., Young, I., Goodfellow, P., et al. (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in an *SRY*-related gene. Nature *372*, 525-530.

19. Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasantes, J., Bricarelli, F., Keutel, J., Hustert, E., et al. (1994). Autosomal sex reversal and

campomelic dysplasia are caused by mutations in and around the *SRY*-related gene *SOX9*. Cell *79*, 1111-1120.

20. Ming, Z., Vining, B., Bagheri-Fam, S., and Harley, V. (2022). SOX9 in organogenesis: shared and unique transcriptional functions. Cellular and Molecular Life Sciences *79*, 522.

21. Lefebvre, V., Huang, W., Harley, V., Goodfellow, P., and deCrombrugghe, B. (1997). SOX9 is a potent activator of the chondrocyte-specific enhancer of the proα1(II) collagen gene. Molecular and Cellular Biology *17*, 2336-2346.

22. Arango, N., Lovell-Badge, R., and Behringer, R. (1999). Targeted mutagenesis of the endogenous mouse *Mis* gene promoter: In vivo definition of genetic pathways of vertebrate sexual development. Cell *99*, 409-419.

23. Pask, A., Harry, J., Graves, J., O'Neill, R., Layfield, S., Shaw, G., and Renfree, M. (2002). *SOX9* has both conserved and novel roles in marsupial sexual differentiation. Genesis *33*, 131-139.

24. Jakob, S., and Lovell-Badge, R. (2011). Sex determination and the control of *Sox9* expression in mammals. Febs Journal *278*, 1002-1009.

25. Endo, H., and Tsuchiya, K. (2006). A new species of Ryukyu spiny rat, *Tokudaia* (Muridae: Rodentia), from Tokunoshima Island, Kagoshima Prefecture, Japan. Mammal Study *31*, 47-57.

26. Murata, C., Yamada, F., Kawauchi, N., Matsuda, Y., and Kuroiwa, A. (2010). Multiple copies of *SRY* on the large Y chromosome of the Okinawa spiny rat, *Tokudaia muenninki*. Chromosome Research *18*, 623-634.

27. Murata, C., Yamada, F., Kawauchi, N., Matsuda, Y., and Kuroiwa, A. (2012). The Y chromosome of the Okinawa spiny rat, *Tokudaia muenninki*, was rescued through fusion with an autosome. Chromosome Research *20*, 111-125.

28. Urunanont, P., Mizushima, S., Uchida, T., Ishimori, K., Matiz-Ceron, L., and Kuroiwa, A. (2025). Multiple *Sry* genes with an A-to-S substitution in the HMG domain retain DNA-binding ability in the Okinawa spiny rat. Genes & Genetic Systems *100*, 25-00021.

29. Ogata, Y., Nishikata, M., Kitada, K., Mizushima, S., Jogahara, T., and Kuroiwa, A. (2019). Spiny rat SRY lacks a long Q-rich domain and is not stable in transgenic mice. Developmental Dynamics *248*, 784-794.

30. Honda, T., Suzuki, H., and Itoh, M. (1977). Unusual sex chromosome constitution found in Amami spinous country-rat, *Tokudaia osimensis osimensis*. Japanese Journal of Genetics.

31. Honda, T., Suzuki, H., Itoh, M., and Hayashi, K. (1978). Karyotypical differences of Amami spinous country-rats, *Tokudaia osimensis osimensis* obtained from 2 neighboring islands. Japanese Journal of Genetics *53*, 297-299.

32. Soullier, S., Hanni, C., Catzeflis, F., Berta, P., and Laudet, V. (1998). Male sex determination in the spiny rat *Tokudaia osimensis* (Rodentia: Muridae) is not *Sry* dependent. Mammalian Genome.

33. Kimura, R., Murata, C., Kuroki, Y., and Kuroiwa, A. (2014). Mutations in the testis-specific enhancer of *SOX9* in the *SRY* independent sex-determining mechanism in the genus *Tokudaia*. Plos One, e108779.

34. Otake, T., and Kuroiwa, A. (2016). Molecular mechanism of male differentiation is conserved in the *SRY*-absent mammal, *Tokudaia osimensis*. Scientific Reports *6*, 32874.

35. Terao, M., Ogawa, Y., Takada, S., Kajitani, R., Okuno, M., Mochimaru, Y., Matsuoka, K., Itoh, T., Toyoda, A., Kono, T., et al. (2022). Turnover of mammal sex chromosomes in the *Sry*-deficient Amami spiny rat is due to male-specific upregulation of *Sox9*. Proceedings of the National Academy of Sciences of the United States of America *119*, e2211574119.

