January 2021

Intron Retention Of JARID2 Regulates Gene Expression During Temperature-Dependent Sex Determination

Jacob Robert Bierstedt

INTRON RETENTION OF JARID2 REGULATES GENE EXPRESSION DURING TEMPERATURE-DEPENDENT SEX DETERMINATION

by

Jacob Robert Bierstedt
Bachelor of Science, University of North Dakota, 2017

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota

May
2021

Name: Jacob Bierstedt

Degree: Master of Science

This document, submitted in partial fulfillment of the requirements for the degree from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

DocuSigned by:
Turk Rhen
32ECCCC21FF5480...
Turk Rhen

DocuSigned by:
Tristan Darland
0825C3322229491...
Tristan Darland

DocuSigned by:
Manu Manu
E0D49834CCAC470...
Manu Manu

This document is being submitted by the appointed advisory committee as having met all the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

DocuSigned by:
Chris Nelson
2E0AE088C733403...
Chris Nelson
Dean of the School of Graduate Studies

5/3/2021
Date

PERMISSION

Title Intron Retention of JARID2 Regulates Gene Expression During Temperature-dependent Sex Determination

Department Biology

Degree Master of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Jacob Bierstedt
05/04/2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I wish to thank the faculty of the Department of Biology and members of my advisory Committee for their guidance during my time at the University of North Dakota. I also wish to thank the members of the Rhen Lab, past and present.

ABSTRACT

Temperature-dependent sex determination (TSD) is the developmental process in which cell fate decisions in the bipotential gonads are influenced by temperature and drive the gonad toward one of two distinct organs, an ovary or testis. In species with TSD, temperature establishes the gene expression patterns required for cell fate decisions and gonad differentiation. However, the precise mechanism that temperature regulates gene expression is unknown. Here we have shown that *Jarid2*, which encodes a Polycomb Repressive Complex 2 (PRC2) accessory protein, exhibits temperature-dependent intron retention that results in three distinct transcripts encoding distinct proteins with functional similarities and differences in TSD. We found in a long-read sequencing assay that the transcripts encoding these proteins exhibited different expression patterns in response to a temperature shift from a male- to female-producing temperature. We also found that overexpression of these proteins in a bipotential gonad cell culture model yielded both similar and differing effects on gene expression, suggesting that they may play unique roles in regulating gene expression in TSD. Our results demonstrate that JARID2 is alternately spliced and the resulting isoforms are likely involved in establishing the transcriptional profiles required for cell fate decisions in TSD. We anticipate that this study will provide a foundation from which to further probe the function of JARID2 in the thermal response during TSD.

To Allie, Twister, Jax, Jager,
and Rizzo

CHAPTER I

MECHANISMS OF TRANSCRIPTIONAL REGULATION DURING SEX DETERMINATION

Introduction

There are two main types of reproduction observed in eukaryotes: asexual and sexual reproduction. The latter relies on the presence of meiosis to produce sperm and eggs and fusion of gametes within populations of organisms. In addition to the ability of sperm to fertilize eggs, there must be compatibility of the genomes of organisms within a population. It is this compatibility combined with the incompatibility of the genomes of other organisms that drives diversification and speciation in eukaryotes through the ability of individuals to reproduce with others of a highly specific genomic makeup (Lynch & Force, 2000; Muller, 1942; Orr, 1996).

In metazoans, sex involves the production of one of two types of haploid gametes (sperm and egg) that fuse with the other type to give rise to a diploid zygote. Production of these gametes relies on the development of sex-specific reproductive organs (testis and ovary) that occurs during a process known as sex determination, which can be simply be described as the process by which a bipotential gonad differentiates into an ovary or a testis. Sex determining mechanisms are diverse, relying on genetic or environmental signals to coordinate gene expression profiles that define the lineage trajectories of cells in the developing gonad.

Temperature-dependent sex determination (TSD) is a form of sex determination in which the primary stimulus governing the fate of the bipotential gonad is the ambient temperature of the embryo during a specific period of development. In this chapter I discuss the transcriptional regulation of TSD and the potential role of JARID2 as a thermosensitive regulator of transcription in the common snapping turtle, *Chelydra serpentina.*

Sex Determining Mechanisms in Vertebrates

The majority of vertebrates reproduce sexually. During sexual reproduction, two haploid gametes fuse during syngamy to create a diploid zygote, which gives rise to every cell in the developing organism. In contrast to offspring produced by asexual reproduction, those produced by sexual reproduction contain a combination of genomic information derived from each parent. This, along with recombination between homologous chromosomes during meiosis, produces new genetic combinations and increases adaptive capacity.

In amniotic vertebrates, morphological differences between the sexes begin to manifest through differentiation of specific cell types in the gonads. During embryogenesis, paired genital ridges form as the result of the proliferation of the coelomic epithelium and the underlying mesenchyme of the mesonephros (DeFalco & Capel, 2009). These parallel ridges appear as thickenings that lie along the caudal-rostral axis of the embryo. Each genital ridge is a single primordium capable of developing into either an ovary or a testis, which is determined by the specification of two main lineages of bipotential somatic cells (Karl & Capel, 1998). Supporting cell precursors give rise to granulosa cells or Sertoli cells in the ovary or testis, respectively (Stevant et al., 2019; Stevant et al., 2018). A second population of bipotential somatic cells, steroidogenic progenitors, give rise to theca cells or Leydig cells in the ovary or testis, respectively (Stevant et al., 2019; Stevant et al., 2018). Differentiation of these two somatic cell lineages into four distinct cell types and the morphogenesis of the ovary or testis is known as sex determination.

The determination of supporting cell fate as granulosa or Sertoli cells is controlled by complex transcriptional networks that tend to be triggered by a few sex-determining factors early in development (Marin & Baker, 1998). The result is highly contrasting profiles of gene expression between pre-Sertoli and pre-granulosa cells (Jin et al., 2001; Ranz, Castillo-Davis, Meiklejohn, &

Hartl, 2003). Evolutionary variation in sex-determining mechanisms is substantial, even among vertebrates, but in all cases definitive gene expression patterns are required for differentiation of key cell types in developing gonads. The gene regulatory networks for determination of cell fate and gonad fate have been studied most intensively in mammals, which have genotypic sex determination (GSD) with distinct sex chromosomes.

In placental mammals, sex is determined by the presence or absence of *Sry*, a gene encoded on the Y-chromosome that invokes changes in the regulatory networks that cause supporting cell differentiation into Sertoli cells, resulting in a male phenotype (Lin & Capel, 2015). Precursors of many of the somatic cells in the bipotential gonad arise from a population of cells expressing NR5A1/SF1 (steroidogenic factor 1). At this stage of development, many genes associated with determination of sexual fate are expressed at the same level in XX and XY gonads. In XY gonads, SRY expression increases briefly and initiates the pathway associated with a male fate. The primary SRY target is SOX9, which increases sharply along with FGF9 and PTGDS. FGF9 exerts positive feedback to maintain SOX9 expression. These gene products in turn regulate the expression of other genes involved in cell differentiation and gonad morphogenesis, including inhibition of Wnt signaling (Y. Kim et al., 2006; Lin & Capel, 2015). In XX gonads, WNT4 and R-Spondin1 signaling stabilize beta-catenin, leading to determination of ovarian fate (Lin & Capel, 2015; Tevosian & Manuylov, 2008). Canonical Wnt signaling stabilizes beta-catenin in the nucleus where it binds T-cell factor (TCF)/lymphoid enhancer factor (LEF) which leads to the transcriptional activation of target genes (Tevosian & Manuylov, 2008). Wnt signaling in XX gonads also antagonizes FGF9 and SOX9 expression. Much of what is understood about the molecular mechanisms of sex determination comes from the study of mammalian sex

determination. Although there are significant similarities in the morphogenetic process of sex determination, there are many differences in sex-determining mechanisms amongst vertebrates.

Sex determination amongst reptiles is especially diverse. The sex of many lizard species is determined by sex chromosomes analogous (but not homologous) to those found in mammals and birds, while other lizard species lack sex chromosomes and commitment to sexual fate is governed by environmental temperature (Rhen, Schroeder, Sakata, Huang, & Crews, 2011; Viets, 1994). Many turtle species, including the snapping turtle *Chelydra serpentina*, and all crocodilians undergo temperature-dependent sex determination (TSD), a process in which sexual phenotype is determined by the ambient temperature of the embryo during development.

In species with TSD, constant temperature incubation of eggs yields predictable offspring sex-ratios. There are three typical patterns of temperature dependent sex determination (Ewert, Jackson, & Nelson, 1994). In some species, low temperatures produce males and high temperatures produce females. In other species, low temperatures produce females and high temperatures produce males. Lastly, and in the case of *Chelydra serpentina*, low and high temperatures produce females and intermediate temperatures produce males (Rhen & Lang, 1998). In all species with TSD, male-producing temperatures (MPT) are the temperatures that produce 100 percent male offspring, female-producing temperatures (FPT) are the temperatures that produce 100 percent female offspring, and pivotal temperatures (PvT) are the temperatures that produce an even sex ratio with 50 percent male and 50 percent female offspring (Rhen & Lang, 1998). The transitional range of temperatures is the range of temperatures between MPTs and FPTs that produce mixed sex ratios.

Embryos of species with temperature-dependent sex determination exhibit complete competence to develop as male or female, with sexual phenotype being primarily dependent on

the incubation temperature of the egg (Rhen, Fagerlie, Schroeder, Crossley, & Lang, 2015; Rhen & Lang, 1994, 1998; Rhen, Metzger, Schroeder, & Woodward, 2007). Studies have shown that there is a degree of heritability that accounts for variation in pivotal temperatures between individuals and populations of the same species, but in most cases, there are temperatures that produce exclusively one sex or the other sex (Bull, Vogt, & Bulmer, 1982; Rhen & Lang, 1998; Rhen et al., 2011).

The temperature sensitive period (TSP) is the time during embryonic development when embryonic gonads are sexually plastic and a change in temperature is capable of having predictable effects on expression of sex determining genes and offspring sex ratios. In *Chelydra serpentina,* the TSP broadly encompasses embryonic stages 13 to 20 (Rhen et al., 2015). However, this species is most sensitive to feminization during a 5-day window starting at embryonic stage 17 to 18 (Rhen et al., 2015). During this period, the gonads are still capable of responding to the polarizing temperature signal. Robust differences in transcriptional patterns manifest between gonads exposed to MPT and those exposed to FPT (Rhen et al., 2015; Rhen et al., 2007). A key set of genes encode proteins that have proven to be markers for sex determination in *Chelydra serpentina*, including FOXL2, SOX9, AR, DMRT1, and others (Rhen et al., 2007; Rhen & Schroeder, 2010). However, these genes do not respond immediately after a change in temperature and therefore are likely to be downstream in the regulatory networks that lead to ovarian or testicular fate. In order to understand the impact of temperature on the gene regulatory networks that govern sex determination, we need to identify and elucidate the function of genes that quickly and directly respond to a temperature signal.

Epigenetic Regulation of Gene Expression and Cell Fate Decisions

Genomic DNA of eukaryotic organisms is organized and condensed into nucleosomes consisting of 146 base pairs of DNA wrapped around an octamer of four types of histone proteins (H2A, H2B, H3, and H4). Nucleosomes coordinate the structure and accessibility of DNA and are subject to a variety of modifications that further determine the organization of chromatin. Epigenetic modifications, such as DNA methylation and histone tail post-translational modifications, lead to mitotically heritable changes in chromatin structure and accessibility of DNA by transcription factors. In order for transcription to occur, the promoter for a particular gene must be accessible to transcription factors. Certain epigenetic modifications, such as trimethylation of histone H3 at Lysine 27 or methylation on the 5^{th} carbon of cytosine residues in genomic DNA, are associated with a heterochromatic state and repressed transcription. Others, such as histone H3 and H4 acetylation at multiple lysine residues and histone H3 Lysine 4 di- or trimethylation, are associated with chromatin accessibility and active transcription. Many histone modifications also serve as substrates or binding sites for proteins that directly or indirectly influence chromatin state or transcription initiation. Histone acetylation is primarily propagated by histone acetyltransferase complexes with specificity that is generally lower than that of histone methyltransferases which propagate broad-scale histone methylation as well as methylation at highly specific loci (Bernstein, Meissner, & Lander, 2007). Histone modification and chromatin accessibility are also associated with chromosomal interactions that together regulate gene expression patterns during development (Apostolou et al., 2013). Genes, cell-type specific signaling, and environmental factors govern epigenetic modifications and the resulting regulation of transcriptional states and genomic topology, which, in turn, are responsible for the establishment and maintenance of cell fate decisions.

One of the most remarkable features of the development of multicellular organisms is the ability of a single cell, the fertilized egg, to give rise to specialized cell-types which leads to tissue formation and ultimately organogenesis. The first cell fate decision in the development of the mouse embryo provides an example of the importance of epigenetic modifications in cell fate commitment. The first decision is the transition from identical cells to a distinction between the pluripotent inner cell mass and the extraembryonic trophectoderm. Following zygotic genome activation, the transcriptional states of the progenitors of the two lineages are nearly identical but once they are separated, inner cells resist differentiation by expressing genes associated with pluripotency: SALL4, OCT4, NANOG, and SOX2. Outer cells express CDX2, which downregulates pluripotency genes that are expressed in the inner cell mass and leads to differentiation of trophectodermal cells (Zernicka-Goetz, Morris, & Bruce, 2009). The coordinated expression of these pathways between cell lineages can be explained by differences in epigenetic modifications between each cell type progenitor. Single-cell sequencing and lineage-tracing methods have found heterogenous expression of epigenetic modifiers in early developmental states (2- and 4-cell stages) that result in differential histone modifications between cells destined to become pluripotent inner cells or trophectoderm cells (Burton et al., 2013; Torres-Padilla, Parfitt, Kouzarides, & Zernicka-Goetz, 2007; Zernicka-Goetz et al., 2009). Differential epigenetic activation of transcription factors regulating CDX2 expression correspond to the initial expression of CDX2 in trophectoderm-destined cells which is followed by the antagonism of pluripotency genes by CDX2 and eventually the loss of lineage plasticity. Regulation of gene expression is multifaceted but can generally be correlated with the accessibility of chromatin. In many cases, such as the specification of early embryonic cells, chromatin modifiers play a role in setting the stage for differential expression of other genes involved in downstream processes (Szutorisz et al.,

2005; Xu et al., 2011). This phenomenon can be observed in many other cell-fate decisions throughout the development of eukaryotic organisms and can explain part of the process of the specification of cell lineages and tissues.

Specification of Gonadal Progenitor Cells

Development of the gonad from a thickened coelomic epithelium to a fully functional ovary or testis requires several cell-fate decisions. The first is the transition from gonadal progenitor cells to supporting cells or steroidogenic cells. This is followed by the differentiation of supporting cells into Sertoli cells in males or granulosa cells in females and the differentiation of steroidogenic cells into Leydig cells in males or theca cells in females. In order to understand these events and the mechanisms driving them, it is important to define key regulatory elements controlling the unique transcriptional profiles of each cell type.

Commitment of coelomic epithelial cells to develop as gonadal progenitor cells is the earliest stage of gonad development and occurs the same way in both sexes. In mammals, the insulin/insulin-like growth factor pathway (IGF pathway) acts to promote expression of key genes involved in initial gonad formation and lineage restriction such as steroidogenic factor 1 (NR5A1) (Luo, Ikeda, & Parker, 1994; Pitetti et al., 2013). Single-cell transcriptomics and traditional methods have shown that expression of NR5A1 is monomorphic between genetic females and males in mammals and variably dimorphic between male- and female-producing temperatures during sex determination in TSD species (Rhen et al., 2007; Stevant et al., 2019) (Fleming, Wibbels, Skipper, & Crews, 1999; Stevant et al., 2018; Western, Harry, Marshall Graves, & Sinclair, 2000). However, NR5A1 is detected throughout gonad development from genital ridge formation to differentiation of the gonad in several species with TSD, mammals, and birds,

suggesting that its function in early gonadogenesis is likely conserved. While the role of NR5A1 in gonad fate determination in TSD should not be discounted, observed differences in expression between species are potentially confounded by several factors including the inclusion of steroidogenic adrenal tissue and temperature-dependent shifts in cell type proportions when entire adrenal-kidney-gonad complexes are analyzed rather than pure gonadal tissue (C. M. Shoemaker & Crews, 2009). Nonetheless, NR5A1 is widely expressed through gonadal cell types and its expression may be used as a marker for the commitment of coelomic epithelial cells to early progenitor cells of the gonad.

At the beginning of female sex determination there is one population of gonadal progenitor cells that is actively proliferating. Included in this population are the cells that will differentiate into pre-granulosa cells and those that will differentiate into steroidogenic cells. In mouse, a single-cell time-course gene expression profiling experiment has shown that the transcriptional profile of early progenitor cells is characterized by up-regulation of many genes involved in mitosis, mesonephros development, epithelial morphogenesis, and stem-cell development (Stevant et al., 2019; Stevant et al., 2018). These genes are subsequently down-regulated during commitment to supporting cell lineages. There are also very few genes displaying sexually dimorphic expression patterns. Because this process is not sex-specific and the same cell lineages exist in species with TSD, it is likely that similar gene regulatory networks control commitment to supporting cell lineages in mammals and species with TSD.

Differentiation of Steroidogenic Lineages in the Gonad

Steroidogenic cells in the adult ovary or testis originate from two populations of cells. The majority arise from WT1 expressing progenitor cells in the gonadal primordium and a subset arise

from mesonephros (Liu, Peng, Matzuk, & Yao, 2015). These cell lineages differ in their origin as well as transcriptionally. The functional differences between each steroidogenic lineage remains to be investigated. In the ovary, gonad primordium-derived cells are the primary source of theca cells that eventually surround primary follicles and express WT1, HSD3B1, and ESR1 (estrogen receptor 1). Theca cells derived from mesonephros tend to localize to the basal lamina and express STAR, CYP11A1, CYP17A1, and LHCGR at higher levels than gonad-derived theca cells (Liu et al., 2015). Similarly, fetal Leydig cells in mammals have been shown to express HSD3B1 and STAR (R. S. Ge et al., 2006; Stevant et al., 2018). There is substantial overlap between the transcriptomes of differentiating steroidogenic cells in mammals. This becomes especially complicated when looking at the differentiation of steroidogenic cells in species with TSD, due to the fact that steroidogenic cells in a gonad are capable of development into both theca and Leydig cells. In addition, estrogen synthesis and signaling play a central role in ovary determination in TSD species and more broadly in non-mammalian vertebrates (Bowden & Paitz, 2018; Crews et al., 1994; Pieau & Dorizzi, 2004). This is a major distinction from placental mammals where estrogen signaling does not play a role in sex determination during embryogenesis but does play a part in the maintenance of ovarian phenotype postnatally (Britt & Findlay, 2002; Couse & Korach, 1999). Therefore, it is important to examine characteristics that distinguish steroidogenic cells from supporting cells during embryogenesis in species with TSD.

In the developing mouse ovary, granulosa cells secrete hedgehog ligands which trigger activation of GLI1 and expression of target genes involved in hedgehog signaling in thecal precursor cells (Wijgerde, Ooms, Hoogerbrugge, & Grootegoed, 2005). Hedgehog signaling is also critical for development of both fetal and adult Leydig cells in mouse testis (Clark, Garland, & Russell, 2000; Yao, Whoriskey, & Capel, 2002). Little work has been done regarding the role

of hedgehog signaling in gonadal differentiation in reptiles, but GLI1 and/or GLI2 expression is likely a conserved marker of steroidogenic cell fate (Finco, LaPensee, Krill, & Hammer, 2015).

There is significant evidence in mammals that supporting cells differentiate prior to steroidogenic lineages in both sexes and that signals from supporting cells control differentiation of steroidogenic lineages. In contrast, there is significant evidence that steroid signaling plays a larger role in controlling the fate of supporting cells in reptiles. Early work in TSD showed that steroidogenic capacity in the gonads exists early in the TSP, suggesting that gonadal cells are capable of steroid signaling prior to becoming mature theca or Leydig cells (White & Thomas, 1992).

Differentiation of Supporting Cell Lineages in the Gonad

The differentiation of unspecified supporting cells to granulosa cells in the developing ovary is controlled by complex gene expression profiles comprised of transient and permanent expression of certain genes. Though there are genes like FOXL2 that serve as markers of granulosa cell fate, there is less known about the transition from early progenitor cells to specified supporting cells to determined supporting cells in both mammals and in reptiles with TSD. Doublesex and mab3-related transcription factor 1 (DMRT1) and CYP11A1, both associated with Sertoli cell specification, are briefly over expressed in mouse at the onset of pre-granulosa cell differentiation (E11.5 and E12.5) (Lei et al., 2007; Stevant et al., 2019; Stevant et al., 2018). DMRT1 is strongly expressed during Sertoli cell differentiation in mammals. In several species with TSD, DMRT1 shows increased expression following a shift from FPT to MPT and decreased expression following a shift from MPT to FPT (Rhen et al., 2007; C. Shoemaker, Ramsey, Queen, & Crews, 2007; Smith, McClive, Western, Reed, & Sinclair, 1999; Torres Maldonado et al., 2002). Though

DMRT1 is expressed at a higher level during and after Sertoli cell differentiation, it is possible that DMRT1 also plays an earlier role in lineage restriction to a supporting cell fate because expression is observed at both female- and male-producing temperatures (Rhen et al., 2007). Leucine-rich repeat-containing G protein-coupled receptor 4 (Lgr4) also is transiently over expressed in the pre-granulosa cell lineage during differentiation. Lgr4 acts as a membrane receptor for RSPO1 and is a key component for RSPO1 potentiation of Wnt signaling. While there are many genes that are transiently over expressed during differentiation of supporting cell lineages, there are many that follow other patterns of expression (Stevant et al., 2019; Stevant et al., 2018).

A large number of genes are expressed from the time of pre-granulosa cell differentiation and are maintained after birth in mammals, including FST, STAR, RUNX1, and FOXL2 (Stevant et al., 2019). The continued expression of these genes may indicate their importance in the maintenance of cellular identity or functional role in cellular phenotype. FOXL2 exhibits temperature-dependent and sexually dimorphic expression and is strongly expressed in mature granulosa cells, but there is also evidence in mammals that cells expressing FOXL2 can give rise to granulosa cells as well as steroidogenic cells during ovarian development (Stevant et al., 2019; Uhlenhaut et al., 2009). In *Trachemys scripta* and *Chelydra serpentina*, turtles with TSD, FOXL2 expression is not dimorphic early in the TSP and becomes strongly divergent between gonads incubated at MPT versus FPT (Rhen et al., 2007; C. M. Shoemaker, Queen, & Crews, 2007). Furthermore, FOXL2 transcripts are localized to the ovarian cortex during differentiation in *Trachemys scripta*, suggesting a potential role in the differentiation of granulosa cell lineages but not excluding the potential role in the differentiation of steroidogenic cell lineages (C. M. Shoemaker et al., 2007). In addition to sexually dimorphic expression in red-eared slider turtles

and rainbow trout, RUNX1 and FOXL2 have been shown to share cellular localization and chromatin occupancy in developing mammalian ovaries, a similar phenotype with knockout of each gene, and synergistic masculinization in double knockout studies (Nicol et al., 2019). These results suggest that both proteins are essential for specification and maintenance of pre-granulosa cell fate in the developing ovary and are conserved across vertebrate lineages (Nicol et al., 2019). RUNX proteins are a highly conserved family of transcription factors that regulate gene expression in developmental context across the entire range of metazoans (Rennert, Coffman, Mushegian, & Robertson, 2003). They have been shown to act as transcriptional activators and repressors dependent on certain cellular and developmental contexts, including the recruitment of Gro/TLE and Forkhead family proteins (Ono et al., 2007; Wildey & Howe, 2009; Yarmus et al., 2006). In contrast with pre-granulosa cell fate, RUNX1 is repressed during Sertoli cell fate commitment (Munger, Natarajan, Looger, Ohler, & Capel, 2013; Stevant et al., 2018).

Other markers of granulosa cell fate include proteins associated with steroidogenesis, HSD3B1 and CYP19A1, which have both been shown to be expressed at FPT in turtles and in differentiated granulosa cells in mammals (Bakhshalizadeh et al., 2017; Estermann et al., 2020). CYP19A1 encodes the aromatase enzyme, which converts androgens to estrogens and has long been of interest in TSD. Expression of CYP19A1 increases in embryos shifted from MPT to FPT in the snapping turtle (Rhen et al., 2007) and inhibition of aromatase activity blocks ovarian development at a temperature that normally produces a very female-biased sex ratio (Rhen and Lang, 1994). Evidence from other TSD reptiles also indicate that gonadal aromatase expression is likely responsible for differential estrogen activity during development of gonads at FPT (Pieau & Dorizzi, 2004; Ramsey & Crews, 2009).

The differentiation of Sertoli cells from gonadal progenitors is similar to the differentiation of granulosa cells and involves a common intermediate specification step that restricts cells to a supporting cell lineage in mammals. There are parallels and differences in gonadal differentiation between mammals and reptiles with TSD, but it is likely that a similar intermediate lineage-restriction step is required in TSD. In mammals, this step is characterized by up-regulation of about 200 genes, with only SRY and SOX9 exhibiting sexual dimorphism (Stevant et al., 2019). Notably, RUNX1, CYP11A1, and DMRT1 fall into this category in mammals. Several genes up-regulated in this intermediate differentiation step remain up-regulated in Sertoli cells following differentiation.

The transcriptomes of Sertoli cells during and after differentiation share similarities with early progenitors and steroidogenic lineages with the exception of a set of key genes that are suspected of driving a male fate in mammals. These genes include NR0B1 (DAX1), DMRT1, GATA4, SOX9, and AMH (Matson et al., 2011; Minkina et al., 2014; Stevant et al., 2018). Interestingly, loss of DMRT1 in Sertoli cells increases expression of CYP19A1 and serum levels of estradiol and knockout of SOX9 and DMRT1 resulted in transdifferentiation of Sertoli cells into granulosa-like cells (Matson et al., 2011; Minkina et al., 2014). Minkina et al. (2014) also found that increased retinoic acid signaling produces synergistic effect with depletion of DMRT1 in Sertoli cells and that decreased retinoic acid signaling, particularly the loss of RARA, suppresses transdifferentiation, suggesting that the feminizing effects of DMRT1 depletion are dependent on retinoic acid signaling (Minkina et al., 2014). Interestingly, they also showed that RARA acts upstream of feminizing genes FOXL2, LRH1, ESR2, and WNT4 in the absence of DMRT1. Retinoic acid signaling is a master cell fate regulator during embryogenesis and is required for spermatogonial differentiation and for oocyte entry into meiosis during mouse embryogenesis

(Bowles & Koopman, 2007; Raverdeau et al., 2012). In the developing testis, retinoic acid (RA) is degraded by CYP26B1 and failure of RA degradation results in a feminizing effect on the testis (Bowles et al., 2018). The dependency of Sertoli cell transdifferentiation on retinoic acid signaling and DMRT1-depletion could give insight into the regulatory mechanisms governing the expression of feminizing genes in TSD. In Sertoli cells, DMRT1 directly represses transcription of feminizing genes (FOXL2, ESR2, WNT4, RSPO1, CYP19A1), which are also regulated by retinoic acid signaling. Minkina et al. (2014) suggested a model in which DMRT1 could act to repress RA-dependent transcription of feminizing genes while allowing Sertoli cells to produce the RA required for spermatogenesis.

Retinoic acid functions as a ligand for retinoic acid receptors that bind to retinoic acid response elements (RAREs) in the promoter regions of genes that are responsive to RA signaling. Interestingly, in the absence of RA, RAREs tend to be bound by polycomb proteins, specifically SUZ12, a component of Polycomb Repressive Complex 2, which deposits the repressive H3K27me3 mark (Gillespie & Gudas, 2007a, 2007b). Upon RA signaling, polycomb proteins dissociate from chromatin and transcription of genes controlled by RARA can be initiated. This model of regulation of feminizing genes, could provide insight to the role of JARID2 in TSD.

JARID2 as a Candidate Regulator of Differentiation During TSD

We have identified a set of genes that respond quickly to a shift in temperature from MPT to FPT during the thermosensitive period in the snapping turtle. Among these are a number of transcription factors, chromatin remodelers, and players in key signaling pathways, all of which may serve roles in commitment to ovarian or testicular fate. One gene that is strongly induced

within 24 hours of a thermal shift (MPT to FPT) is the Jumonji family, ARID-domain containing protein JARID2. JARID2 has been shown to associate with Polycomb Repressive Complex 2 and regulate cell fate determination. JARID2 and Polycomb Repressive Complex 2 play a role in the deposition and maintenance of H3K27me3 repressive chromatin domains, which have been shown in several cases to play a role in the temperature sensitive regulation of genes involved in TSD (C. Ge et al., 2018; Matsumoto, Hannigan, & Crews, 2016).

In *T. scripta*, H3K27me3 levels are elevated at the *Dmrt1* promoter at FPT which correlates with low occupancy of the lysine-specific demethylase KDM6B. Additionally, siRNA interference of KDM6B resulted in decreased KDM6B occupancy at the *Dmrt1* promoter and increased levels of H3K27me3, which resulted in decreased expression of DMRT1 (C. Ge et al., 2018). Furthermore, overexpression of DMRT1 was sufficient to restore the male pathway in the absence of KDM6B. The results from this study suggest an integral role for H3K27me3 in the regulation of genes involved in TSD and KDM6B plays a role in the regulation of H3K27me3. The sharp increase in JARID2 expression in response to a thermal shift and temperature-sensitive H3K27me3 levels together give reason to further investigate the role of JARID2 in the regulation of gene expression during temperature-dependent sex determination.

Polycomb group proteins (PcG proteins) are important transcriptional regulators during development and embryonic stem cell differentiation (Margueron & Reinberg, 2011). They are remarkably well-conserved throughout eukaryotes, from single-celled eukaryotes to vertebrates (Shaver, Casas-Mollano, Cerny, & Cerutti, 2010; Whitcomb, Basu, Allis, & Bernstein, 2007). The first polycomb group (PcG) gene was initially discovered as a repressor of Hox genes in *Drosophila* (Slifer, 1942). Researchers have since discovered that PcG protein evolution, even amongst vertebrates, is quite complex. Many vertebrate lineages have diverged to have

functionally distinct copies of certain genes, such as Enhancer of Zeste 1 and 2, *Ezh1* and *Ezh2* (Margueron et al., 2008). In some plants, there have been as many as 12 duplication events for particular PcG proteins (Hennig & Derkacheva, 2009). It has been proposed that the diversification of PcG proteins has led to the evolution of developmental patterns in gene expression that ultimately lead to differentiation and tissue specification (Whitcomb et al., 2007). PcG complexes vary in content and complexity across organisms, creating challenges in understanding their function and importance. However, several core proteins have been identified that are highly conserved among multicellular eukaryotes.

PcG proteins catalyze the post-translational modification of histone tails which results in the formation of heterochromatic domains that may repress expression of nearby genes. These chromatin modifications are carried out by two major polycomb group protein complexes, Polycomb Repressive Complex 1 (PRC1) and Polycomb Repressive Complex 2 (PRC2). PRC1 catalyzes mono-ubiquitination of histone H2A (Endoh et al., 2012). Methylation of histone H3 on lysine 27 (K27) is performed chiefly by PRC2 (Cao et al., 2002; Kuzmichev, Nishioka, Erdjument-Bromage, Tempst, & Reinberg, 2002). PRC2 consists of four main subunits: SUZ12, EZH2, EED, and RBBP4/RBBP7. EED is responsible for binding di- and tri-methylated H3K27 and allosterically activating EZH2, which is responsible for the histone methyltransferase activity of PRC2. Together, EED and EZH2 propagate H3K27 methylation (Margueron et al., 2009; Margueron et al., 2008; Oksuz et al., 2018). None of the core subunits of PRC2 contain a DNA-binding domain, yet PRC2 has been shown to localize at specific sites and repress nearby genes, with target loci varying temporally and between cell types (Margueron & Reinberg, 2011). The coordination of PRC2 activity is essential to differentiation, but the exact mechanism of this coordination is largely unknown in higher order vertebrates (Margueron & Reinberg, 2011; Oksuz

et al., 2018). It is currently proposed that recruitment of PRC2 occurs via association with accessory proteins and long non-coding RNAs (H. Kim, Kang, & Kim, 2009; Margueron & Reinberg, 2011; Rinn et al., 2007).

PRC2 has been shown to associate with various proteins in transient interactions that play varying roles in chromatin remodeling. Among these proteins are JARID2, AEBP2, ATRX, MTF2 (Grijzenhout et al., 2016; H. Kim et al., 2009; H. Li et al., 2017; Margueron et al., 2008; Peng et al., 2009; Sarma et al., 2014). PRC2 is capable of enzymatic activity without associating with an accessory protein, but its specific recruitment and enzymatic activity may be enhanced when the complex is bound to an accessory protein (Oksuz et al., 2018). JARID2, a Jumonji family ARID-domain containing protein void of demethylase activity, has been shown to play a role in recruitment and modulation of PRC2 enzymatic activity in a fashion that leads to the down-regulation of certain genes involved in differentiation (Adhikari & Davie, 2018; Adhikari, Mainali, & Davie, 2019; da Rocha et al., 2014; Kaneko et al., 2014; G. Li et al., 2010; Oksuz et al., 2018; Sanulli et al., 2015). In the absence of JARID2, PRC2 fails to localize to some of its canonical targets and propagate H3K27me2/3, which can result in misregulation of genes involved in differentiation (Adhikari & Davie, 2018; G. Li et al., 2010; Oksuz et al., 2018). This suggests that JARID2-PRC2 is essential to the programming of cell-fate decisions during embryogenesis.

JARID2, interacts with PRC2 and facilitates the initial interaction between PRC2 and chromatin required for *de novo* deposition of H3K27 methylation at CpG islands (Chen, Jiao, Shubbar, Yang, & Liu, 2018; Hojfeldt et al., 2018; Oksuz et al., 2018). JARID2-PRC2 complexes localize along the chromatin, forming foci of polycomb activity that result in the propagation of H3K27 methylation. Following initial deposition of H3K27me3, PRC2 is capable of self-propagation and spreading this epigenetic modification along the chromosome (Margueron et al.,

2009; Oksuz et al., 2018). While sites of JARID2-mediated PRC2 recruitment remain stable in certain vertebrate cell types and models, they lack the predictable nature of the Polycomb Response Elements (PREs) observed in Drosophila (Bauer, Trupke, & Ringrose, 2016; J. Muller & Kassis, 2006). Several putative mammalian PREs have been identified but their capacities to recruit PRC2 seem to be lineage specific and the effects of PRC2 spreading versus *de novo* deposition are difficult to sort out. Advances in single-cell technology will likely make this task easier (Mohn et al., 2008; Woo, Kharchenko, Daheron, Park, & Kingston, 2010). While the exact loci that PRC2 is recruited to varies between model systems, the reliance on accessory proteins, such as JARID2, remains constant.

One potential link between TSD and PRC2-mediated transcriptional repression is the occupancy of PRC2 at genes that may be stimulated by retinoic acid signaling. As previously discussed, these genes include FOXL2, ESR2, CYP19A1, WNT4, and RSPO1 in mammals (Minkina et al., 2014). It has been shown in mammalian cell lines that PRC2 associates with retinoic acid response elements, and that retinoic acid causes PRC2 to dissociate from chromatin. Additionally, a decrease in PRC2 translates to a decrease in H3K27me3 upon addition of retinoic acid (Gillespie & Gudas, 2007b). The idea that CYP19A1 and FOXL2 expression are regulated by the interplay between PRC2 and retinoic acid signaling is an attractive hypothesis in relation to TSD research.

If genes involved with sex determination are regulated by the intersection of DMRT1, retinoic acid signaling, and PRC2-mediated transcriptional repression, one would expect to observe differential H3K27me3 levels in the promoters of these genes. In *T. scripta*, a turtle with TSD, elevated H3K27me3 levels were observed at the CYP19A1 promoter in a shift from MPT to FPT and DMRT1 knockdown resulted in increased expression of FOXL2 and CYP19A1, as well

as sex-reversed phenotypes at each temperature (C. Ge et al., 2017; Matsumoto et al., 2016). A sharp increase in JARID2 expression due to a shift from MPT to FPT could correlate with increased PRC2 occupancy and H3K27me3 levels at genes regulated by retinoic acid signaling and DMRT1, which would result in decreased transcription.

JARID2 has been shown to display temperature-sensitive alternate splicing resulting in the retention of introns in JARID2 transcripts in *Pogona vitticeps*, *Alligator missipiensis*, and *Trachemys scripta* (Deveson et al., 2017). Intron retention may alter the function of JARID2 or its interaction with PRC2. It is also possible that temperature influences expression of different isoforms of JARID2 and that these isoforms may cause differential expression of other genes in the snapping turtle.

Conclusion

TSD is an example of phenotypic plasticity in which many complex regulatory mechanisms control the fates of cells that eventually differentiate and contribute to the development of an ovary or testis. Many of these decisions rely on complex patterns of gene expression which may be partially established and maintained by epigenetic modifications. One such modification, H3K27me3, is likely to be responsible for the regulation of several genes involved in sex determination. JARID2, which is also implicated in TSD, plays a role in the establishment of H3K27me3 in a highly coordinated fashion and could potentially be a mechanism of establishing temperature-dependent patterns of gene expression required for gonad differentiation and sex determination.

This chapter has explored the roles of temperature and JARID2 in temperature-dependent sex determination. We have summarized the findings of studies that investigate the relationships between the regulation of gene expression and the differentiation of gonadal cell types that underlie

the development of ovaries and testes. We have discussed the epigenetic regulation of gene expression by polycomb-mediated histone modification and its implications in genotypic and temperature-dependent sex determination. We propose a model by which JARID2 isoforms may differ in function and may regulate the expression of genes involved in sex determination by facilitating recruitment of PRC2 to the promoters of those genes.