36. Gonen, N., Futtner, C., Wood, S., Garcia-Moreno, S., Salamone, I., Samson, S., Sekido, R., Poulat, F., Maatouk, D., and Lovell-Badge, R. (2018). Sex reversal following deletion of a single distal enhancer of *Sox9*. Science *360*, 1469-1473.

37. Symon, A., and Harley, V. (2017). SOX9: A genomic view of tissue specific expression and action. International Journal of Biochemistry & Cell Biology *87*, 18-22.

38. Sreenivasan, R., Gonen, N., and Sinclair, A. (2022). SOX genes and their role in disorders of sex development. Sexual Development *16*, 80-91.

39. Sekido, R., and Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific *Sox9* enhancer. Nature *453*, 930-934.

40. Gonen, N., Quinn, A., O'Neill, H., Koopman, P., and Lovell-Badge, R. (2017). Normal levels of *Sox9* expression in the developing mouse testis depend on the TES/TESCO enhancer, but this does not act alone. Plos Genetics *13*, e1006520.

41. Hirata, Y., Mizushima, S., Mitsukawa, S., Kon, M., Kuroki, Y., Jogahara, T., Shinohara, N., and Kuroiwa, A. (2024). Identification of a new enhancer that promotes *Sox9* expression by a comparative analysis of mouse and *Sry*-deficient Amami spiny rat. Cytogenetic and Genome Research *163*, 307-316.

42. Ridnik, M., Schoenfelder, S., and Gonen, N. (2021). Cis-regulatory control of mammalian sex determination. Sexual Development *15*, 317-334.

43. Barrionuevo, F., Hurtado, A., Kim, G., Real, F., Bakkali, M., Kopp, J., Sander, M., Scherer, G., Burgos, M., and Jimenez, R. (2016). *Sox9* and *Sox8* protect the adult testis from male-to-female genetic reprogramming and complete degeneration. eLife *5*, e15635.

44. Kobayashi, T., Yamada, F., Hashimoto, T., Abe, S., Matsuda, Y., and Kuroiwa, A. (2007). Exceptional minute sex-specific region in the X0 mammal, Ryukyu spiny rat. Chromosome Research *15*, 175-187.

45. Yee, S., and Rigby, P. (1993). The regulation of myogenin gene-expression during the embryonic-development of the mouse. Genes & Development *7*, 1277-1289.

46. Croft, B., Ohnesorg, T., Hewitt, J., Bowles, J., Quinn, A., Tan, J., Corbin, V., Pelosi, E., van den Bergen, J., Sreenivasan, R., et al. (2018). Human sex reversal is caused by duplication or deletion of core enhancers upstream of *SOX9*. Nature Communications *9*, 5319.

47. Hogan, M., Roberson, M., and Austin, R. (1989). DNA flexibility variation may dominate DNase I cleavage. Proceedings of the National Academy of Sciences of the United States of America *86*, 9273-9277.

48. Evans, C., and Aguilera, R. (2003). DNase II: genes, enzymes and function. Gene *322*, 1-15.

49. Kamiya, T., Kai, W., Tasumi, S., Oka, A., Matsunaga, T., Mizuno, N., Fujita, M., Suetake, H., Suzuki, S., Hosoya, S., et al. (2012). A trans-species missense SNP in *Amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (fugu). Plos Genetics *8*, e1002798.

50. Schieffer, K., Agarwal, V., LaHaye, S., Miller, K., Koboldt, D., Lichtenberg, T., Leraas, K., Brennan, P., Kelly, B., Crist, E., et al. (2021). *YAP1-FAM118B* fusion defines a rare subset of childhood and young adulthood meningiomas. American Journal of Surgical Pathology *45*, 329-340.

51. Boyer, L., Latek, R., and Peterson, C. (2004). The SANT domain: a unique histone-tail-binding module? Nature Reviews Molecular Cell Biology *5*, 158-163.

52. Konig, P., Giraldo, R., Chapman, L., and Rhodes, D. (1996). The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. Cell *85*, 125-136.

53. Hapgood, J., Riedemann, J., and Scherer, S. (2001). Regulation of gene expression by GC-rich DNA *cis*-elements. Cell Biology International *25*, 17-31.