References

Adhikari, A., & Davie, J. (2018). JARID2 and the PRC2 complex regulate skeletal muscle differentiation through regulation of canonical Wnt signaling. *Epigenetics Chromatin, 11*(1), 46. doi:10.1186/s13072-018-0217-x

Adhikari, A., Mainali, P., & Davie, J. K. (2019). JARID2 and the PRC2 complex regulates the cell cycle in skeletal muscle. *J Biol Chem*. doi:10.1074/jbc.RA119.010060

Apostolou, E., Ferrari, F., Walsh, R. M., Bar-Nur, O., Stadtfeld, M., Cheloufi, S., . . . Hochedlinger, K. (2013). Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell, 12*(6), 699-712. doi:10.1016/j.stem.2013.04.013

Bakhshalizadeh, S., Amidi, F., Alleyassin, A., Soleimani, M., Shirazi, R., & Shabani Nashtaei, M. (2017). Modulation of steroidogenesis by vitamin D3 in granulosa cells of the mouse model of polycystic ovarian syndrome. *Syst Biol Reprod Med, 63*(3), 150-161. doi:10.1080/19396368.2017.1296046

Bauer, M., Trupke, J., & Ringrose, L. (2016). The quest for mammalian Polycomb response elements: are we there yet? *Chromosoma, 125*(3), 471-496. doi:10.1007/s00412-015-0539-4

Bernstein, B. E., Meissner, A., & Lander, E. S. (2007). The mammalian epigenome. *Cell, 128*(4), 669-681. doi:10.1016/j.cell.2007.01.033

Bowden, R. M., & Paitz, R. T. (2018). Temperature fluctuations and maternal estrogens as critical factors for understanding temperature-dependent sex determination in nature. *J Exp Zool A Ecol Integr Physiol, 329*(4-5), 177-184. doi:10.1002/jez.2183

Bowles, J., Feng, C. W., Ineson, J., Miles, K., Spiller, C. M., Harley, V. R., . . . Koopman, P. (2018). Retinoic Acid Antagonizes Testis Development in Mice. *Cell Rep, 24*(5), 1330-1341. doi:10.1016/j.celrep.2018.06.111

Bowles, J., & Koopman, P. (2007). Retinoic acid, meiosis and germ cell fate in mammals. *Development, 134*(19), 3401-3411. doi:10.1242/dev.001107

Britt, K. L., & Findlay, J. K. (2002). Estrogen actions in the ovary revisited. *J Endocrinol, 175*(2), 269-276. doi:10.1677/joe.0.1750269

Bull, J. J., Vogt, R. C., & Bulmer, M. G. (1982). Heritability of Sex Ratio in Turtles with Environmental Sex Determination. *Evolution, 36*(2), 333-341. doi:10.1111/j.1558-5646.1982.tb05049.x

Burton, A., Muller, J., Tu, S., Padilla-Longoria, P., Guccione, E., & Torres-Padilla, M. E. (2013). Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Rep, 5*(3), 687-701. doi:10.1016/j.celrep.2013.09.044

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., . . . Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science, 298*(5595), 1039-1043. doi:10.1126/science.1076997

Chen, S., Jiao, L., Shubbar, M., Yang, X., & Liu, X. (2018). Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for Chromatin Binding. *Mol Cell, 69*(5), 840-852 e845. doi:10.1016/j.molcel.2018.01.039

Clark, A. M., Garland, K. K., & Russell, L. D. (2000). Desert hedgehog (Dhh) gene is required in the mouse testis for formation of adult-type Leydig cells and normal development of peritubular cells and seminiferous tubules. *Biol Reprod, 63*(6), 1825-1838. doi:10.1095/biolreprod63.6.1825

Couse, J. F., & Korach, K. S. (1999). Estrogen receptor null mice: what have we learned and where will they lead us? *Endocr Rev, 20*(3), 358-417. doi:10.1210/edrv.20.3.0370

Crews, D., Bergeron, J. M., Bull, J. J., Flores, D., Tousignant, A., Skipper, J. K., & Wibbels, T. (1994). Temperature-dependent sex determination in reptiles: proximate mechanisms, ultimate outcomes, and practical applications. *Dev Genet, 15*(3), 297-312. doi:10.1002/dvg.1020150310

da Rocha, S. T., Boeva, V., Escamilla-Del-Arenal, M., Ancelin, K., Granier, C., Matias, N. R., . . . Heard, E. (2014). Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. *Mol Cell, 53*(2), 301-316. doi:10.1016/j.molcel.2014.01.002

DeFalco, T., & Capel, B. (2009). Gonad morphogenesis in vertebrates: divergent means to a convergent end. *Annu Rev Cell Dev Biol, 25*, 457-482. doi:10.1146/annurev.cellbio.042308.13350

Deveson, I. W., Holleley, C. E., Blackburn, J., Marshall Graves, J. A., Mattick, J. S., Waters, P. D., & Georges, A. (2017). Differential intron retention in Jumonji chromatin modifier genes is implicated in reptile temperature-dependent sex determination. *Sci Adv, 3*(6), e1700731. doi:10.1126/sciadv.1700731

Endoh, M., Endo, T. A., Endoh, T., Isono, K., Sharif, J., Ohara, O., . . . Koseki, H. (2012). Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity. *PLoS Genet, 8*(7), e1002774. doi:10.1371/journal.pgen.1002774

Estermann, M. A., Williams, S., Hirst, C. E., Roly, Z. Y., Serralbo, O., Adhikari, D., . . . Smith, C. A. (2020). Insights into Gonadal Sex Differentiation Provided by Single-Cell Transcriptomics in the Chicken Embryo. *Cell Rep, 31*(1), 107491. doi:10.1016/j.celrep.2020.03.055

Ewert, A. M., Jackson, D. R., & Nelson, C. E. (1994). Patterns of Temperature-Dependent Sex Determination in Turtles. *The Journal Of Experimental Zoology*(270), 3-15.

Finco, I., LaPensee, C. R., Krill, K. T., & Hammer, G. D. (2015). Hedgehog signaling and steroidogenesis. *Annu Rev Physiol, 77*, 105-129. doi:10.1146/annurev-physiol-061214-111754

Fleming, A., Wibbels, T., Skipper, J. K., & Crews, D. (1999). Developmental expression of steroidogenic factor 1 in a turtle with temperature-dependent sex determination. *Gen Comp Endocrinol, 116*(3), 336-346. doi:10.1006/gcen.1999.7360

Ge, C., Ye, J., Weber, C., Sun, W., Zhang, H., Zhou, Y., . . . Capel, B. (2018). The histone demethylase KDM6B regulates temperature-dependent sex determination in a turtle species. *Science, 360*(6389), 645-648. doi:10.1126/science.aap8328

Ge, C., Ye, J., Zhang, H., Zhang, Y., Sun, W., Sang, Y., . . . Qian, G. (2017). Dmrt1 induces the male pathway in a turtle species with temperature-dependent sex determination. *Development, 144*(12), 2222-2233. doi:10.1242/dev.152033

Ge, R. S., Dong, Q., Sottas, C. M., Papadopoulos, V., Zirkin, B. R., & Hardy, M. P. (2006). In search of rat stem Leydig cells: identification, isolation, and lineage-specific development. *Proc Natl Acad Sci U S A, 103*(8), 2719-2724. doi:10.1073/pnas.0507692103

Gillespie, R. F., & Gudas, L. J. (2007a). Retinoic acid receptor isotype specificity in F9 teratocarcinoma stem cells results from the differential recruitment of coregulators to retinoic response elements. *J Biol Chem, 282*(46), 33421-33434. doi:10.1074/jbc.M704845200

Gillespie, R. F., & Gudas, L. J. (2007b). Retinoid regulated association of transcriptional co-regulators and the polycomb group protein SUZ12 with the retinoic acid response elements of Hoxa1, RARbeta(2), and Cyp26A1 in F9 embryonal carcinoma cells. *J Mol Biol, 372*(2), 298-316. doi:10.1016/j.jmb.2007.06.079

Grijzenhout, A., Godwin, J., Koseki, H., Gdula, M. R., Szumska, D., McGouran, J. F., . . . Cooper, S. (2016). Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. *Development, 143*(15), 2716-2723. doi:10.1242/dev.123935

Hennig, L., & Derkacheva, M. (2009). Diversity of Polycomb group complexes in plants: same rules, different players? *Trends Genet, 25*(9), 414-423. doi:10.1016/j.tig.2009.07.002

Hojfeldt, J. W., Laugesen, A., Willumsen, B. M., Damhofer, H., Hedehus, L., Tvardovskiy, A., . . . Helin, K. (2018). Accurate H3K27 methylation can be established de novo by SUZ12-directed PRC2. *Nat Struct Mol Biol, 25*(3), 225-232. doi:10.1038/s41594-018-0036-6

Jin, W., Riley, R. M., Wolfinger, R. D., White, K. P., Passador-Gurgel, G., & Gibson, G. (2001). The contributions of sex, genotype and age to transcriptional variance in Drosophila melanogaster. *Nat Genet, 29*(4), 389-395. doi:10.1038/ng766

Kaneko, S., Bonasio, R., Saldana-Meyer, R., Yoshida, T., Son, J., Nishino, K., . . . Reinberg, D. (2014). Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Mol Cell, 53*(2), 290-300. doi:10.1016/j.molcel.2013.11.012

Karl, J., & Capel, B. (1998). Sertoli cells of the mouse testis originate from the coelomic epithelium. *Dev Biol, 203*(2), 323-333. doi:10.1006/dbio.1998.9068

Kim, H., Kang, K., & Kim, J. (2009). AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. *Nucleic Acids Res, 37*(9), 2940-2950. doi:10.1093/nar/gkp149

Kim, Y., Kobayashi, A., Sekido, R., DiNapoli, L., Brennan, J., Chaboissier, M. C., . . . Capel, B. (2006). Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination. *PLoS Biol, 4*(6), e187. doi:10.1371/journal.pbio.0040187

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., & Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev, 16*(22), 2893-2905. doi:10.1101/gad.1035902

Lei, N., Hornbaker, K. I., Rice, D. A., Karpova, T., Agbor, V. A., & Heckert, L. L. (2007). Sex-specific differences in mouse DMRT1 expression are both cell type- and stage-dependent during gonad development. *Biol Reprod, 77*(3), 466-475. doi:10.1095/biolreprod.106.058784

Li, G., Margueron, R., Ku, M., Chambon, P., Bernstein, B. E., & Reinberg, D. (2010). Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev, 24*(4), 368-380. doi:10.1101/gad.1886410

Li, H., Liefke, R., Jiang, J., Kurland, J. V., Tian, W., Deng, P., . . . Wang, Z. (2017). Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature, 549*(7671), 287-291. doi:10.1038/nature23881

Lin, Y. T., & Capel, B. (2015). Cell fate commitment during mammalian sex determination. *Curr Opin Genet Dev, 32*, 144-152. doi:10.1016/j.gde.2015.03.003

Liu, C., Peng, J., Matzuk, M. M., & Yao, H. H. (2015). Lineage specification of ovarian theca cells requires multicellular interactions via oocyte and granulosa cells. *Nat Commun, 6*, 6934. doi:10.1038/ncomms7934

Luo, X., Ikeda, Y., & Parker, K. L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell, 77*(4), 481-490. doi:10.1016/0092-8674(94)90211-9

Lynch, M., & Force, A. G. (2000). The Origin of Interspecific Genomic Incompatibility via Gene Duplication. *Am Nat, 156*(6), 590-605. doi:10.1086/316992

Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., 3rd, . . . Gamblin, S. J. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature, 461*(7265), 762-767. doi:10.1038/nature08398

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C. L., . . . Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell, 32*(4), 503-518. doi:10.1016/j.molcel.2008.11.004

Margueron, R., & Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature, 469*(7330), 343-349. doi:10.1038/nature09784

Marin, I., & Baker, B. S. (1998). The evolutionary dynamics of sex determination. *Science, 281*(5385), 1990-1994. doi:10.1126/science.281.5385.1990

Matson, C. K., Murphy, M. W., Sarver, A. L., Griswold, M. D., Bardwell, V. J., & Zarkower, D. (2011). DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature, 476*(7358), 101-104. doi:10.1038/nature10239

Matsumoto, Y., Hannigan, B., & Crews, D. (2016). Temperature Shift Alters DNA Methylation and Histone Modification Patterns in Gonadal Aromatase (cyp19a1) Gene in Species with Temperature-Dependent Sex Determination. *PLoS One, 11*(11), e0167362. doi:10.1371/journal.pone.0167362

Minkina, A., Matson, C. K., Lindeman, R. E., Ghyselinck, N. B., Bardwell, V. J., & Zarkower, D. (2014). DMRT1 protects male gonadal cells from retinoid-dependent sexual transdifferentiation. *Dev Cell, 29*(5), 511-520. doi:10.1016/j.devcel.2014.04.017

Mohn, F., Weber, M., Rebhan, M., Roloff, T. C., Richter, J., Stadler, M. B., . . . Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell, 30*(6), 755-766. doi:10.1016/j.molcel.2008.05.007

Muller. (1942). Isolating mechanisms, evolution, and temperature. *Biol. Symp., 6*, 71-125.

Muller, J., & Kassis, J. A. (2006). Polycomb response elements and targeting of Polycomb group proteins in Drosophila. *Curr Opin Genet Dev, 16*(5), 476-484. doi:10.1016/j.gde.2006.08.005

Munger, S. C., Natarajan, A., Looger, L. L., Ohler, U., & Capel, B. (2013). Fine time course expression analysis identifies cascades of activation and repression and maps a putative regulator of mammalian sex determination. *PLoS Genet, 9*(7), e1003630. doi:10.1371/journal.pgen.1003630

Nicol, B., Grimm, S. A., Chalmel, F., Lecluze, E., Pannetier, M., Pailhoux, E., . . . Yao, H. H. (2019). RUNX1 maintains the identity of the fetal ovary through an interplay with FOXL2. *Nat Commun, 10*(1), 5116. doi:10.1038/s41467-019-13060-1

Oksuz, O., Narendra, V., Lee, C. H., Descostes, N., LeRoy, G., Raviram, R., . . . Reinberg, D. (2018). Capturing the Onset of PRC2-Mediated Repressive Domain Formation. *Mol Cell, 70*(6), 1149-1162 e1145. doi:10.1016/j.molcel.2018.05.023

Ono, M., Yaguchi, H., Ohkura, N., Kitabayashi, I., Nagamura, Y., Nomura, T., . . . Sakaguchi, S. (2007). Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature, 446*(7136), 685-689. doi:10.1038/nature05673

Orr, H. A. (1996). Dobzhansky, Bateson, and the genetics of speciation. *Genetics, 144*(4), 1331-1335. Retrieved from **https://www.ncbi.nlm.nih.gov/pubmed/8978022**

Peng, J. C., Valouev, A., Swigut, T., Zhang, J., Zhao, Y., Sidow, A., & Wysocka, J. (2009). Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell, 139*(7), 1290-1302. doi:10.1016/j.cell.2009.12.002

Pieau, C., & Dorizzi, M. (2004). Oestrogens and temperature-dependent sex determination in reptiles: all is in the gonads. *J Endocrinol, 181*(3), 367-377. doi:10.1677/joe.0.1810367

Pitetti, J. L., Calvel, P., Romero, Y., Conne, B., Truong, V., Papaioannou, M. D., . . . Nef, S. (2013). Insulin and IGF1 receptors are essential for XX and XY gonadal differentiation and adrenal development in mice. *PLoS Genet, 9*(1), e1003160. doi:10.1371/journal.pgen.1003160

Ramsey, M., & Crews, D. (2009). Steroid signaling and temperature-dependent sex determination-Reviewing the evidence for early action of estrogen during ovarian determination in turtles. *Semin Cell Dev Biol, 20*(3), 283-292. doi:10.1016/j.semcdb.2008.10.004

Ranz, J. M., Castillo-Davis, C. I., Meiklejohn, C. D., & Hartl, D. L. (2003). Sex-dependent gene expression and evolution of the Drosophila transcriptome. *Science, 300*(5626), 1742-1745. doi:10.1126/science.1085881

Raverdeau, M., Gely-Pernot, A., Feret, B., Dennefeld, C., Benoit, G., Davidson, I., . . . Ghyselinck, N. B. (2012). Retinoic acid induces Sertoli cell paracrine signals for spermatogonia differentiation but cell autonomously drives spermatocyte meiosis. *Proc Natl Acad Sci U S A, 109*(41), 16582-16587. doi:10.1073/pnas.1214936109

Rennert, J., Coffman, J. A., Mushegian, A. R., & Robertson, A. J. (2003). The evolution of Runx genes I. A comparative study of sequences from phylogenetically diverse model organisms. *BMC Evol Biol, 3*, 4. doi:10.1186/1471-2148-3-4

Rhen, T., Fagerlie, R., Schroeder, A., Crossley, D. A., 2nd, & Lang, J. W. (2015). Molecular and morphological differentiation of testes and ovaries in relation to the thermosensitive period of gonad development in the snapping turtle, Chelydra serpentina. *Differentiation, 89*(1-2), 31-41. doi:10.1016/j.diff.2014.12.007

Rhen, T., & Lang, J. W. (1994). Temperature-dependent sex determination in the snapping turtle: manipulation of the embryonic sex steroid environment. *Gen Comp Endocrinol, 96*(2), 243-254. doi:10.1006/gcen.1994.1179

Rhen, T., & Lang, J. W. (1998). Among-Family Variation for Environmental Sex Determination in Reptiles. *Evolution, 52*(5), 1514-1520. doi:10.1111/j.1558-5646.1998.tb02034.x

Rhen, T., Metzger, K., Schroeder, A., & Woodward, R. (2007). Expression of putative sex-determining genes during the thermosensitive period of gonad development in the snapping turtle, Chelydra serpentina. *Sex Dev, 1*(4), 255-270. doi:10.1159/000104775

Rhen, T., & Schroeder, A. (2010). Molecular mechanisms of sex determination in reptiles. *Sex Dev, 4*(1-2), 16-28. doi:10.1159/000282495

Rhen, T., Schroeder, A., Sakata, J. T., Huang, V., & Crews, D. (2011). Segregating variation for temperature-dependent sex determination in a lizard. *Heredity (Edinb), 106*(4), 649-660. doi:10.1038/hdy.2010.102

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., . . . Chang, H. Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell, 129*(7), 1311-1323. doi:10.1016/j.cell.2007.05.022

Sanulli, S., Justin, N., Teissandier, A., Ancelin, K., Portoso, M., Caron, M., . . . Margueron, R. (2015). Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation. *Mol Cell, 57*(5), 769-783. doi:10.1016/j.molcel.2014.12.020

Sarma, K., Cifuentes-Rojas, C., Ergun, A., Del Rosario, A., Jeon, Y., White, F., . . . Lee, J. T. (2014). ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. *Cell, 159*(4), 869-883. doi:10.1016/j.cell.2014.10.019

Shaver, S., Casas-Mollano, J. A., Cerny, R. L., & Cerutti, H. (2010). Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga Chlamydomonas. *Epigenetics, 5*(4), 301-312. doi:10.4161/epi.5.4.11608

Shoemaker, C., Ramsey, M., Queen, J., & Crews, D. (2007). Expression of Sox9, Mis, and Dmrt1 in the gonad of a species with temperature-dependent sex determination. *Dev Dyn, 236*(4), 1055-1063. doi:10.1002/dvdy.21096

Shoemaker, C. M., & Crews, D. (2009). Analyzing the coordinated gene network underlying temperature-dependent sex determination in reptiles. *Semin Cell Dev Biol, 20*(3), 293-303. doi:10.1016/j.semcdb.2008.10.010

Shoemaker, C. M., Queen, J., & Crews, D. (2007). Response of candidate sex-determining genes to changes in temperature reveals their involvement in the molecular network underlying temperature-dependent sex determination. *Mol Endocrinol, 21*(11), 2750-2763. doi:10.1210/me.2007-0263

Slifer, E. H. (1942). A mutant stock of Drosophila with extra sex combs. *The Journal Of Experimental Zoology, 90*, 31-40.

Smith, C. A., McClive, P. J., Western, P. S., Reed, K. J., & Sinclair, A. H. (1999). Conservation of a sex-determining gene. *Nature, 402*(6762), 601-602. doi:10.1038/45130

Stevant, I., Kuhne, F., Greenfield, A., Chaboissier, M. C., Dermitzakis, E. T., & Nef, S. (2019). Dissecting Cell Lineage Specification and Sex Fate Determination in Gonadal Somatic Cells Using Single-Cell Transcriptomics. *Cell Rep, 26*(12), 3272-3283 e3273. doi:10.1016/j.celrep.2019.02.069

Stevant, I., Neirijnck, Y., Borel, C., Escoffier, J., Smith, L. B., Antonarakis, S. E., . . . Nef, S. (2018). Deciphering Cell Lineage Specification during Male Sex Determination with Single-Cell RNA Sequencing. *Cell Rep, 22*(6), 1589-1599. doi:10.1016/j.celrep.2018.01.043

Szutorisz, H., Canzonetta, C., Georgiou, A., Chow, C. M., Tora, L., & Dillon, N. (2005). Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. *Mol Cell Biol, 25*(5), 1804-1820. doi:10.1128/MCB.25.5.1804-1820.2005

Tevosian, S. G., & Manuylov, N. L. (2008). To beta or not to beta: canonical beta-catenin signaling pathway and ovarian development. *Dev Dyn, 237*(12), 3672-3680. doi:10.1002/dvdy.21784

Torres Maldonado, L. C., Landa Piedra, A., Moreno Mendoza, N., Marmolejo Valencia, A., Meza Martinez, A., & Merchant Larios, H. (2002). Expression profiles of Dax1, Dmrt1, and Sox9 during temperature sex determination in gonads of the sea turtle Lepidochelys olivacea. *Gen Comp Endocrinol, 129*(1), 20-26. doi:10.1016/s0016-6480(02)00511-7

Torres-Padilla, M. E., Parfitt, D. E., Kouzarides, T., & Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature, 445*(7124), 214-218. doi:10.1038/nature05458

Uhlenhaut, N. H., Jakob, S., Anlag, K., Eisenberger, T., Sekido, R., Kress, J., . . . Treier, M. (2009). Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. *Cell, 139*(6), 1130-1142. doi:10.1016/j.cell.2009.11.021

Viets, B. E. E., M.A.. Talent, L.G.. Craig E Nelson, C.E.. (1994). Sex-Determining Mechanisms in Squamate Reptiles. *The Journal Of Experimental Zoology*(270), 45-56.

Western, P. S., Harry, J. L., Marshall Graves, J. A., & Sinclair, A. H. (2000). Temperature-dependent sex determination in the American alligator: expression of SF1, WT1 and DAX1 during gonadogenesis. *Gene, 241*(2), 223-232. doi:10.1016/s0378-1119(99)00466-7

Whitcomb, S. J., Basu, A., Allis, C. D., & Bernstein, E. (2007). Polycomb Group proteins: an evolutionary perspective. *Trends Genet, 23*(10), 494-502. doi:10.1016/j.tig.2007.08.006

White, R. B., & Thomas, P. (1992). Adrenal-kidney and gonadal steroidogenesis during sexual differentiation of a reptile with temperature-dependent sex determination. *Gen Comp Endocrinol, 88*(1), 10-19. doi:10.1016/0016-6480(92)90189-q

Wijgerde, M., Ooms, M., Hoogerbrugge, J. W., & Grootegoed, J. A. (2005). Hedgehog signaling in mouse ovary: Indian hedgehog and desert hedgehog from granulosa cells induce target gene expression in developing theca cells. *Endocrinology, 146*(8), 3558-3566. doi:10.1210/en.2005-0311

Wildey, G. M., & Howe, P. H. (2009). Runx1 is a co-activator with FOXO3 to mediate transforming growth factor beta (TGFbeta)-induced Bim transcription in hepatic cells. *J Biol Chem, 284*(30), 20227-20239. doi:10.1074/jbc.M109.027201

Woo, C. J., Kharchenko, P. V., Daheron, L., Park, P. J., & Kingston, R. E. (2010). A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell, 140*(1), 99-110. doi:10.1016/j.cell.2009.12.022

Xu, C. R., Cole, P. A., Meyers, D. J., Kormish, J., Dent, S., & Zaret, K. S. (2011). Chromatin "prepattern" and histone modifiers in a fate choice for liver and pancreas. *Science, 332*(6032), 963-966. doi:10.1126/science.1202845

Yao, H. H., Whoriskey, W., & Capel, B. (2002). Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. *Genes Dev, 16*(11), 1433-1440. doi:10.1101/gad.981202

Yarmus, M., Woolf, E., Bernstein, Y., Fainaru, O., Negreanu, V., Levanon, D., & Groner, Y. (2006). Groucho/transducin-like Enhancer-of-split (TLE)-dependent and -independent transcriptional regulation by Runx3. *Proc Natl Acad Sci U S A, 103*(19), 7384-7389. doi:10.1073/pnas.0602470103

Zernicka-Goetz, M., Morris, S. A., & Bruce, A. W. (2009). Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. *Nat Rev Genet, 10*(7), 467-477. doi:10.1038/nrg2564

CHAPTER II

JARID2 INTRON-RETENTION REGULATES GENE EXPRESSION IN TSD

INTRODUCTION

Temperature-dependent sex determination (TSD) is a developmental phenomenon in which the bipotential gonads are driven by temperature towards an ovarian or testicular fate. In its most basic sense, the fate of the bipotential gonad relies on cell fate decisions of gonadal progenitor cells to first become supporting or steroidogenic lineages. Supporting cells then develop into Sertoli or granulosa cells, while steroidogenic progenitors become Leydig or theca cells. These cell types ultimately establish the morphology and functionality of the testis or ovary. Cell fate decisions are governed by complex gene regulatory networks that are often triggered by a particular signal. In mammals, this signal is the expression of sex determining region Y (SRY), a gene that confers testicular fate. In many turtles, including the common snapping turtle *Chelydra serpentina*, and all crocodilians, this signal is the ambient temperature of the embryo during a precise window of development known as the temperature sensitive period (TSP).

In species with TSD, the temperature of embryos during the TSP is capable of yielding predictable offspring sex ratios. In the case of the snapping turtle, females are produced at the extremes of the viable range of incubation temperatures and males are produced at intermediate temperatures. Temperatures at which 100 percent of offspring develop testes are known as male-producing temperatures (MPTs). Conversely, temperatures that induce ovarian development in 100 percent of offspring are known as female-producing temperatures (FPTs). The exact temperatures for masculinization or feminization of embryos depends on the latitude of the population of turtles being studied (Ewert, Lang, & Nelson, 2005). In this study, we have used snapping turtles from a northern population in which an intermediate temperature of 26.5° C

(MPT) produces exclusively males and a high temperature of 31° C (FPT) produces exclusively females (Rhen & Lang, 1998).

In TSD, a temperature signal is responsible for establishing highly contrasting gene expression profiles that result in differentiation of supporting cells and steroidogenic cells, and ultimately the differentiation of the bipotential gonad into an ovary or testis (Radhakrishnan, Literman, Neuwald, Severin, & Valenzuela, 2017; Rhen, Metzger, Schroeder, & Woodward, 2007; Yatsu et al., 2016). Much of the work in TSD species has focused on genes that are known to play conserved roles in gonadogenesis and sex determination in vertebrates, including FOXL2, DMRT1, SOX9, CYP19A1, and AMH (Morrish & Sinclair, 2002; Rhen et al., 2007; Rhen & Schroeder, 2010; Shoemaker, Ramsey, Queen, & Crews, 2007). While precise regulation of sex-determining genes is required to commit the bipotential gonad to a particular fate, the exact mechanism by which temperature is transduced into a signal that regulates these genes has remained elusive.

TSD is an example of developmental plasticity in which the environment has a major effect on phenotype. The sexual fate of an organism with TSD primarily relies on their response to an environmental stimulus, suggesting that differences in gene expression and phenotype are the result of epigenetic gene regulation. There are many epigenetic mechanisms for regulation of gene expression including DNA methylation, histone post-translational modifications, and regulation by non-coding RNA molecules. Though epigenetic regulators work simultaneously to regulate gene expression, certain epigenetic modifications are sufficient to explain differences in gene expression.

jumonji, AT rich interactive domain 2 (JARID2) is a candidate gene that may play such a role in TSD. JARID2 regulates gene expression during development in mouse and fly and has been identified as an accessory protein to Polycomb Repressive Complex 2 (PRC2) (Landeira & Fisher, 2011; G. Li et al., 2010; Peng et al., 2009; Sasai, Kato, Kimura, Takeuchi, & Yamaguchi, 2007). PRC2 is responsible for the deposition and maintenance of trimethylation of Histone H3 Lysine 27 (H3K27me3) which is associated with transcriptional repression, though none of the core components of PRC2 are capable of recognizing a specific DNA sequence (G. Li et al., 2010). JARID2 functions to recruit PRC2 to specific loci, which results in the repression of key developmental genes (Margueron & Reinberg, 2011; Oksuz et al., 2018). Because JARID2 and PRC2 play major roles in regulating gene expression, the hypothesis that Jarid2 plays a role in the thermal response during TSD is plausible.

Indeed, JARID2 was down-regulated by a shift from FPT to MPT in American alligators *Alligator missippiensis*, suggesting that higher expression of JARID2 may occur at temperatures that produce females whether those temperatures are warmer or cooler (Yatsu et al., 2016). Instances of differential expression and intron retention in JARID2 have also been observed during sex determination in the Australian bearded dragon (*Pogona vitticeps*), the American alligator (*Alligator missippiensis*), and the red-eared slider turtle (*Trachemys scripta*) (Deveson et al., 2017). Differential intron retention is an important regulatory mechanism that may result in reduced protein expression if the transcript is degraded or retained in the nucleus (Wong, Au, Ritchie, & Rasko, 2016). It is also possible that intron retention creates messages encoding different protein isoforms with novel functions.

In this study, we describe expression patterns of JARID2 in bipotential gonads of the common snapping turtle, *Chelydra serpentina*, during TSD. We also use long-read direct cDNA

sequencing (i.e., Nanopore technology) to identify and describe three distinct isoforms of JARID2 that arise from intron retention during mRNA processing. We observed that JARID2 expression increased dramatically in bipotential gonads of snapping turtles following a shift from a MPT (26.5°C) to a FPT (31°C) during the TSP. In addition to exhibiting temperature-sensitive expression levels, our data demonstrate that JARID2 mRNA molecules are alternatively spliced in a temperature-dependent manner.

The molecular mechanisms that govern TSD in reptiles have remained elusive for decades for a variety of reasons. One major limitation has been a lack of an appropriate system to perform functional studies of genes via overexpression or RNA interference-mediated knockdown. Thus, a major goal of this study was to develop and validate an *in vitro* model for manipulation of gene expression in embryonic turtle cells during the thermosensitive period for sex determination. A second goal of this study was to investigate the function of different JARID2 isoforms by overexpressing them in primary culture of gonadal cells isolated from embryos during the TSP. Here we describe the effects of temperature on transcriptome-wide patterns of gene expression in dissociated primary gonad cell culture and the effects of overexpression of all three JARID2 isoforms in this system.

MATERIALS AND METHODS

Nanopore Direct cDNA Sequencing

Reliable detection of alternative splicing events in long multi-exonic transcripts is complicated when using short-read sequencing technology. We therefore used Nanopore direct cDNA sequencing to detect alternatively spliced transcripts and quantify gene expression during snapping turtle TSD. We sampled gonads from snapping turtle embryos during a thermal shift experiment that allows us to investigate the effects of temperature on the transcriptome of

bipotential gonads during the TSP. We incubated snapping turtle eggs at a male-producing temperature (MPT) of 26.5ºC until embryos reached stage 17 at which point eggs either remained at the MPT or were shifted to a female-producing temperature (FPT) of 31ºC. We collected embryonic gonads 24 hours and 48 hours after the start of the temperature shift. We pooled 20 gonad pairs per sample balanced across 4 clutches and extracted total RNA using 1 ml of TRIzol Reagent (Invitrogen; catalog number:15596026). We then sequenced mRNA-enriched cDNA libraries on a GridION instrument according to the Nanopore direct cDNA sequencing protocol.

We generated an average of 1.6 million reads per library. The resulting libraries were demultiplexed and trimmed with Porechop (v0.2.4, **https://github.com/rrwick/Porechop**) and were aligned to the *C. serpentina* genome with Minimap2 (H. Li, 2018). Gene-wise read counts were obtained with featureCounts (v1.6.4) in long read mode for differential expression analysis (Liao, Smyth, & Shi, 2014). Transcript variants/isoforms were predicted and isoform-wise read counts were obtained with FLAIR (Tang et al., 2020). We performed differential gene expression analysis using DESeq2 between temperature treatments (Love, Huber, & Anders, 2014). We also performed differential isoform expression analysis (DESeq2) and differential transcript usage analysis (FLAIR, diff_iso_usage.py). Multiple testing corrections for differential expression analysis and differential transcript usage analysis were performed using the Benjamini-Hochberg method.

Quantification of JARID2 Splicing with qRT-PCR

We further investigated the temperature-dependent splicing of JARID2 by performing quantitative real-time PCR (qRT-PCR) to quantify levels of intron retention. We designed primers across the exon/exon and exon/intron boundaries that would allow us to measure the rates at which introns are retained in JARID2 transcripts. We extracted total RNA from 12 samples consisting of

embryonic gonads incubated at constant MPT or shifted to FPT and hatchling testes and ovaries at 30-days of age (n=3). For each embryonic sample we pooled the gonads from two embryos. Each hatchling sample consisted of RNA from a single pair of gonads. RNA extraction was performed with the PicoPure RNA Isolation Kit (Life Technologies; catalog number: KIT0204) with an on-column DNAse I treatment to degrade residual DNA. Purified total RNA was screened for genomic DNA contamination via qPCR prior to cDNA synthesis. We observed no amplification from RNA, demonstrating that RNA samples were free of genomic DNA. Pure RNA (142 ng/sample) was then reverse transcribed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems; catalog number 4368814) with Oligo (dT) priming. We then performed qPCR to quantify the rates of intron retention between incubation temperatures during TSD as well as between sexes of hatchling turtles. We generated standard curves for each primer pair to calculate absolute abundance of each exon/exon and exon/intron boundary.

JARID2 Structure and Function

We used InterProScan (v5.36-75.0) to analyze the predicted amino acid sequences from the JARID2 isoforms to identify the functional domains present in each isoform (Jones et al., 2014). To understand the evolution of JARID2 in vertebrates and species with TSD, we performed proteome-wide phylogenetic analysis based on multiple-sequence alignment and resolved gene trees with OrthoFinder (Emms & Kelly, 2019). We did not include the truncated isoforms of JARID2 in this analysis. We analyzed the orthogroup containing the canonical protein for JARID2 as well as orthologs in several turtles, birds, mammals, and crocodilians.

Molecular Cloning of JARID2

We cloned all three JARID2 isoforms into expression vectors to produce JARID2-mCherry fusion proteins that could be used to study the subcellular localization and the function of snapping

turtle JARID2. Primers were designed to amplify each isoform beginning at the start codon and ending at the last amino acid, but not including the stop codon for each isoform (Table 1). We constructed mRNA-enriched cDNA libraries from pooled RNA extracted from embryonic gonads at various stages of the TSP (Stages 17, 18, 19, and 20) incubated at MPT and FPT. We then amplified and purified cDNA for each JARID2 isoform and cloned them into pEF1alpha-N1-mCherry plasmids (Takara Bio; catalog number: 631969) to create JARID2-mCherry fusion proteins driven by the human elongation factor 1 alpha (EF1a) promoter. Prior transient transfection experiments with this vector shows that this promoter drives very high expression of the mCherry transcript and protein in snapping turtle gonadal cells. We validated the sequence of each JARID2-mCherry isoform fusion gene by Sanger sequencing. Sequencing reactions were prepared with the BrightDye Terminator Cycle Sequencing Kit (MCLAB; catalog number BDT3-100) using 300 ng of plasmid as template. We trimmed the resulting reads for quality and used CLC Genomics Workbench (Version 11.0; CLC bio, Cambridge, MA) to perform a multiple sequence alignment and determine consensus sequences for each isoform.

Fluorescence Microscopy

We analyzed the localization of fluorescently tagged JARID2 isoforms in primary cultures of snapping turtle ovarian cells. We euthanized hatchling turtles (<1 year old) and dissected ovaries under sterile conditions. We minced the gonads and dissociated tissue in 1% Trypsin-EDTA (Gibco; catalog number 15090046) at 30°C until a single cell suspension was achieved. We then seeded cells onto a glass coverslip and incubated cells in Leibovitz medium (L-15, no phenol red) (Gibco; catalog number 21083027) at 30°C for 48 hours. Cells in different wells were transfected with plasmid DNA containing a gene encoding one of three predicted isoforms of the JARID2 protein in-frame with the mCherry fluorescent protein. The isoforms of the JARID2 protein

(described in detail below) are denoted by the predicted amino acid length of the resulting protein. We transfected each well with either the mCherry plasmid (control), the JARID2-1231-mCherry fusion plasmid, the JARID2-1077-mCherry fusion plasmid, or JARID2-318-mCherry fusion plasmid. Following a five-day incubation at 30°C, cells were labeled with Tubulin Tracker Green (Invitrogen; catalog number: T34075), NucRed Live 647 (Invitrogen; catalog number: R37106) and imaged on a Zeiss 510 META Laser Scanning Microscope. Tubulin Tracker Green is a cell-permeant tubulin stain with excitation and emission maxima at 494nm and 522nm, respectively. NucRed Live 647 is a cell-permeant DNA stain suitable for live-cell imaging with excitation and emission maxima at 638nm and 686nm, respectively. mCherry is a bright, commercially available fluorescent protein with excitation and emission maxima at 587nm and 610nm, respectively. To visualize mCherry, we used a Helium-Neon laser at 543nm for excitation with main dichroic HFT UV/488/543/633, secondary dichroic NFT 545, and a band-pass 585-615nm emission filter. To visualize NucRed Live 647, we used a Helium-Neon laser at 633nm for excitation with main dichroic HFT UV/488/543/633, secondary dichroic NFT 545, and a long-pass 650nm emission filter. To visualize Tubulin Tracker Green, we used an Argon laser at 488nm for excitation with main dichroic HFT UV/488/543/633, secondary dichroic NFT 545, and a band-pass 505-550nm emission filter. We verified that the fluorophores could be clearly distinguished in this configuration by conducting preliminary experiments in which images were captured of cells without fluorophores and cells with all combinations of fluorophores and assessed for background fluorescence and channel bleed-through.