54. Dey, B., Thukral, S., Krishnan, S., Chakrobarty, M., Gupta, S., Manghani, C., and Rani, V. (2012). DNA-protein interactions: methods for detection and analysis. Molecular and Cellular Biochemistry *365*, 279-299.

55. Geyer, P., and Corces, V. (1992). DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. Genes & Development *6*, 1865-1873.

56. Muravyova, E., Golovnin, A., Gracheva, E., Parshikov, A., Belenkaya, T., Pirrotta, V., and Georgiev, P. (2001). Loss of insulator activity by paired Su(Hw) chromatin insulators. Science *291*, 495-498.

57. Upadhyay, G., Chowdhury, A., Vaidyanathan, B., Kim, D., and Saleque, S. (2014). Antagonistic actions of Rcor proteins regulate LSD1 activity and cellular differentiation. Proceedings of the National Academy of Sciences of the United States of America *111*, 8071-8076.

58. Solari, F., Bateman, A., and Ahringer, J. (1999). The caenorhabditis elegans genes egl-27 and egr-1 are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning. Development *126*, 2483-2494.

59. Aasland, R., Stewart, A., and Gibson, T. (1996). The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional corepressor N-CoR and TFIIIB. Trends in Biochemical Sciences *21*, 87-88.

60. Müller, L., Hainberger, D., Stolz, V., and Ellmeier, W. (2018). NCOR1-a new player on the field of T cell development. Journal of Leukocyte Biology *104*, 1061-1068.

61. Grüne, T., Brzeski, J., Eberharter, A., Clapier, C., Corona, D., Becker, P., and Müller, C. (2003). Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. Molecular Cell *12*, 449-460.

62. Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S., and Nishimura, Y. (1994). Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. CELL *79*, 639-648.

63. Tahirov, T., Sasaki, M., Inoue-Bungo, T., Fujikawa, A., Sato, K., Kumasaka, T., Yamamoto, M., and Ogata, K. (2001). Crystals of ternary protein-DNA complexes composed of DNA-binding domains of c-Myb or v-Myb, C/EBPα or C/EBPβ and *tom-1*A promoter fragment. Acta Crystallographica Section D Biological Crystallography *57*, 1655-1658.

64. Chen, Z., Paquette, A., and Anderson, D. (1998). NRSF/REST is required *in vivo* for repression of multiple neuronal target genes during embryogenesis. Nature Genetics *20*, 136-142.

65. Negrini, S., Prada, I., D'Alessandro, R., and Meldolesi, J. (2013). REST: an oncogene or a tumor suppressor? Trends in Cell Biology *23*, 289-295.

66. Qureshi, I., Gokhan, S., and Mehler, M. (2010). REST and CoREST are transcriptional and epigenetic regulators of seminal neural fate decisions. Cell Cycle *9*, 4477-4486.

67. Shi, Y., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. Molecular Cell *19*, 857-864.

68. Bernstein, B., Humphrey, E., Erlich, R., Schneider, R., Bouman, P., Liu, J., Kouzarides, T., and Schreiber, S. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. Proceedings of the National Academy of Sciences of the United States of America *99*, 8695-8700.

69. Yang, P., Wang, Y., Chen, J., Li, H., Kang, L., Zhang, Y., Chen, S., Zhu, B., and Gao, S. (2011). RCOR2 is a subunit of the LSD1 complex that regulates ESC property and substitutes for SOX2 in reprogramming somatic cells to pluripotency. Stem Cells *29*, 791-801.

70. Myrick, D., Christopher, M., Scott, A., Simon, A., Donlin-Asp, P., Kelly, W., and Katz, D. (2017). KDM1A/LSD1 regulates the differentiation and maintenance of spermatogonia in mice. Plos One *12*, e0177473.

71. Miyamoto, Y., Taniguchi, H., Hamel, F., Silversides, D., and Viger, R. (2008). A GATA4/WT1 cooperation regulates transcription of genes required for mammalian sex determination and differentiation. BMC Molecular Biology *9*, 44.

72. Viger, R., Guittot, S., Anttonen, M., Wilson, D., and Heikinheimo, M. (2008). Role of the GATA family of transcription factors in endocrine development, function, and disease. Molecular Endocrinology *22*, 781-798.

73. Kim, G., Sock, E., Buchberger, A., Just, W., Denzer, F., Hoepffner, W., German, J., Cole, T., Mann, J., Seguin, J., et al. (2015). Copy number variation of two separate regulatory regions upstream of *SOX9* causes isolated 46,XY or 46,XX disorder of sex development. Journal of Medical Genetics *52*, 240-247.