In vitro Overexpression of JARID2 isoforms

We collected snapping turtle eggs from three nests in northern Minnesota with Department of Natural Resources special permit number 29256. We processed eggs, recorded egg mass and

diameter, and incubated eggs in moist vermiculite at MPT until Stage 17 (nearing the end of the TSP) (Figure 5A) as previously described (Rhen et al., 2007). We euthanized embryos via rapid decapitation and microdissected bipotential gonads from the underlying adrenal-kidney tissue. We pooled 8 gonads per sample and dissociated gonads in 1% Trypsin-EDTA at room temperature, until a single-cell suspension was achieved (roughly 15 minutes). Cells were centrifuged at 125 x g for 10 minutes, supernatant was removed, and cells were resuspended in 600 ml of L-15 media supplemented with 10% FBS and Antibiotic Antimycotic Solution (Sigma; catalog number: A5955). Dissociated cells were seeded in separate wells of 24-well plates in 600 ml of L-15 media and incubated at MPT for 48 hours to allow cells to adhere and begin proliferating. After this initial incubation, we separated samples into experimental groups and transfected cells with control mCherry plasmid or JARID2-mCherry fusion plasmid using Lipofectamine 3000 Transfection Reagent (Invitrogen; catalog number: L3000001) according the manufacturer's instructions. Five experimental transfection groups included vehicle control (Lipofectamine reagents, no DNA), mCherry (transfected with plasmid expressing only mCherry fluorescent protein), or one of three JARID2-mCherry isoforms (transfected with plasmid containing each JARID2 isoform). Cells in each transfection group were incubated at MPT or FPT for a fully factorial design. After 5 days, we analyzed confluency and transfection efficiency via brightfield and fluorescence microscopy and collected cells for RNA extraction. Cells were collected by removing media and gently washing once with 500 □l L-15, followed by a 3-minute incubation in 0.25% Trypsin-EDTA/PBS, after which cells were sufficiently detached from the well surface. Suspended cells were then centrifuged for 5 minutes at 500 x g, the supernatant was removed, and the cells were resuspended in PicoPure extraction buffer. RNA was extracted with the PicoPure RNA Isolation Kit (Life Technologies; catalog number: KIT0204) and subjected to on-column DNase I treatment to digest

genomic DNA and plasmid DNA. RNA quality was assessed by Bioanalyzer and all samples had a RIN above 8.7. mRNA-enriched libraries were prepared with the Illumina TruSeq paired-end stranded mRNA library preparation kit from 400 ng of total RNA. The resulting libraries were sequenced for 150 cycles on a NovaSeq6000 instrument. The sequencing data was trimmed for quality and adapter content using Trimmomatic and processed libraries were aligned to the *C. serpentina* genome using HISAT2 (Bolger, Lohse, & Usadel, 2014; Das et al., 2020; Kim, Paggi, Park, Bennett, & Salzberg, 2019). Gene-wise counts were obtained using featureCounts in paired-end mode (Liao et al., 2014).

We sequenced 32 samples, 16 at FPT and 15 at MPT. At each temperature, each group consisted of 3 samples, with the exception of the JARID2-1231 group which had 4 samples. One sample, an mCherry control at MPT, yielded a low number of reads and poor alignment rate and was excluded from subsequent analyses. The final number of samples is indicated in (Figure 5C).

Expression Data Analysis

We performed differential expression analysis using a combinatorial approach with three well-known differential expression softwares, DESeq2, edgeR, and limma, to identify genes that display expression differences between treatment groups (Love et al., 2014; Ritchie et al., 2015; Robinson, McCarthy, & Smyth, 2010). To ensure that changes in gene expression were the result of JARID2 overexpression and not the mCherry protein, we first compared cells transfected with the mCherry plasmid and vehicle treated groups with DESeq2. We identified 0 significant genes (Benjamini-Hochberg adjusted p-value < 0.05), suggesting that the mCherry-containing plasmid had minimal effects on gene expression in turtle cells. We therefore combined the mCherry transfected group and the vehicle-treated group into a single control for subsequent analyses. To ensure that the overexpression was sufficient, we aligned the reads to the mCherry coding sequence

and quantified mCherry between groups (Figure 5D). We then performed comparisons between incubation temperatures, plasmid transfection groups (control and JARID2 isoforms), and tested for an interaction between temperature and plasmid. We identified a list of genes that were differentially expressed with a raw p-value < 0.01 using all three programs (i.e., the results were robust to statistical model used by the software). Genes in this list were further analyzed with a two-way ANOVA using the FPKM data for each gene. Genes with a p-value < 0.05 for temperature, plasmid transfection, and/or the temperature by plasmid interaction from the two-way ANOVA were called significantly differentially expressed.

We performed Gene Ontology (GO) enrichment analysis for genes that were differentially expressed using GOATOOLS (find_enrichment.py) and a custom snapping turtle GO annotation (Das et al., 2020; Klopfenstein et al., 2018). GOATOOLS offers advantages over other commonly used software, because it allows for efficient parsing of the GO hierarchy. This becomes especially useful in non-model organisms, where it is common for researchers to use a list of gene symbols from their species, adopt GO terms annotated to genes with the same symbol in another species, and analyze these annotations against a background set from the other species. We also performed *de novo* motif discovery and enrichment analysis with HOMER2 using proximal promoters (as defined as 1kb upstream TSS) from differentially expressed genes (Heinz et al., 2010). Sequences for proximal promoters were acquired using the *C. serpentina* genome annotation and gfftobed (https://github.com/jacobbierstedt/gfftobed). We used HOMER2 to compare promoters of focal genes to promoter regions of all other snapping turtle protein-coding genes to identify putative transcription factor binding sites that are enriched in the focal genes.

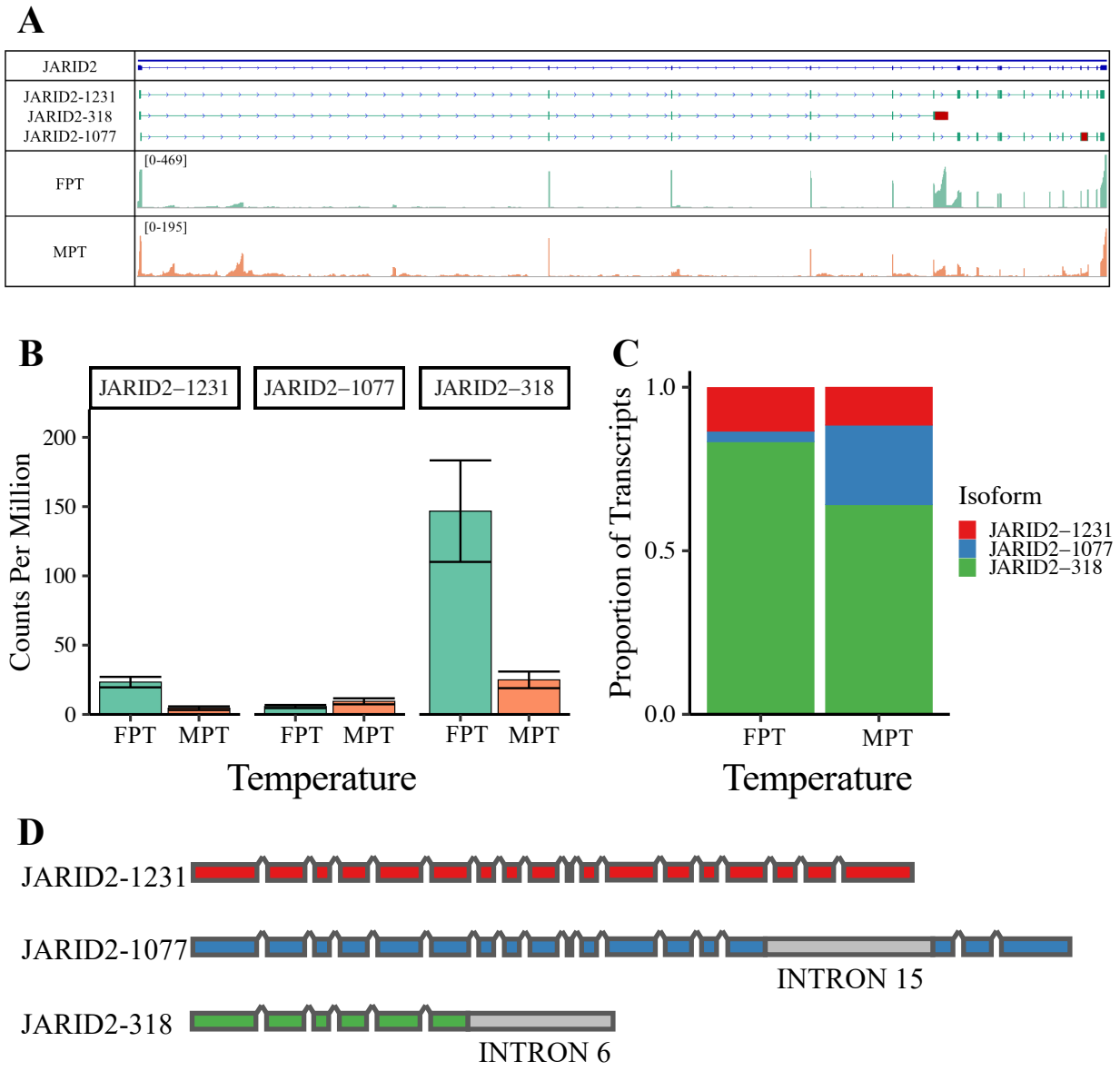


Figure 1: JARID2 occurs in three distinct transcript variants during the TSP
A. Genomic tracks showing intron/exon structure of the JARID2 gene (Blue), tracks showing the intron/exon structure of the JARID2 transcripts (Green) with retained introns (Red), and splice-aware total read coverage of Nanopore reads at MPT and FPT. B. Mean expression bar plots for each isoform based on full-length Nanopore cDNA reads (error bars indicate mean standard error). C. Stacked bar plot showing each isoform's proportion of total JARID2 transcription. D. Schematic diagram showing intron retention of Intron 15 and Intron 6 (not to scale).

RESULTS

JARID2 mRNA is Spliced into Three Main Transcripts

The full transcriptome-wide analysis of long-read data will be published in a later paper. Here we take advantage of long-read sequencing data to directly identify and quantify expression of novel JARID2 isoforms. We identified three JARID2 isoforms that had full-length read support (Figure 1A-D). When analyzing read counts at the gene level, JARID2 was differentially expressed between temperatures with higher expression at the FPT than at the MPT (Benjamini-Hochberg adjusted p-value < 1e-6). However, there were also significant differences between temperatures at the isoform level. The first isoform of JARID2 contains 18 exons and would be translated into a protein comprised of 1231 residues, which we refer to as JARID2-1231 (Figure 1A). The JARID2-1231 isoform was differentially expressed with significantly higher expression at the FPT than at the MPT (Figure 1B). A second isoform of JARID2 retains intron 15 near the 3' end of the transcript (Figure 1A,D). Intron 15 contains a premature stop codon and this transcript would be translated into a protein containing 1077 amino acids, which we call JARID2-1077. The JARID2-1077 isoform was neither significantly differentially expressed nor was there a statistically significant shift in proportions between temperatures, but it tended to make up a higher proportion of total transcripts at MPT (Figure 1C). Though not statistically significant, the pattern can also be seen in two short-read RNA-seq datasets and qPCR results (below). A third isoform of JARID2 retains intron 6 with a stop codon that would result in a protein comprised of 318 amino acids, which we refer to as JARID2-318 (Figure 1A,D). The JARID2-318 isoform was strongly differentially expressed between temperatures and exhibited significantly higher expression at the FPT (Figure 1B). It is clear that the total transcriptional output of the JARID2 gene varies between temperatures, with higher expression at FPT. In addition, we observe temperature-dependent

splicing of JARID2 during the TSP in the snapping turtle with a shift in isoform ratios between MPT (JARID2-1077) versus FPT (JARID2-318) (Figure 1C).

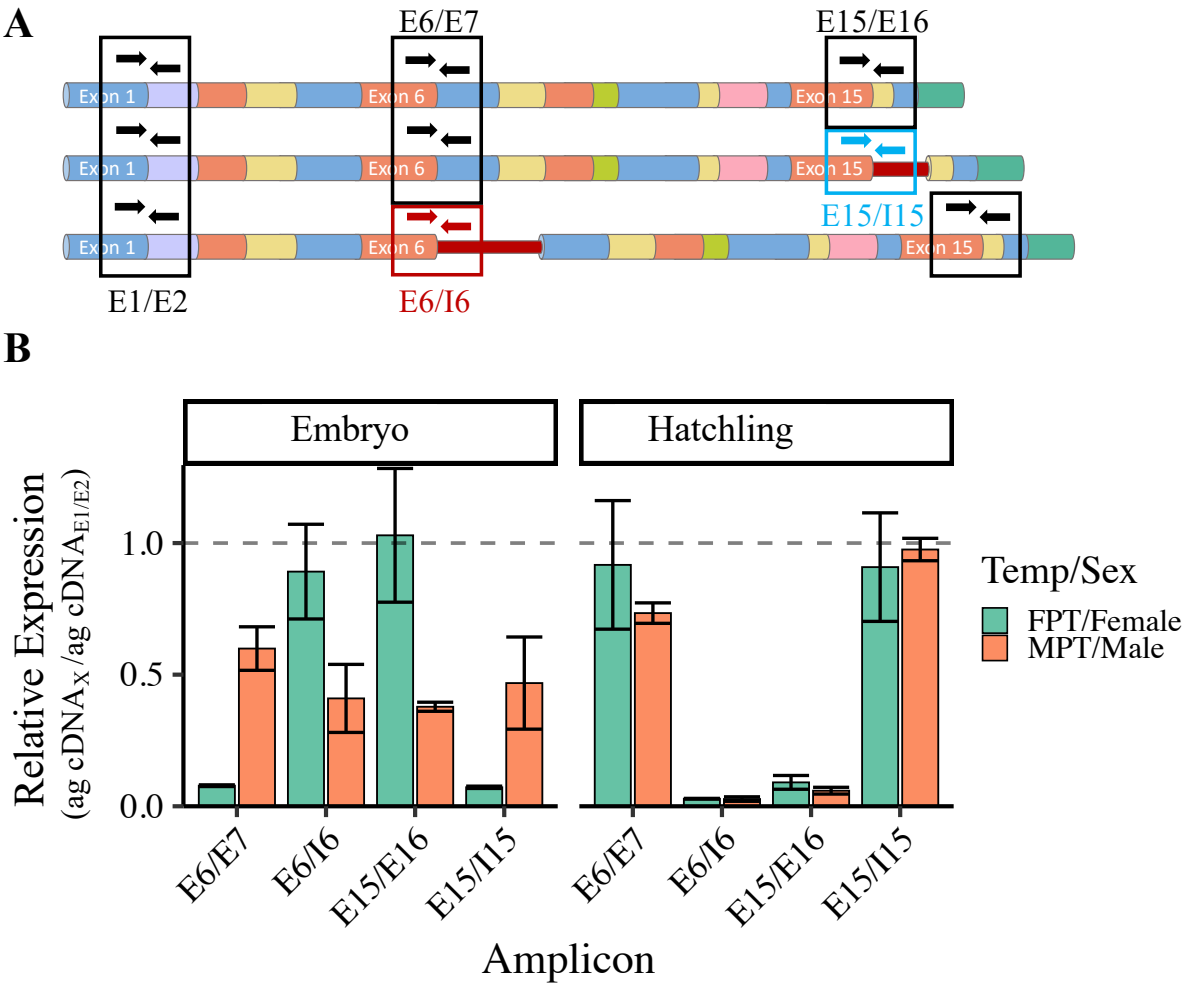


Figure 2: Quantification of JARID2 splicing with qPCR
Schematic and bar plots showing relative expression of JARID2 fragments spanning intron/exon and exon/exon boundaries in embryonic and hatchling gonads at MPT and FPT in embryonic samples and in male or female hatchling samples. A. Schematic diagram showing primers designed to span splice junctions of JARID2 transcripts. B. Bar plot showing expression for each primer pair normalized to expression for a primer pair spanning junction between Exon 1 and Exon 2 of JARID2 (ag $cDNA_X$/ag $cDNA_{E1/E2}$).

Confirmation of JARID2 Intron Retention via qPCR

qPCR with primers spanning exon-intron and exon-exon boundaries supported results from long-read sequencing (Figure 2A). Intron 6 was more highly retained in embryonic gonads at the FPT than at the MPT (Figure 2B). We also found that expression calculated from a primer pair spanning the Exon 15 and Exon 16 junction was higher at FPT, but primer pairs spanning E15/I15 and I15/E16 were not significantly different between temperatures (Figure 2B). Differences between Nanopore and qPCR could be due to small sample sizes.

We observe that Intron 6 was retained at higher levels in embryonic gonads during the TSP than in hatchling testes and ovaries (Figure 2B). In hatchlings, Intron 6 is not retained preferentially in either sex and expression of transcripts containing Intron 6 is lower than transcripts with Exon 6/Exon 7 splicing (Figure 2B). Conversely, we observe that Intron 15 was retained at higher levels than Exon15/Exon 16 spliced transcripts in hatchlings, with no difference between the sexes (Figure 2B). This suggests that the JARID2 mRNA with a retained Intron 6 may be important during TSD but may not be important for the function of differentiated gonads.

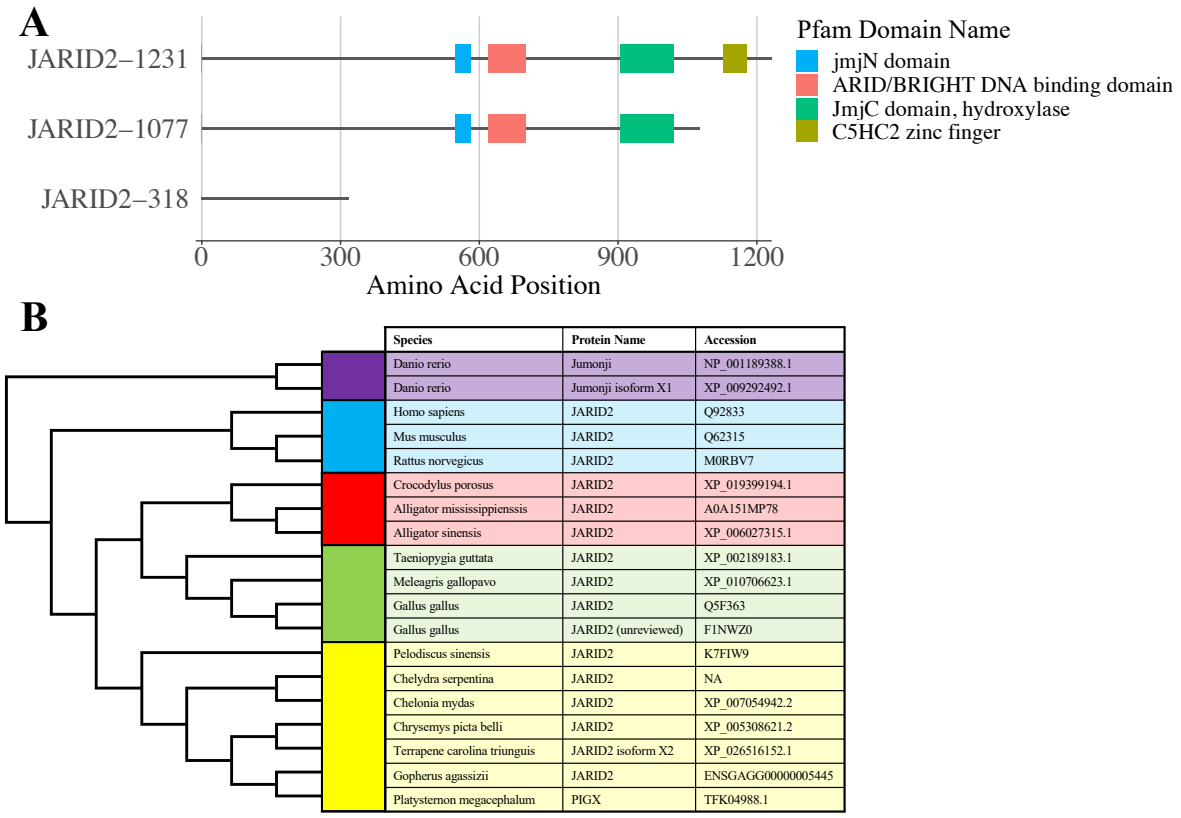
Structure of JARID2 Isoforms

We examined the primary sequence of the predicted snapping turtle JARID2 isoforms to identify putative functional domains (Figure 3A). The canonical JARID2-1231 isoform contains 1231 amino acids and four domains, including *JmjN* (residues 549-582; Pfam: PF02375), *ARID/Bright DNA binding domain* (residues 620-700; Pfam: PF01388), *JmjC* (residues 904-1019; Pfam: PF02373), and *C5HC2 zinc finger domain* (residues 1127-1177; Pfam: PF02928). These four domains are consistent with the structure of JARID2 in other vertebrates. In contrast, the JARID2-1077 isoform contains 1077 amino acids and the first three domains but is missing the

zinc finger domain. The last domain is lost because of a premature stop codon in intron 15. The JARID2-318 transcript encodes a 318 amino acid protein that is truncated by a premature stop codon in intron 6 and is missing all four domains: *JmjN/C*, *ARID*, *JmjC* and *C5HC2 zinc finger* domains. However, all three JARID2 isoforms still contain predicted nuclear localization signals in the amino terminus of the protein. The common amino terminus also contains key domains that interact with core PRC2 proteins, nucleosomes, RNA, as well as a functional domain that allosterically enhances PRC2 activity.

JARID2 Evolution

To understand the evolutionary relationships between JARID2 proteins, we performed a phylogenetic analysis with the proteomes of multiple species (Figure 3B). Notably, we observe that JARID2 falls into a separate orthogroup from other *Jumonji*-family proteins. We also observed that snapping turtle JARID2 is most closely related to orthologs in turtles. However, significant homology exists between JARID2 in all vertebrates. The amino acid sequence of JARID2 is highly conserved amongst turtles, with all turtle species analyzed having greater than 97% amino acid identity to snapping turtle with the exception of big-headed turtle JARID2, which only shares 92% amino acid identity. JARID2 in crocodilians and birds is highly conserved with identities to snapping turtle ranging from 90%-93%. Mammalian and snapping turtle JARID2 are less conserved but still display 86%-87% amino acid identity. Zebrafish JARID2 is more divergent and only shares 66% amino acid identity with snapping turtle JARID2.



Species	Protein Name	Accession
Danio rerio	Jumonji	NP_001189388.1
Danio rerio	Jumonji isoform X1	XP_009292492.1
Homo sapiens	JARID2	Q92833
Mus musculus	JARID2	Q62315
Rattus norvegicus	JARID2	M0RBV7
Crocodylus porosus	JARID2	XP_019399194.1
Alligator mississippienssis	JARID2	A0A151MP78
Alligator sinensis	JARID2	XP_006027315.1
Taeniopygia guttata	JARID2	XP_002189183.1
Meleagris gallopavo	JARID2	XP_010706623.1
Gallus gallus	JARID2	Q5F363
Gallus gallus	JARID2 (unreviewed)	F1NWZ0
Pelodiscus sinensis	JARID2	K7FIW9
Chelydra serpentina	JARID2	NA
Chelonia mydas	JARID2	XP_007054942.2
Chrysemys picta belli	JARID2	XP_005308621.2
Terrapene carolina triunguis	JARID2 isoform X2	XP_026516152.1
Gopherus agassizii	JARID2	ENSGAGG00000005445
Platysternon megacephalum	PIGX	TFK04988.1

Figure 3: JARID2 Functional Domains and Orthogroup Tree
A. Diagram showing functional domains of JARID2 identified by InterProScan5 from N- to C-terminus. Black line indicates amino acid sequence and colored rectangles indicate locations of Pfam domains. B. Phylogenetic tree of JARID2 orthogroup from OrthoFinder. Colors indicate taxonomic groups: Zebrafish (purple), mammals (blue), crocodilians (red), birds (green), turtles (yellow).

JARID2 Isoforms are Localized to the Nuclei of Gonadal Cells in Culture

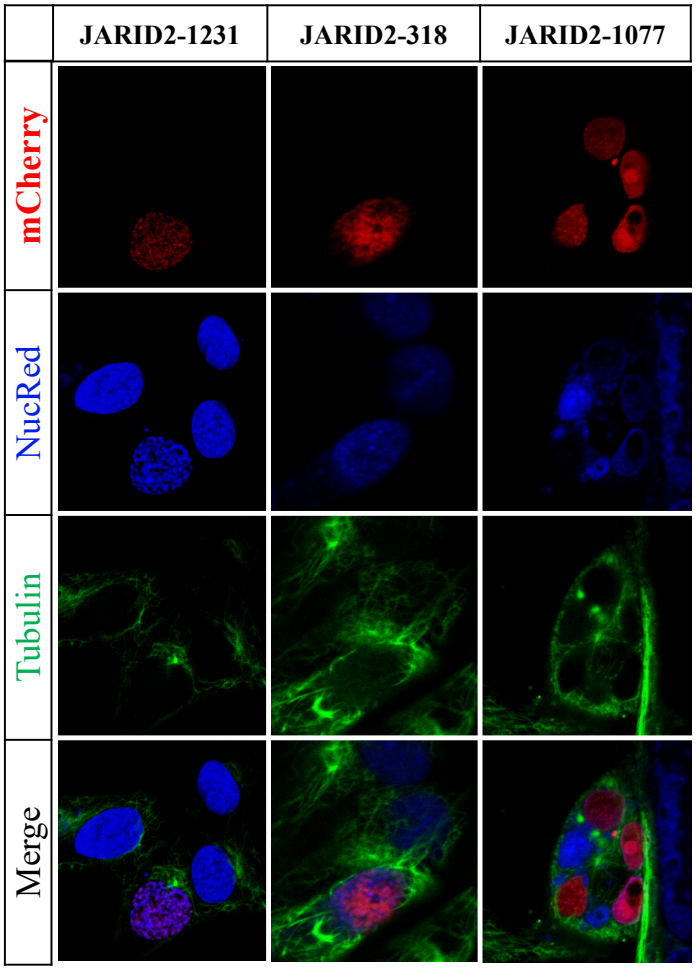


Figure 4: JARID2 isoforms are localized to the nucleus
Live-cell confocal fluorescence microscopy images showing nuclear localization of JARID2-mCherry fusion proteins in gonadal cell cultures. JARID2-mCherry fusion protein (red); DNA is stained with NucRed Live 647 (blue); tubulin is stained with Tubulin Tracker Green (green); bottom panels show merge.

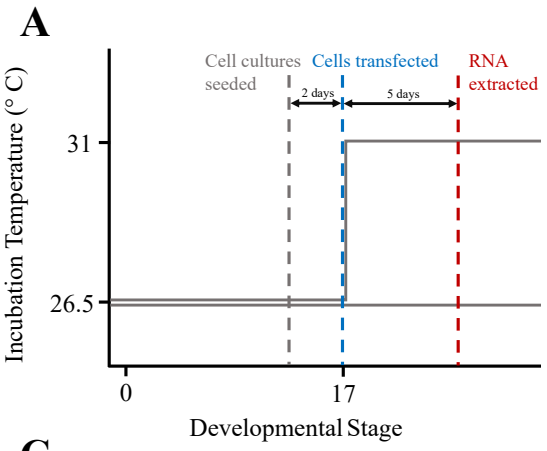
Cells expressing the control mCherry plasmid showed uniformly distributed fluorescence throughout the cytoplasm and nucleus, while cells expressing all three JARID2-mCherry fusion constructs displayed fluorescence that was colocalized with DNA stain in the nucleus (Figure 4). These results indicate that the cloned JARID2 isoforms were successfully translated and that the fusion proteins were properly folded. These findings also suggest the presence of one or more nuclear localization signals (NLS) within the N-terminal region of JARID2 (the first 301 amino acids that are common to all three isoforms). We examined JARID2 sequences using an online tool to find nuclear localization signals (**https://rostlab.org/services/nlsdb/**). Three motifs in the common amino terminus of all three JARID2 isoforms are likely to be NLSs: PSRKRPR (from 104-110), RKRPRL (from 106-

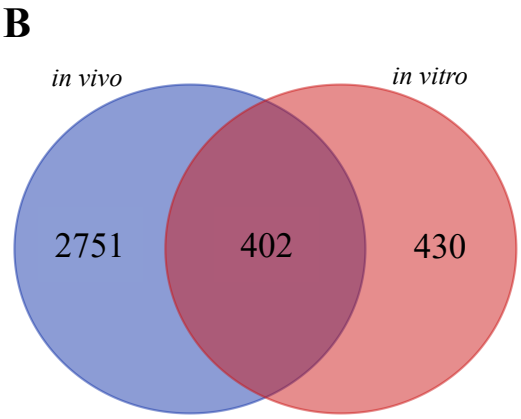
111), DLSKRK (from 146-151), SKRKPK (from 148-153), and PRKGK (from 210-214). A fourth NLS motif, PPKKMK, is found from 456-461 in the two longer JARID2 isoforms.

Temperature Effects on Gene Expression in Gonadal Cells Cultured *in vitro*

To test whether dissociated gonadal cells from embryos respond to changes in ambient temperature *in vitro*, we analyzed the effects of a temperature shift in control samples that were not transfected or transfected with the mCherry plasmid. A total of 832 genes were differentially expressed between cells incubated at the MPT and cells shifted to the FPT. In a previous study, we identified 3154 differentially expressed genes between gonads collected from embryos that had been incubated at the MPT versus gonads from embryos shifted to the FPT. Of the 832 genes differentially expressed *in vitro*, 402 (~48%) were in common with the *in vivo* set of differentially expressed genes, suggesting a significant transcriptional overlap between bipotential gonads *in vivo* and dissociated gonadal cells in culture (Figure 5B).

To understand the functional consequences of JARID2 overexpression, it is essential to identify genes that respond in the same way to temperature in dissociated gonadal cells *in vitro* and in gonads from embryos *in vivo* as well as genes that respond differently. Many genes associated with sex-determination were expressed at a detectable level but did not exhibit differential expression between temperatures. These genes include WT1, NR5A1, SOX9, KDM6B, WNT2B, WNT4, ESR1, DAX1, HSD3B1, HSD17B1, and WNT5A. Though these genes did not show temperature-dependent expression, it is important to note that they are expressed, and that differential expression *in vivo* may depend on direct cell to cell signaling or interactions with extracellular matrix proteins that are disrupted when gonads are dissociated and cells are grown *in vitro* in a monolayer. Nevertheless, numerous genes previously shown to





C

RNA Sequencing Samples					
	Vehicle	**mCherry**	**JARID2-1231**	**JARID2-1077**	**JARID2-318**
FPT	3	3	4	3	3
MPT	3	2	4	3	3

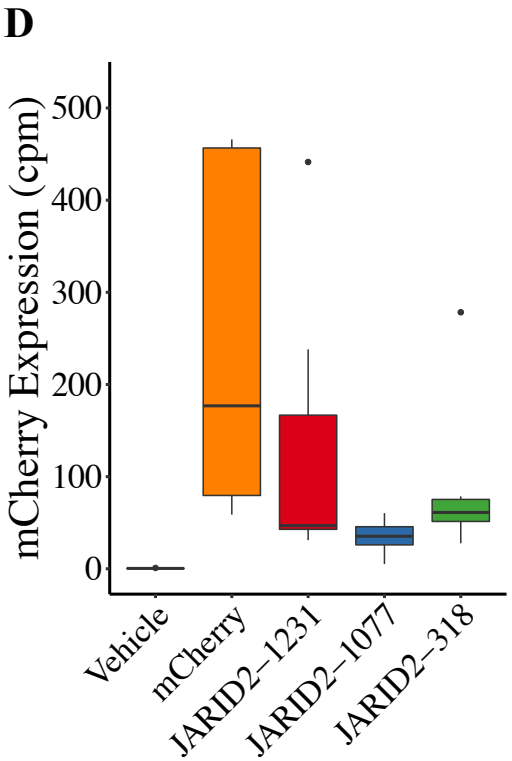


Figure 5: Experimental design and validation
A. Diagram showing experimental design with developmental time on the x-axis and incubation temperature on the y-axis. Snapping turtle eggs were incubated until approximately 2 days before stage 17 at which point embryos were euthanized and bipotential gonads were dissected, dissociated, and plated in a 24-well plate (grey dashed line). Cell cultures were incubated at MPT (26.5° C) for two days, at which point they were transfected with one of the experimental transfection groups (blue dashed line). Cell cultures were incubated at either MPT or FPT for 5 days and were collected for RNA extraction (red dashed line). B. Venn Diagram showing genes expressed between temperature shifts *in vitro* (red) and *in vivo* (blue). Substantial overlap suggests that dissociated gonads maintain much of the ability to respond to temperature, though some key components of the temperature response may be missing. C. Table showing sampling strategy for overexpression experiment. Columns indicate transfection treatment groups; rows indicate temperature regimes. D. Boxplot showing the expression of mRNAs originating from the plasmid backbone (counts per million).

exhibit differential expression between temperatures *in vivo* were also differentially expressed *in vitro*. To understand the transcriptional response to temperature in primary cell culture, we classified genes into two groups that display higher expression at FPT or MPT, respectively (Figure 6A; Figure 7A).

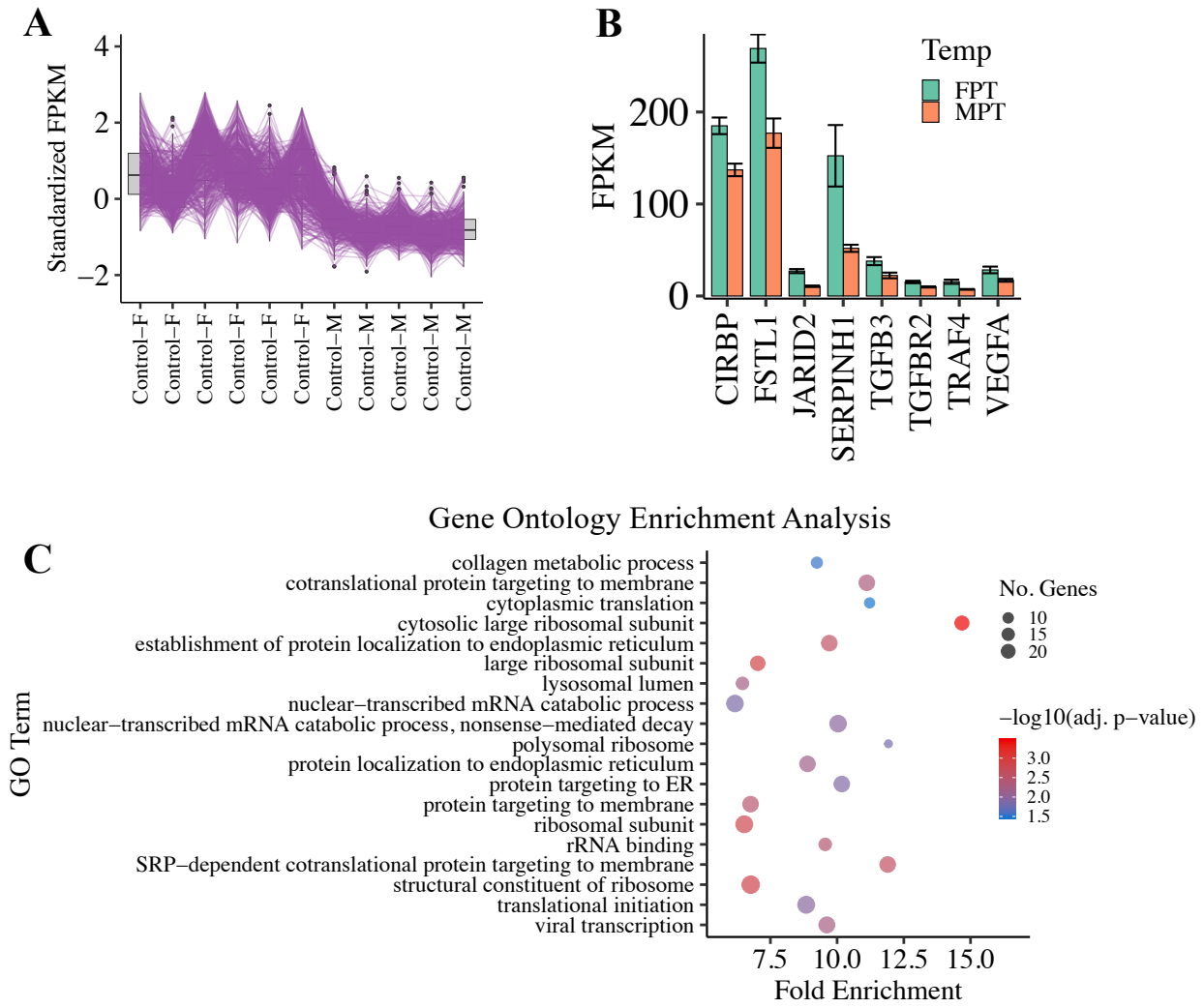


Figure 6: Genes upregulated at FPT
A. Parallel line plot showing pattern of genes that are upregulated at FPT by sample. Y-axis is Z-score standardized FPKM values (n=381). B. Bar plot showing mean expression of several genes upregulated at FPT (error bars indicate standard error). C. Plot showing enriched Gene Ontology terms for genes upregulated at FPT. Color scale shows Bonferroni adjusted p-values.

We identified 381 genes that were upregulated at FPT (Figure 6A). 127 of these genes were found to be upregulated in the previous *in vivo* study at FPT (Table 2). Notably, JARID2 and CIRBP fell into this category and were strongly upregulated at FPT. Other genes in this category include SERPINH1, NFKB2, VEGFA, CYP27B1, TGFB3, TGFBR2, SPARC, FSTL1, and GLI1 (Figure 6B). The NFKB2 pathway is involved in the immune and stress response and was shown to exhibit female-specific expression in *C. picta* (Radhakrishnan et al., 2017). Several genes in the NFKB2 signaling pathway were upregulated at FPT, including TRAF4, TNIP2, IKBKE, GSTP1, RACK1. VEGFA, which has been shown to play a role in the establishment of sex-specific vasculature, was upregulated at FPT among other genes associated with vasculogenesis (Bott, Clopton, & Cupp, 2008). SPARC, FSTL1, EEF2, and COL1A1 were shown to be associated with latitudinal differences in TSD patterns (Roush & Rhen, 2018). GLI1 is a steroidogenic cell marker that is a product of Hedgehog signaling and is upregulated at FPT (Huang & Yao, 2010; Wijgerde, Ooms, Hoogerbrugge, & Grootegoed, 2005). The most highly enriched GO categories for genes that were upregulated at FPT include terms related to extracellular matrix, regulation of translation, response to stimulus, and RNA catabolic process (Figure 6C).