74. Benko, S., Gordon, C., Mallet, D., Sreenivasan, R., Thauvin-Robinet, C., Brendehaug, A., Thomas, S., Bruland, O., David, M., Nicolino, M., et al. (2011). Disruption of a long distance regulatory region upstream of *SOX9* in isolated disorders of sex development. Journal of Medical Genetics *48*, 825-830.

75. Hyon, C., Chantot-Bastaraud, S., Harbuz, R., Bhouri, R., Perrot, N., Peycelon, M., Sibony, M., Rojo, S., Piguel, X., Bilan, F., et al. (2015). Refining the regulatory region upstream of *SOX9* associated with 46,XX testicular disorders of sex development (DSD). American Journal of Medical Genetics Part A *167*, 1851-1858.

76. Karl, J., and Capel, B. (1998). Sertoli cells of the mouse testis originate from the coelomic epithelium. Developmental Biology *203*, 323-333.

77. Capel, B. (2000). The battle of the sexes. Mechanisms of Development *92*, 89-103.

78. Bullejos, M., and Koopman, P. (2001). Spatially dynamic expression of *Sry* in mouse genital ridges. Developmental Dynamics *221*, 201-205.

79. Kidokoro, T., Matoba, S., Hiramatsu, R., Fujisawa, M., Kanai-Azuma, M., Taya, C., Kurohmaru, M., Kawakami, H., Hayashi, Y., Kanai, Y., and Yonekawa, H. (2005). Influence on spatiotemporal patterns of a male-specific *Sox9* activation by ectopic *Sry* expression during early phases of testis differentiation in mice. Developmental Biology *278*, 511-525.

80. daSilva, S., Hacker, A., Harley, V., Goodfellow, P., Swain, A., and Lovell-Badge, R. (1996). *Sox9* expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. Nature Genetics *14*, 62-68.

81. Sekido, R., Bar, I., Narváez, V., Penny, G., and Lovell-Badge, R. (2004). SOX9 is up-regulated by the transient expression of SRY specifically in Sertoli cell precursors. Developmental Biology *274*, 271-279.

82. Kim, Y., Kobayashi, A., Sekido, R., DiNapoli, L., Brennan, J., Chaboissier, M., Poulat, F., Behringer, R., Lovell-Badge, R., and Capel, B. (2006). *Fgf9* and *Wnt4* act as antagonistic signals to regulate mammalian sex determination. Plos Biology *4*, 1000-1009, e187.

83. Yao, H., Whoriskey, W., and Capel, B. (2002). Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. Genes & Development *16*, 1433-1440.

84. Ottolenghi, C., Omari, S., Garcia-Ortiz, J., Uda, M., Crisponi, L., Forabosco, A., Pilia, G., and Schlessinger, D. (2005). *Foxl2* is required for commitment to ovary differentiation. Human Molecular Genetics *14*, 2053-2062.

85. Chassot, A., Ranc, F., Gregoire, E., Roepers-Gajadien, H., Taketo, M., Camerino, G., de Rooij, D., Schedl, A., and Chaboissier, M. (2008). Activation of β-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. Human Molecular Genetics *17*, 1264-1277.

86. Stévant, I., Neirijnck, Y., Borel, C., Escoffier, J., Smith, L., Antonarakis, S., Dermitzakis, E., and Nef, S. (2018). Deciphering cell lineage specification during male sex determination with single-cell RNA sequencing. Cell Reports *22*, 1589-1599.

87. Stévant, I., Kuehne, F., Greenfield, A., Chaboissier, M., Dermitzakis, E., and Nef, S. (2019). Dissecting cell lineage specification and sex fate determination in gonadal somatic cells using single-cell transcriptomics. Cell Reports *26*, 3272-3283.

88. Liu, C., Bingham, N., Parker, K., and Yao, H. (2009). Sex-specific roles of β-catenin in mouse gonadal development. Human Molecular Genetics *18*, 405-417.

89. Hacker, A., Capel, B., Goodfellow, P., and Lovellbadge, R. (1995). Expression of *Sry*, the mouse sex-determining gene. Development *121*, 1603-1614.

90. de Santa Barbara, P., Moniot, B., Poulat, F., and Berta, P. (2000). Expression and subcellular localization of SF-1, SOX9, WT1, and AMH proteins during early human testicular development. Developmental Dynamics *217*, 293-298.

91. Wu, N., Yu, A., Zhu, H., and Lin, X. (2012). Effective silencing of *Sry* gene with rna interference in developing mouse embryos resulted in feminization of XY gonad. Journal of Biomedicine and Biotechnology *2012*, 343891.

92. She, Z., and Yang, W. (2017). *Sry* and *SoxE* genes: How they participate in mammalian sex determination and gonadal development? Seminars in Cell & Developmental Biology *63*, 13-22.

93. Schmidt, D., Ovitt, C., Anlag, K., Fehsenfeld, S., Gredsted, L., Treier, A., and Treier, M. (2004). The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. Development *131*, 933-942.

94. Gregoire, E., Lavery, R., Chassot, A., Akiyama, H., Treier, M., Behringer, R., and Chaboissier, M. (2011). Transient development of ovotestes in XX *Sox9* transgenic mice. Developmental Biology *349*, 65-77.

95. Karim, F., Urness, L., Thummel, C., Klemsz, M., Mckercher, S., Celada, A., Vanbeveren, C., Maki, R., Gunther, C., Nye, J., and Graves, B. (1990). The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. Genes & Development *4*, 1451-1453.

96. Potter, M., Buijs, A., Kreider, B., van Rompaey, L., and Grosveld, G. (2000). Identification and characterization of a new human ETS-family transcription factor, *TEL2*, that is expressed in hematopoietic tissues and can associate with *TEL1*/*ETV6*. BLOOD *95*, 3341-3348.

97. Buijs, A., van Rompaey, L., Molijn, A., Davis, J., Vertegaal, A., Potter, M., Adams, C., van Baal, S., Zwarthoff, E., Roussel, M., and Grosveld, G. (2000). The MN1-TEL fusion protein, encoded by the translocation (12;22)(p13;q11) in myeloid leukemia, is a transcription factor with transforming activity. Molecular and Cellular Biology *20*, 9281-9293.

98. Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S., and Elledge, S. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes & Development *23*, 837-848.

99. Rauluseviciute, I., Riudavets-Puig, R., Blanc-Mathieu, R., Castro-Mondragon, J., Ferenc, K., Kumar, V., Lemma, R., Lucas, J., Chéneby, J., Baranasic, D., et al. (2023). JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. Nucleic Acids Research *52*, D174-D182.

100. Marco, E., García-Nieto, R., and Gago, F. (2003). Assessment by molecular dynamics simulations of the structural determinants of DNA-binding specificity for transcription factor Sp1. Journal of Molecular Biology *328*, 9-32.

101. Flores, O., Ha, I., and Reinberg, D. (1990). Factors involved in specific transcription by mammalian RNA polymerase II - purification and subunit composition of transcription factor IIF. Journal of Biological Chemistry *265*, 5629-5634.

102. Hisatake, K., Ohta, T., Takada, R., Guermah, M., Horikoshi, M., Nakatani, Y., and Roeder, R. (1995). Evolutionary conservation of human TATA-binding-polypeptide-associated factors TAF(II)31 and TAF(II)80 and interactions of

TAF(II)80 with other TAFs and with general transcription factors. Proceedings of the National Academy of Sciences of the United States of America *92*, 8195-8199.

103. Forget, D., Robert, F., Grondin, G., Burton, Z., Greenblatt, J., and Coulombe, B. (1997). RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box. Proceedings of the National Academy of Sciences of the United States of America *94*, 7150-7155.

104. Gutierrez-Hartmann, A., Duval, D., and Bradford, A. (2007). ETS transcription factors in endocrine systems. Trends in Endocrinology and Metabolism *18*, 150-158.

105. Oikawa, T., and Yamada, T. (2003). Molecular biology of the Ets family of transcription factors. Gene *303*, 11-34.

106. Oettgen, P., Akbarali, Y., Boltax, J., Best, J., Kunsch, C., and Libermann, T. (1996). Characterization of NERF, a novel transcription factor related to the Ets factor ELF-1. Molecular and Cellular Biology *16*, 5091-5106.

107. Cho, J., Akbarali, Y., Zerbini, L., Gu, X., Boltax, J., Wang, Y., Oettgen, P., Zhang, D., and Libermann, T. (2004). Isoforms of the Ets transcription factor NERF/ELF-2 physically interact with AML1 and mediate opposing effects on AML1-mediated transcription of the B cell-specific *blk* gene. Journal of Biological Chemistry *279*, 19512-19522.