A total of 451 genes were upregulated at MPT, with 258 of these genes also upregulated in the previous *in vivo* study at MPT (Table 3; Figure 7A). Interestingly, FST (*Follistatin*), which is associated with expression in granulosa cells in mammals, was upregulated at MPT. Other genes in this category include: PARP1, which encodes a protein that interacts with SRY and other SOX-family transcription factors in mammals (Lai et al., 2012; Y. Li, Oh, & Lau, 2006), HMGB3 was identified as a candidate for positive selection in TSD (Roush & Rhen, 2018). UBE3A encodes the E6-AP protein that acts as a coactivator for many steroid hormone receptors (Nawaz et al.,

1999) (Figure 7B). The MAPK cascade was highly enriched at MPT, consistent with the role of MAPK in mammalian sex determination and observations in *C. picta* (Radhakrishnan et al., 2017; Warr et al., 2012). The most highly enriched GO categories for genes that were upregulated at MPT include terms related to regulation of the cell cycle, Wnt signaling, MAPK cascade, steroid hormone receptor binding, and regulation of stem cell differentiation (Figure 7C).

A number of important genes in TSD were not expressed at either temperature in this study. Most notably, FOXL2, AMH, DMRT1, and CYP19A1 expression levels were below a limit of accurate detection and quantification (<5 FPKM). We hypothesize that these genes are downstream in the sex-determination pathway and are dependent on mechanisms that require direct cell-cell signaling, and therefore, are not expressed in the *in vitro* model.

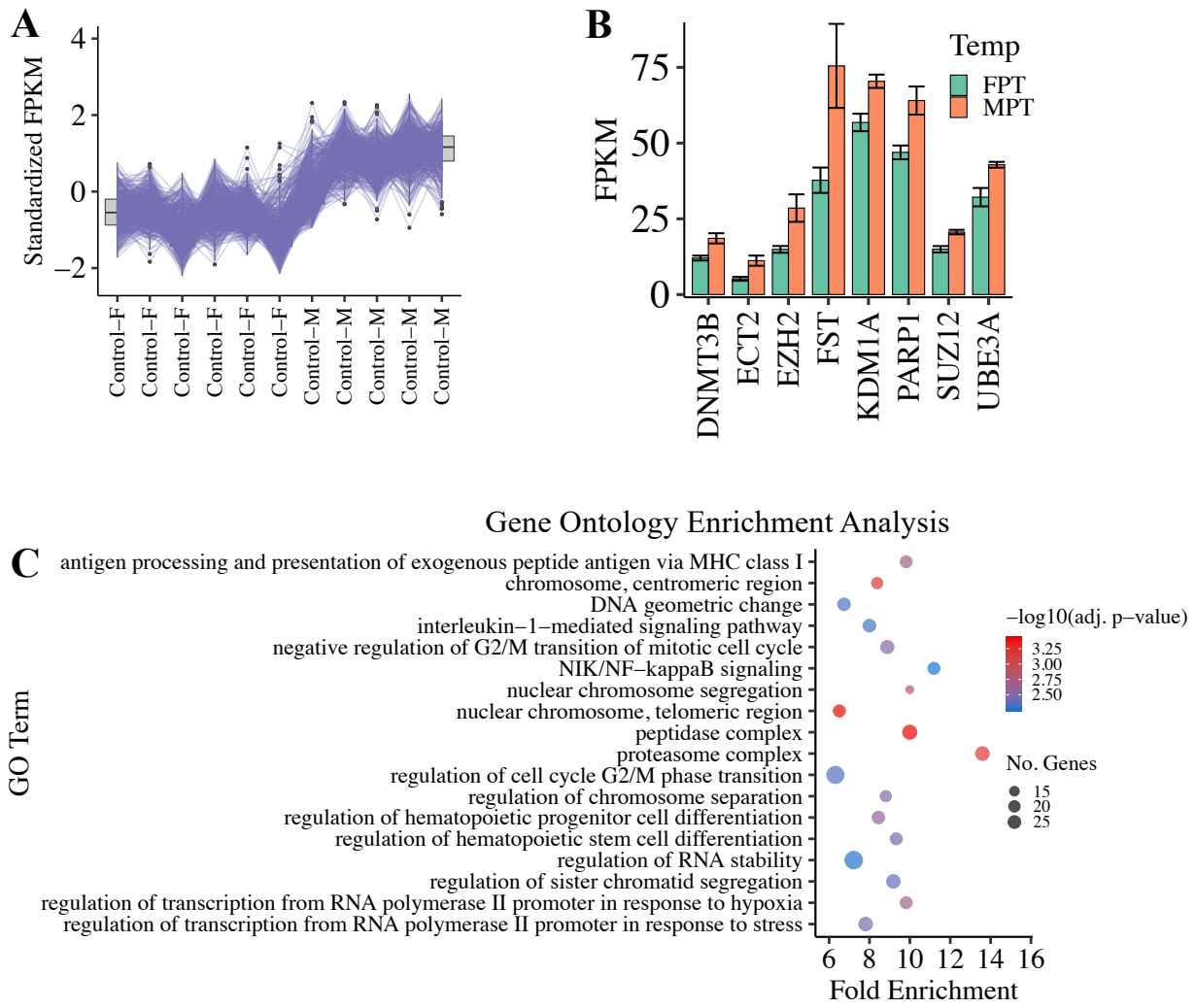


Figure 7: Genes upregulated at MPT

A. Parallel line plot showing pattern of genes that are upregulated at MPT by sample. Y-axis is Z-score standardized FPKM values (n=451). B. Bar plot showing mean expression of several genes upregulated at MPT (error bars indicate standard error). C. Plot showing enriched Gene Ontology terms for genes upregulated at MPT. Color scale shows Bonferroni adjusted p-values.

Overexpression of JARID2-1231-mCherry in Embryonic Gonadal Cells *in vitro*

Endogenous JARID2 expression responds sharply within 24 hours of a shift in incubation temperature, with greater expression in embryonic gonads and gonad cell cultures incubated at FPT than at MPT. Here, we overexpressed JARID2 in primary gonad cell cultures to observe its effect on expression of other genes across differing thermal regimes. Following incubation at MPT until stage 17, cell cultures were transfected with a JARID2 overexpression vector containing one of three JARID2 isoforms and then kept at MPT or shifted to FPT. To understand the effects of overexpression of each isoform, we compared JARID2 transfected cells to control cells (i.e., non-transfected and mCherry transfected cells).

We identified 30 genes that were differentially expressed between control cells and cells overexpressing JARID2-1231 (the canonical full-length isoform). There were 16 genes upregulated and 14 genes downregulated by JARID2-1231 (Figure 8A-D). Several interesting genes were upregulated by JARID2-1231: DDX11, which is associated with rRNA transcription and cell growth (Sun et al., 2015), SCUBE3, which is a TGF-beta ligand associated with cell migration and epithelial-mesenchymal transition (Wu et al., 2011); FNTA, which interacts with TGF-beta type I receptor (Wang et al., 1996); RGS22 a protein associated with mammalian spermatogenesis (Hu et al., 2008); and PDZRN3, shown to negatively regulate osteoblast differentiation through inhibition of Wnt signaling (Honda, Yamamoto, Ishii, & Inui, 2010) (Figure 8B). Among the 14 genes that were significantly downregulated by JARID2-1231 was TGFB3, which is expressed in mammalian testis and ovary at different stages of development (Cupp, Kim, & Skinner, 1999; Memon, Anway, Covert, Uzumcu, & Skinner, 2008). Other notable genes that were downregulated by JARID2-1231 were MAP3K12, FMO2, COL18A1, and TEX9 (Figure 8D).

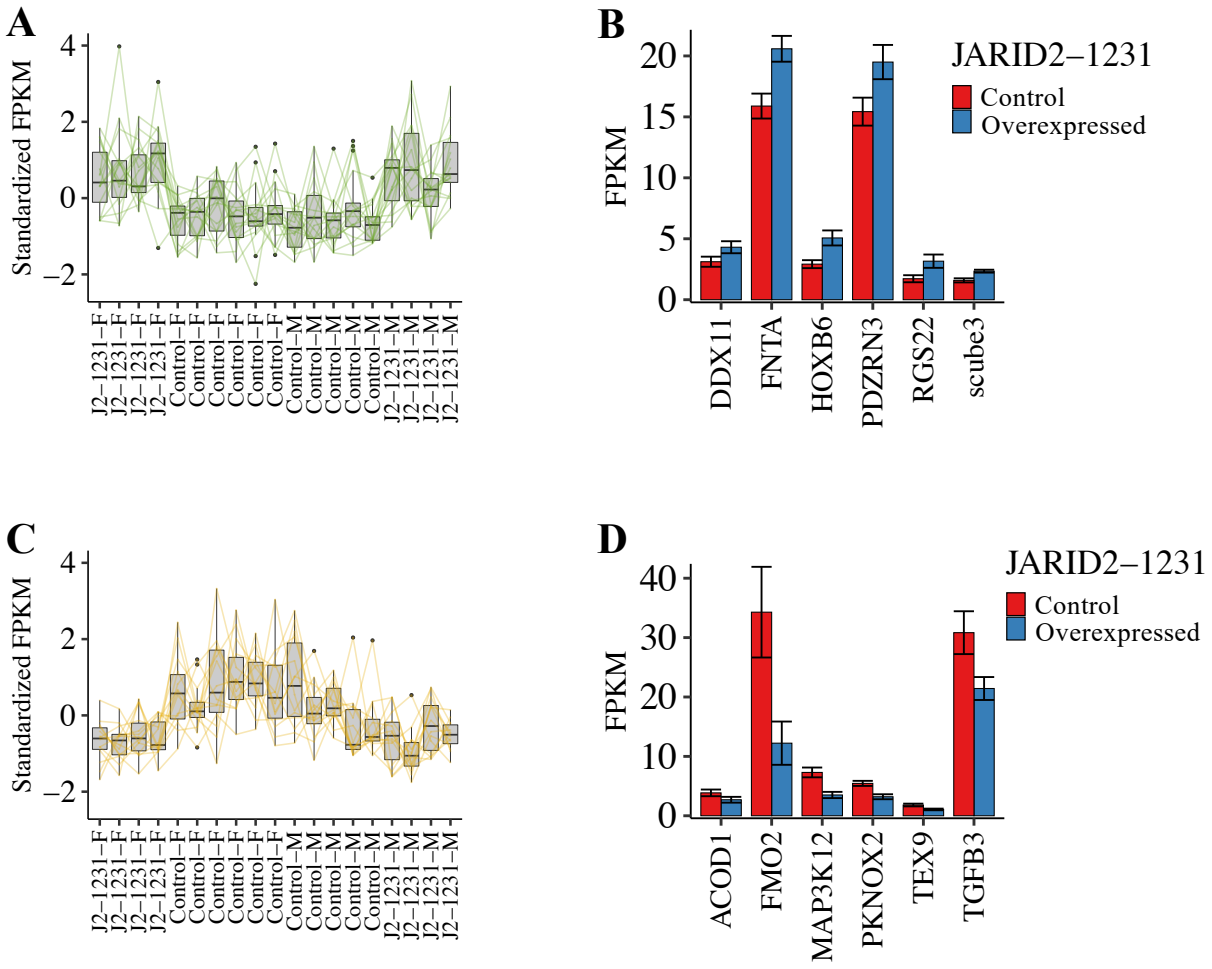


Figure 8: JARID2-1231 regulates gene expression *in vitro*
A. Parallel line plot showing pattern of genes that are upregulated by overexpression of JARID2-1231 by sample. Y-axis is Z-score standardized FPKM values (n=16). B. Bar plot showing mean expression of several genes upregulated by overexpression of JARID2-1231 (error bars indicate standard error). C. Parallel line plot showing pattern of genes that are downregulated by overexpression of JARID2-1231 by sample. Y-axis is Z-score standardized FPKM values (n=14). D. Bar plot showing mean expression of several genes downregulated by overexpression of JARID2-1231 (error bars indicate standard error).

We also identified genes that displayed a significant interaction between JARID2-1231 and temperature; in other words, these genes responded differentially to JARID2-1231 overexpression at MPT versus FPT. Notably, RDH16, HOXA5, HOXA9, and BMP3 displayed slight FPT-biased expression in control samples but were downregulated by JARID2-1231 at FPT and upregulated

by JARID2-1231 at MPT. JARID2-1231 overexpression also caused a decrease in temperature effects of PARP1. In summary, 45 genes displayed temperature-dependent expression and were also regulated by JARDI2-1231 overexpression, while 51 genes displayed a significant JARID2-1231 by temperature interaction.

Overexpression of JARID2-1077-mCherry in Embryonic Gonadal Cells *in vitro*

JARID2 is expressed at a lower level in embryonic gonads and cell cultures incubated at MPT during the thermosensitive period. However, a second isoform of JARID2 (JARID2-1077) accounts for a significantly larger proportion of JARID2 expression at MPT. We therefore performed differential expression analysis to assess the effects of overexpressing JARID2-1077 in primary gonadal cells at MPT and at FPT. We identified 203 genes that showed significant changes in expression in response to overexpression of JARID2-1077. A total of 126 genes were upregulated (37) or downregulated (89) by overexpression of JARID2-1077 (Figure 9A-E). The remaining 77 genes displayed an interaction between JARID2-1077 and incubation temperature. Gene ontology analysis showed enriched terms related to extracellular matrix, collagen, and cell adhesion (Figure 9E).

Overexpression of the JARID2-1077 isoform upregulated 37 genes in cultured gonadal cells (Figure 9A). We observed an increase in RUNX1 expression. RUNX1 is a ortholog of the *runt* gene essential for ovarian determination in *Drosophila* (Chuang, Ito, & Ito, 2013; Duffy & Gergen, 1991) and is co-expressed with FOXL2 in gonads from a wide range of vertebrate species (Nicol et al., 2019). RUNX1 is also briefly overexpressed in supporting cell lineages during differentiation of the mammalian gonad (Stevant et al., 2019; Stevant et al., 2018). Additionally, overexpression of JARID2-1077 upregulated SBNO2, AMIGO2, and KLHL2 (Figure 9B).

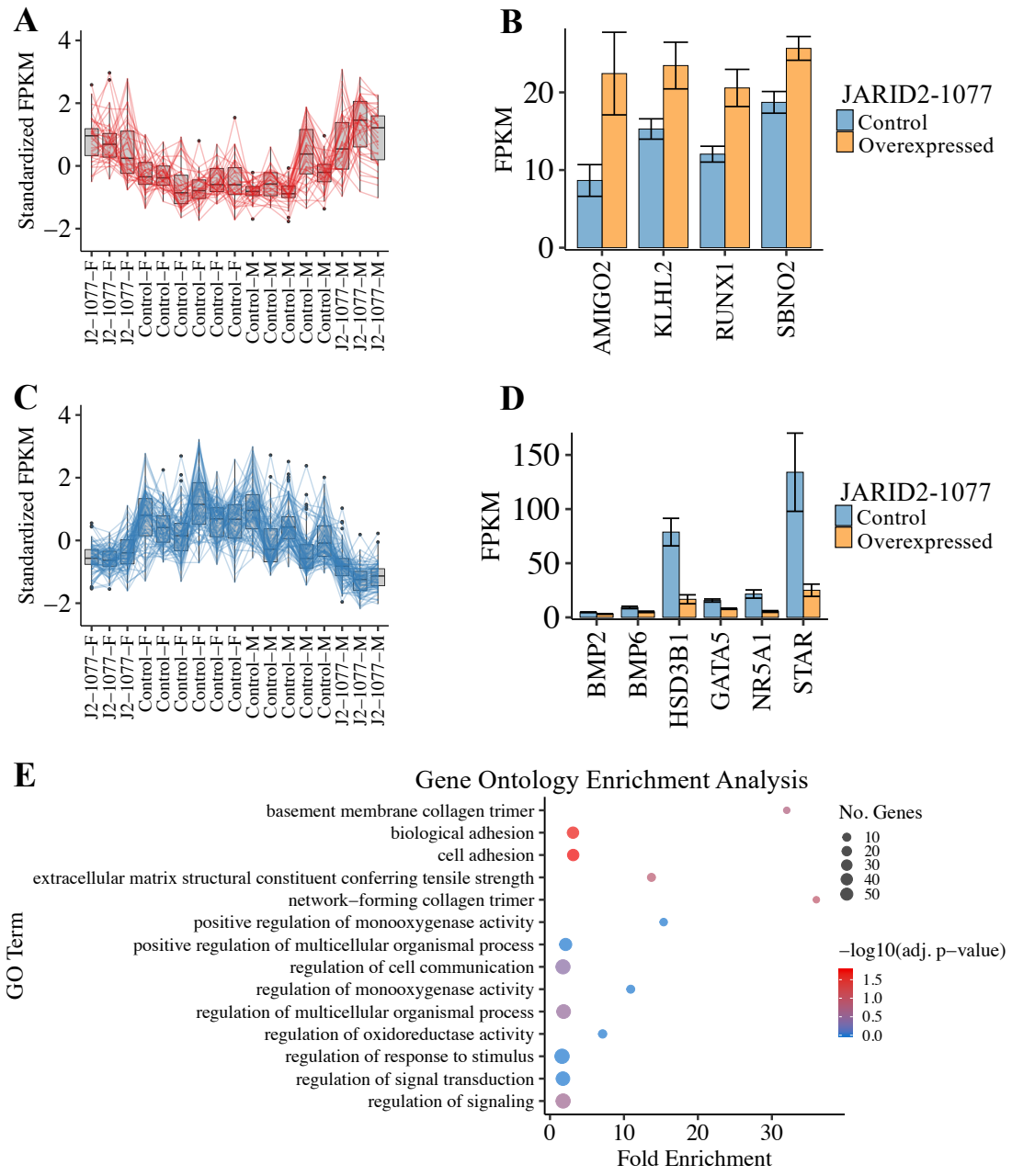


Figure 9: JARID2-1077 regulates expression of genes involved in steroidogenesis
A. Parallel line plot showing pattern of genes that are upregulated by overexpression of JARID2-1077 by sample. Y-axis is Z-score standardized FPKM values (n=37). B. Bar plot showing mean expression of several genes upregulated by overexpression of JARID2-1077 (error bars indicate standard error). C. Parallel line plot showing pattern of genes that are downregulated by overexpression of JARID2-1077 by sample. Y-axis is Z-score standardized FPKM values (n=89). D. Bar plot showing mean expression of several genes downregulated by overexpression of JARID2-1077 (error bars indicate standard error). E. Plot showing the most highly enriched Gene Ontology terms for all genes that were differentially expressed by JARID2-1077 overexpression (upregulated, downregulated, and temperature interaction). Color scale shows Bonferroni adjusted p-values.

Overexpression of JARID2-1077 downregulated 89 genes at both temperatures (MPT or FPT) (Figure 9C). The effects of JARID2-1077 were biased towards transcriptional repression, rather than activation. Most notably, gonadal progenitor cell marker NR5A1 (SF-1) and steroidogenic genes HSD3B1, FDXR, FDX2, and STAR were strongly downregulated by JARID2-1077 overexpression at both temperatures (Figure 9D). HSD3B1 and STAR are NR5A1 (SF-1) targets and play vital roles in the biosynthesis of all sex steroids (Leers-Sucheta, Morohashi, Mason, & Melner, 1997; Nakamoto et al., 2012). FDXR and FDX2 are both involved in the electron transport chain, though with differing roles in steroidogenesis. Interestingly, it has been shown in mammalian adrenocortical and granulosa cell-derived KGN cells that FDXR expression is regulated by NR5A1 (SF-1) binding an intronic enhancer. Furthermore, knockdown of NR5A1 (SF-1) decreased FDXR expression as well as the expression of STAR (Ehrlund et al., 2012; Imamichi et al., 2014; Sheftel et al., 2010). The observed effects of JARID2-1077 expression on steroidogenic genes suggests that cells overexpressing JARID2-1077 were less likely to differentiate into steroidogenic lineages. It could also suggest that steroidogenic cells are not producing steroids at this point in development at MPT. BMP2, which coordinates FST expression with FOXL2 in mammals (Kashimada et al., 2011), and BMP6, an oocyte-derived BMP ligand that regulates FSH action (Otsuka, Moore, & Shimasaki, 2001), were both downregulated by JARID2-1077 overexpression at both incubation temperatures. The bone morphogenetic protein family are members of the TGF-beta superfamily and have been relatively unstudied in reptiles with TSD. Other downregulated genes included WNT2B, GATA5, BAMBI, SMOC1, and FMO2. Gene ontology analysis revealed enriched terms related to the extracellular matrix and collagen-containing extracellular matrix.