108. Guan, F., Bailey, C., Metierre, C., O'Young, P., Gao, D., Khoo, T., Holst, J., and Rasko, J. (2017). The antiproliferative ELF2 isoform, ELF2B, induces apoptosis

in vitro and perturbs early lymphocytic development in vivo. Journal of Hematology & Oncology *10*, 75.

109. Ji, H., Gupta, A., Okamoto, S., Blum, M., Tan, L., Goldring, M., Lacy, E., Roy, A., and Terhorst, C. (2002). T cell-specific expression of the murine *CD3*δ promoter. Journal of Biological Chemistry *277*, 47898-47906.

110. Zhang, B., Tomita, Y., Qiu, Y., He, J., Morii, E., Noguchi, S., and Aozasa, K. (2007). E74-like factor 2 regulates *valosin-containing protein* expression. Biochemical and Biophysical Research Communications *356*, 536-541.

111. Golub, T., Barker, G., Lovett, M., and Gilliland, D. (1994). Fusion of PDGF receptor β to a novel *ets*-like gene, *tel*, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. Cell *77*, 307-316.

112. Wang, L., Kuo, F., Fujiwara, Y., Gilliland, D., Golub, T., and Orkin, S. (1997). Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. EMBO Journal *16*, 4374-4383.

113. Yamagata, T., Maki, K., Waga, K., and Mitani, K. (2006). TEL/ETV6 induces apoptosis in 32D cells through p53-dependent pathways. Biochemical and Biophysical Research Communications *347*, 517-526.

114. Shen, Z., Wei, L., Yu, Z., Yao, Z., Cheng, J., Wang, Y., Song, X., and Li, M. (2021). The roles of TRIMs in antiviral innate immune signaling. Frontiers in Cellular and Infection Microbiology *11*, 628275.

115. Isomura, T., Tamiyakoizumi, K., Suzuki, M., Yoshida, S., Taniguchi, M., Matsuyama, M., Ishigaki, T., Sakuma, S., and Takahashi, M. (1992). RFP is a

DNA binding protein associated with the nuclear matrix. Nucleic Acids Research *20*, 5305-5310.

116. Shimono, Y., Murakami, H., Hasegawa, Y., and Takahashi, M. (2000). RET finger protein is a transcriptional repressor and interacts with enhancer of polycomb that has dual transcriptional functions. Journal of Biological Chemistry *275*, 39411-39419.

117. Bloor, A., Kotsopoulou, E., Hayward, P., Champion, B., and Green, A. (2005). RFP represses transcriptional activation by bHLH transcription factors. Oncogene *24*, 6729-6736.

118. Wang, T., Xia, C., Dong, Y., Chen, X., Wang, J., and Du, J. (2018). Trim27 confers myeloid hematopoiesis competitiveness by up-regulating myeloid master genes. Journal of Leukocyte Biology *104*, 799-809.

119. Zhao, D., Qiang, L., Lei, Z., Ge, P., Lu, Z., Wang, Y., Zhang, X., Qiang, Y., Li, B., Pang, Y., et al. (2024). TRIM27 elicits protective immunity against tuberculosis by activating TFEB-mediated autophagy flux. Autophagy *20*, 1483-1504.

120. Cao, T., Shannon, M., Handel, M., and Etkin, L. (1996). Mouse ret finger protein (*rfp*) proto-oncogene is expressed at specific stages of mouse spermatogenesis. Developmental Genetics *19*, 309-320.

121. Tezel, G., Shimono, Y., Murakumo, Y., Kawai, K., Fukuda, T., Iwahashi, N., and Takahashi, M. (2002). Role for O-glycosylation of RFP in the interaction with

enhancer of polycomb. Biochemical and Biophysical Research Communications *290*, 409-414.

122. Zhuang, X., Tang, W., Feng, X., Liu, C., Zhu, J., Yan, J., Liu, D., Liu, P., and Qiao, J. (2016). Trim27 interacts with Slx2, is associated with meiotic processes during spermatogenesis. Cell Cycle *15*, 2576-2584.

123. Wolf, I., and Rohrschneider, L. (1999). Fiz1, a novel zinc finger protein interacting with the receptor tyrosine kinase Flt3. Journal of Biological Chemistry *274*, 21478-21484.

124. Mali, R., Peng, G., Zhang, X., Dang, L., Chen, S., and Mitton, K. (2008). FIZ1 is part of the regulatory protein complex on active photoreceptor-specific gene promoters *in vivo*. BMC Molecular Biology *9*, 87.

125. Cohen, H., Bossone, S., Zhu, G., McDonald, G., and Sukhatme, V. (1997). Sp1 is a critical regulator of the *Wilms*' *tumor-1* gene. Journal of Biological Chemistry *272*, 2901-2913.

126. Assumpçao, J., Ferraz, L., Benedetti, C., Maciel-Guerra, A., Guerra, G., Marques-de-Faria, A., Baptista, T., and de Mello, M. (2005). A naturally occurring deletion in the *SRY* promoter region affecting the Sp1 binding site is associated with sex reversal. Journal of Endocrinological Investigation *28*, 651-656.

127. Roeder, R. (1996). Nuclear RNA polymerases: role of general initiation factors and cofactors in eukaryotic transcription. Methods in Enzymology *273*, 165-171.

128. Kanai, Y., and Koopman, P. (1999). Structural and functional characterization of the mouse *Sox9* promoter: implications for campomelic dysplasia. Human Molecular Genetics *8*, 691-696.

129. DiTacchio, L., Bowles, J., Shin, S., Lim, D., Koopman, P., and Janknecht, R. (2012). Transcription factors ER71/ETV2 and SOX9 participate in a positive feedback loop in fetal and adult mouse testis. Journal of Biological Chemistry *287*, 23657-23666.

130. Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T., Zon, L., Yuan, G., et al. (2016). Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. Developmental Cell *36*, 9-23.

131. Okuno, M., Mochimaru, Y., Matsuoka, K., Yamabe, T., Matiz-Ceron, L., Jogahara, T., Toyoda, A., Kuroiwa, A., and Itoh, T. (2023). Chromosomal-level assembly of *Tokudaia osimensis*, *Tokudaia tokunoshimensis*, and *Tokudaia muenninki* genomes. Scientific Data *10*, 927.

132. Burgin, C., Colella, J., Kahn, P., and Upham, N. (2018). How many species of mammals are there? Journal of Mammalogy *99*, 1-14.

133. Graves, J. (2006). Sex chromosome specialization and degeneration in mammals. Cell *124*, 901-914.

134. Cortez, D., Marin, R., Toledo-Flores, D., Froidevaux, L., Liechti, A., Waters, P., Grützner, F., and Kaessmann, H. (2014). Origins and functional evolution of Y chromosomes across mammals. Nature *508*, 488-493.

135. Grützner, F., Rens, W., Tsend-Ayush, E., El-Mogharbel, N., O'Brien, P., Jones, R., Ferguson-Smith, M., and Graves, J. (2004). In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. Nature *432*, 913-917.

136. Wallis, M., Waters, P., and Graves, J. (2008). Sex determination in mammals – before and after the evolution of *SRY*. Cellular and Molecular Life Sciences *65*, 3182-3195.

137. Rens, W., Grützner, F., O'Brien, P., Fairclough, H., Graves, J., and Ferguson-Smith, M. (2004). Resolution and evolution of the duck-billed platypus karyotype with an $X_1Y_1X_2Y_2X_3Y_3X_4Y_4X_5Y_5$ male sex chromosome constitution. Proceedings of the National Academy of Sciences of the United States of America *101*, 16257-16261.

138. Veyrunes, F., Waters, P., Miethke, P., Rens, W., McMillan, D., Alsop, A., Grüzner, F., Deakin, J., Whittington, C., Schatzkamer, K., et al. (2008). Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Research *18*, 965-973.

139. Just, W., Rau, W., Vogel, W., Akhverdian, M., Fredga, K., Graves, J., and Lyapunova, E. (1995). Absence of *Sry* in species of the vole *Ellobius*. Nature Genetics *11*, 117-118.

140. Kolomiets, O., Vorontsov, N., Lyapunova, E., and Mazurova, T. (1991). Ultrastructure, meiotic behavior, and evolution of sex chromosomes of the genus *Ellobius*. Genetica *84*, 179-189.

141. Romanenko, S., Sitnikova, N., Serdukova, N., Perelman, P., Rubtsova, N., Bakloushinskaya, I., Lyapunova, E., Just, W., Ferguson-Smith, M., Yang, F., and Graphodatsky, A. (2007). Chromosomal evolution of Arvicolinae (Cricetidae, Rodentia). II. The genome homology of two mole voles (genus *Ellobius*), the field vole and golden hamster revealed by comparative chromosome painting. Chromosome Research *15*, 891-897.

142. Bagheri-Fam, S., Sreenivasan, R., Bernard, P., Knower, K., Sekido, R., Lovell-Badge, R., Just, W., and Harley, V. (2012). *Sox9* gene regulation and the loss of the XY/XX sex-determining mechanism in the mole vole *Ellobius lutescens*. Chromosome Research *20*, 191-199.

143. Matthey, R. (1958). New type of chromosomal sex determination in the mammals *Ellobius lutescens* Th. and *Microtus* (*Chilotus*) *oregoni* Bachm. (Muridae, Microtinae). Experientia *14*, 240-241.

144. Aitken, R., and Graves, J. (2002). Human spermatozoa: The future of sex. Nature *415*, 963.

145. Hughes, J., Skaletsky, H., Pyntikova, T., Graves, T., van Daalen, S., Minx, P., Fulton, R., McGrath, S., Locke, D., Friedman, C., et al. (2010). Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. Nature *463*, 536-539.

146. Hughes, J., Skaletsky, H., Brown, L., Pyntikova, T., Graves, T., Fulton, R., Dugan, S., Ding, Y., Buhay, C., Kremitzki, C., et al. (2012). Strict evolutionary

conservation followed rapid gene loss on human and rhesus Y chromosomes. Nature *483*, 82-86.

Main Publication

Mitsukawa, **S.**, Mizushima, S., Kimura, Y., Kuroiwa, A. (2025) Identification of candidate transcription factors that regulate *Sox9* expression through the testis-specific enhancer in the Amami spiny rat, (*Tokudaia osimensis*), an XO/XO mammal. Genes and Genetic systems *100*, 25-00024.

Other publication

Hirata, Y., Mizushima, S., **Mitsukawa**, **S.**, Kon, M., Kuroki, Y., Jogahara, T., Shinohara, N., Kuroiwa, A. (2023) Identification of a new enhancer that promotes *Sox9* expression by a comparative analysis of mouse and *Sry*-deficient Amami Spiny Rat. Cytogenetic and Genome Research *163*, 307-316.

Presentations

Oral presentations

1. **〇光川祥一朗**, 水島秀成, 黒岩麻里 「アマミトゲネズミ精巣特異的エンハンサーを介して *SOX9* 発現を制御する転写因子の同定」, 日本動物学会北海道支部第 66 回大会, オンライン開催, 2022 年 3 月 19 日

2. **〇光川 祥一朗**, 水島 秀成, 黒岩 麻里 「アマミトゲネズミにおける精巣特異的エンハンサーを介した *SOX9* 遺伝子制御の研究」, 日本動物学会第 93 回 早稲田大会, 早稲田大学早稲田キャンパス, 2022 年 9 月 10 日

3. **〇光川 祥一朗**, 水島 秀成, 黒岩 麻里 「アマミトゲネズミ精巣特異的エンハンサーを介して *SOX9* を制御する転写因子の同定」, 日本遺伝学会 第 94 回大会, 北海道大学工学部, 2022 年 9 月 15 日

4. **〇光川 祥一朗**, 水島 秀成, 黒岩 麻里 「Y 染色体をもたない哺乳類種における精巣特異的エンハンサーを介した *Sox9* 遺伝子制御の研究」, 染色体学会 第 74 回年会, オンライン開催, 2023 年 10 月 21 日

5. **〇光川 祥一朗**, 水島 秀成, 黒岩 麻里 「アマミトゲネズミ精巣特異的エンハンサーを介して *Sox9* 発現を制御する転写因子の探索」, 日本動物学会北海道支部第 68 回大会, 北海道大学理学部, 2024 年 3 月 16 日(土)

6. **〇光川 祥一朗**, 水島 秀成, 黒岩 麻里 「アマミトゲネズミにおける精巣特異的エンハンサーを介して *Sox9* 遺伝子を制御するトランス因子の同定」, 日本動物学会第 95 回 長崎大会, 長崎大学文教キャンパス, 2024 年 9 月 12 日

Poster presentations

1. **〇光川 祥一朗**, 水島 秀成, 黒岩 麻里 「Studies on the *Sox9* gene regulation through a testis-specific enhancer in a mammalian species lacking the *Sry* gene」, 第 45 回 日本分子生物学会年会, 幕張メッセ, 2022 年 11 月 30 日

2. **〇光川 祥一朗**, 水島 秀成, 黒岩 麻里 「アマミトゲネズミ *Sox9* 遺伝子を制御する新規転写因子の探索」, 第 47 回 日本分子生物学会年会, マリンメッセ福岡, 2024 年 11 月 28 日