To investigate the relationship between NR5A1 (SF-1) expression and the expression of steroidogenic genes HSD3B1, STAR, FDXR, and FDX2, we analyzed the correlation between the expression levels of each. We found strong positive correlations among NR5A1, HSD3B1, FDXR, and FDX2. This, along with evidence for the regulatory relationships between these genes in mammals, suggests a conserved regulatory relationship in snapping turtles. We hypothesize that JARID2-1077 overexpression resulted in lower NR5A1 (SF-1) expression, which could contribute to a lack of differentiation and/or reduced steroidogenic capacity of interstitial lineages in gonads at MPT.

To further understand JARID2-1077 regulation of these genes, we performed *de novo* motif finding and enrichment analysis in the promotor regions of genes that were upregulated or downregulated (Figure 12E). In upregulated genes, we identified one motif that was significantly enriched that showed homology to NKX2 motifs in human. In downregulated genes, we identified two motifs that were significantly enriched. These motifs shared homology to known motifs for retinoic acid receptor alpha (RARA/NR1B1) and forkhead box transcription factors. This is surprising because we observed no change in NR1B1 expression and FOXO6 was the only Forkhead family member that was differentially expressed. The function of these genes remains poorly understood in TSD. We speculate that a mechanism other than the expression of RARA and FOX transcription factors is having a differential impact on the expression of downstream genes.

A set of 77 genes displayed a significant interaction between JARID2-1077 and temperature. ADAMTS1 was downregulated at MPT by the JARID2-1077 isoform and this gene has previously been shown to be regulated by progesterone and dihydrotestosterone in human endometrial stromal cells (Wen, Zhu, Murakami, Leung, & MacCalman, 2006). Overexpression

of JARID2-1077 caused a slight increase in PLCL1 expression at FPT, but a sharp decrease at MPT, to the point where expression was reaching the detection limit. PLCL1 is associated with gonadal steroid hormones in mice. In PLCL1-KO mice, progesterone levels were significantly reduced compared to controls (Matsuda & Hirata, 2017; Tsutsumi et al., 2011). Overexpression of JARID2-1077 decreased NEURL1B expression at MPT. NEURL1B is an E3 ubiquitin ligase involved in the internalization and degradation of Notch ligands and is required for Notch signaling activation in signal-sending cells (Rullinkov et al., 2009; Song et al., 2006).

Overexpression of JARID2-318-mCherry in Embryonic Gonadal Cells *in vitro*

The JARID2-318 isoform is expressed at a higher level at the FPT than at the MPT. We identified 582 genes that responded to JARID2-318 overexpression and classified these genes into three groups based on expression pattern. A total of 178 genes were upregulated by JARID2-318 (Figure 10), while 193 genes were downregulated by JARID2-318 (Figure 11). Another 211 genes displayed a significant JARID2-318 by temperature interaction.

The majority of genes that were upregulated by JARID2-318 overexpression are unique to this group. Interestingly, RUNX1 and RUNX2 were significantly upregulated by JARID2-318. Similarly upregulated was STIP1 (a.k.a. HOP), an adapter protein that mediates the association between chaperone proteins HSP70 and HSP90, which may be implicated in the response to temperature in TSD (Hartl, 1996; Kohno et al., 2010; Pratt, 1997). SNAI2 was also upregulated by JARID2-318 and is a well-known marker of epithelial-mesenchymal transition, a regulator of differentiation, and is a downstream target of SPARC (Fenouille et al., 2012; Nieto, 2002). In many cases SNAI2 expression exhibits an inverse relationship with cellular differentiation, with high levels of SNAI2 in undifferentiated cells and decreased levels of SNAI2 as differentiation

occurs (Mistry, Chen, Wang, Zhang, & Sen, 2014; Vrenken et al., 2020). Upregulation of SNAI2 could be indicative of reduced differentiation capacity as a result of JARID2-318 overexpression. Gene ontology enrichment analysis showed enriched terms for cytoplasmic pattern recognition receptor signaling pathway, negative regulation of gene expression, and negative regulation of biological process (Figure 10C).

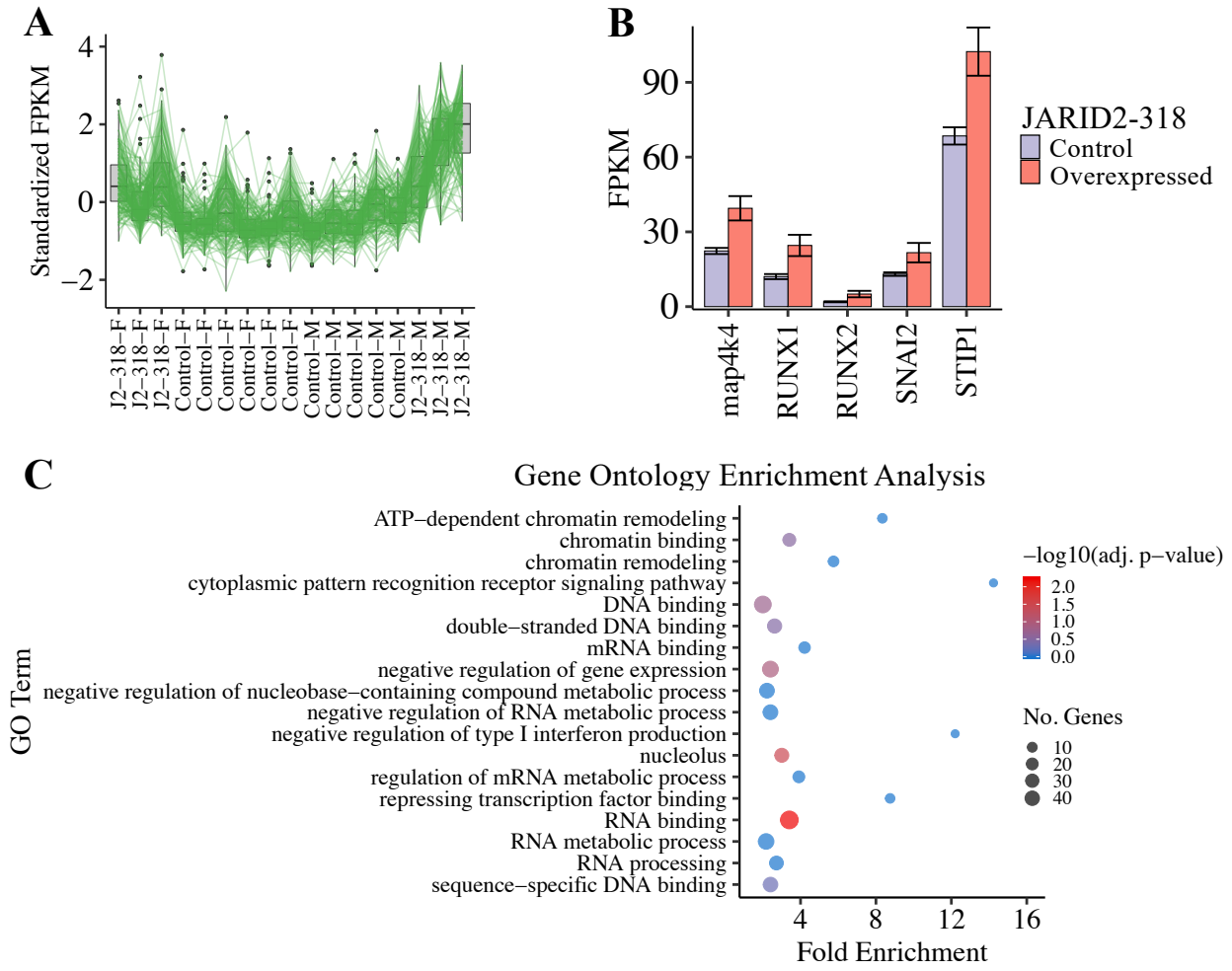


Figure 10: JARID2-318 upregulates genes associated with pluripotency
A. Parallel line plot showing pattern of genes that are upregulated by overexpression of JARID2-318 by sample. Y-axis is Z-score standardized FPKM values (n=178). B. Bar plot showing mean expression of several genes upregulated by overexpression of JARID2-318 (error bars indicate standard error). C. Plot showing the most highly enriched Gene Ontology terms in genes that were upregulated by the overexpression of JARID2-318. Color scale shows Bonferroni adjusted p-values.

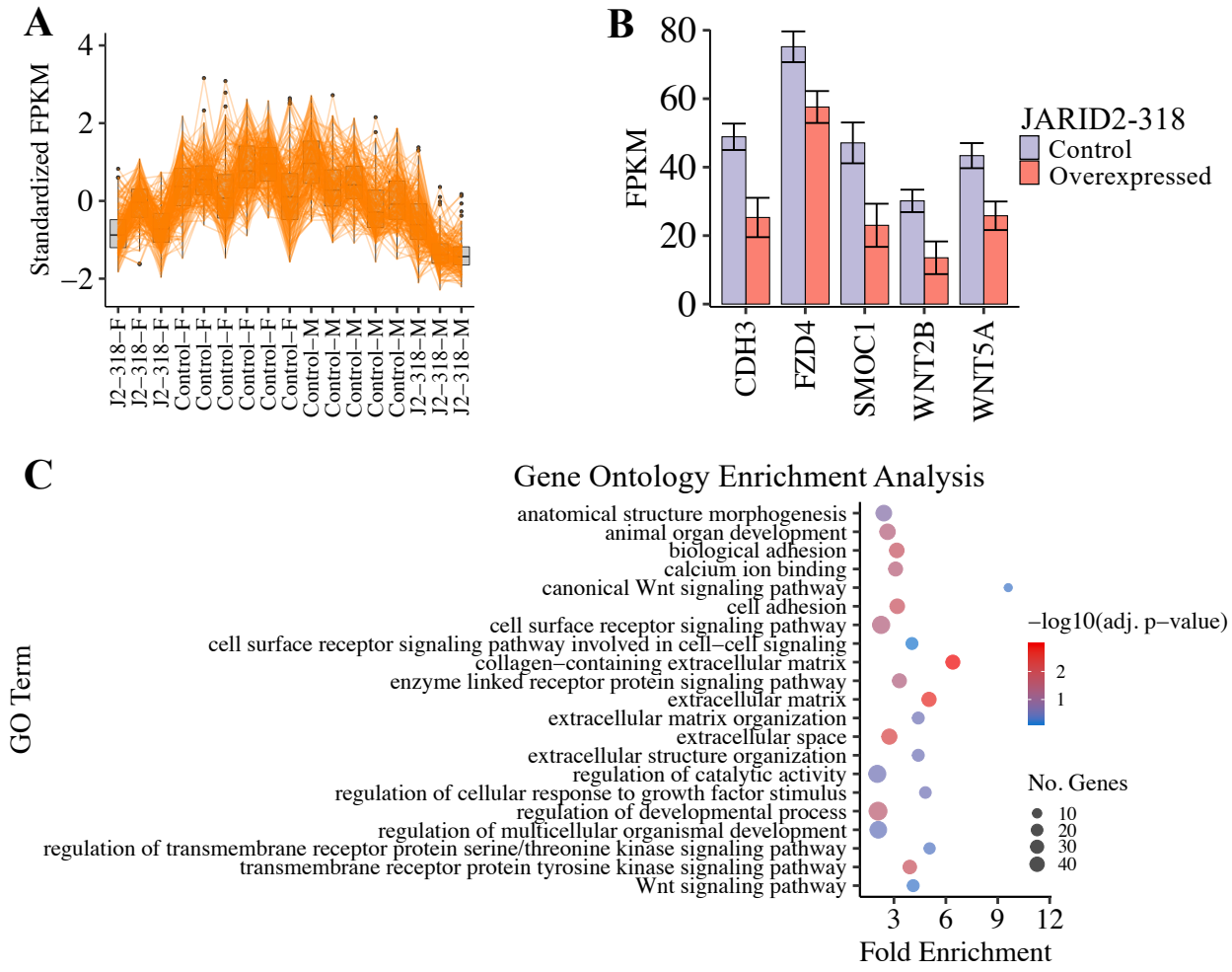


Figure 11: JARID2-318 downregulates several genes involved in the Wnt signaling pathway A. Parallel line plot showing pattern of genes that are downregulated by overexpression of JARID2-318 by sample. Y-axis is Z-score standardized FPKM values (n=193). B. Bar plot showing mean expression of several genes downregulated by overexpression of JARID2-318 (error bars indicate standard error). C. Plot showing the most highly enriched Gene Ontology terms in genes that were downregulated by the overexpression of JARID2-318. Color scale shows Bonferroni adjusted p-values.

Overexpression of JARID2-318 downregulated 193 genes (Figure 11A). BMP6, CDH3, WNT2B, COL8A1, GATA5, BAMBI, and SMOC1 all were downregulated by JARID2-318 overexpression (Figure 11B). Several gene ontology terms related to Wnt signaling were significantly enriched among genes that were downregulated by JARID2-318 (Figure 11C). Genes

associated with these terms were WNT5A, FZD4, WNT2B, and BAMBI. WNT5A has been shown to be a marker of steroidogenic cell precursors in mammals (Stevant et al., 2019; Stevant et al., 2018) and Wnt signaling is implicated in sex-determination, though the effects of Wnt signaling are likely to differ between GSD and TSD.

In addition to genes that were regulated the same way at both FPT and MPT, we identified 211 genes that displayed a significant interaction between JARID2-318 and temperature. NFKB2 exhibited temperature dependent expression in control samples, with elevated expression at FPT compared to MPT. Overexpression of JARID2-318 increased NFKB2 expression at MPT but decreased NFKB2 expression at FPT. The opposite pattern was observed for SMAD5, which displayed marginally higher expression at MPT than at FPT. Overexpression of JARID2-318 significantly decreased SMAD5 expression at MPT, but increased SMAD5 at FPT. SMAD5 is involved in BMP receptor signaling and a double knockout study of SMAD1 and SMAD5 in female mice decreased fertility and caused metastatic granulosa cell tumors. In male mice, Sertoli-Leydig tumors developed but fertility was unchanged (Feng & Derynck, 2005; Pangas et al., 2008). A decrease in SMAD5 expression could be compensated for by other SMADs because of redundancy between SMADs. However, the downregulation of SMAD5 by JARID2-318 in a temperature-dependent manner suggests some role for BMP signaling and SMADs in the developing turtle gonad. Together these results suggest that the JARID2-318 isoform may play a role in regulating gonadal differentiation during TSD.

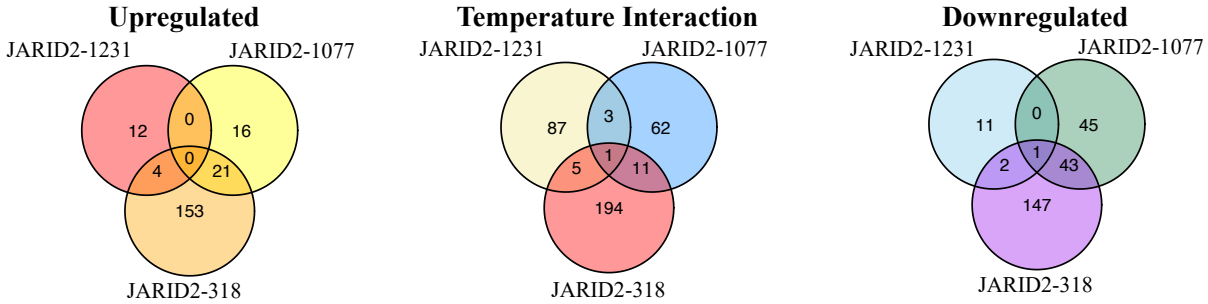
Overexpression of JARID2 Isoforms Have Common and Unique Effects on Gene Expression

A primary goal of this study was to elucidate the effects of JARID2 isoforms on gene expression in gonadal cells from embryonic turtles and the implications for TSD. In order to do

this, we investigated whether the effects of overexpression were unique or common between each isoform. Interestingly, we find that many effects were unique to each isoform. We found that genes that were upregulated by overexpression were largely unique, with minimal overlap between upregulated gene sets of each isoform treatment group. However, there was some overlap, 4 of 16 differentially expressed genes upregulated by JARID2-1231 were also upregulated by JARID2-318. These genes were SCUBE3, TRIM7, PDZRN3, and PPP1R16B. We observed a higher proportion of overlap between JARID2-1077 upregulated genes and JARID2-318 upregulated genes, with 21 genes common between the two groups, including RUNX1, MAP4K4, BASP1.

A similar pattern was observed for the downregulated genes (Figure 12A), in which we observed substantially more overlap between JARID2-1077 and JARID2-318 gene sets compared to JARID2-1231 overlap. This further suggests some functional similarities of the JARID2-318 and JARID2-1077 proteins, despite their dramatic differences in domain composition. The majority of genes that displayed a significant interaction between temperature and isoform overexpression were unique to each isoform, which may suggest that the isoforms have differing effects in response to temperature.

A



B

JARID2-1077 Enriched Motifs	Match	P-value	% in DE	% in background
	RARA	1e-10	12.36%	0.66%
	Forkhead box	1e-9	20.22%	3.04%
	NKX2	1e-8	29.7%	2.69%

Figure 12: JARID2 isoforms have both unique and shared effects on transcriptional regulation

A. Venn Diagrams comparing differentially expressed gene sets for each isoform. B. Table showing motifs enriched in the proximal promoters of genes regulated by JARID2-1077. A motif matching NKX2 was enriched in the promoters of genes upregulated by JARID2-1077 and motifs matching RARA and Forkhead box were enriched in the promoters downregulated by JARID2-1077. We did not identify significantly enriched motifs in the promoters of genes regulated by the other isoforms.

DISCUSSION

JARID2 Exists in Three Distinct Variants In TSD

Here we have identified JARID2 as one of the earliest genes to exhibit temperature driven changes in expression in gonads during the critical period for sex determination in the common snapping turtle. We also report temperature-dependent expression of three mRNA variants that encode distinct proteins: JARID2-1231, JARID2-1077, and JARID2-318. JARID2-1231 encodes the canonical JARID2 protein and shows a strong increase in expression in gonads of embryos shifted from a MPT to a FPT. One variant retains Intron 15, which contains an early stop codon and yields a truncated protein we named JARID2-1077. JARID2-1077 contains most of the functional domains found in canonical JARID2-1231. The C-terminal zinc finger is the only domain missing from JARID2-1077. The other mRNA variant retains Intron 6 and also contains an early stop codon. This variant produces a truncated protein we called JARID2-318. The shortest variant lacks the four annotated functional domains of the full-length canonical JARID2-1231: *JmjN* (residues 549-582; Pfam: PF02375), *ARID/Bright DNA binding domain* (residues 620-700; Pfam: PF01388), *JmjC* (residues 904-1019; Pfam: PF02373), and *C5HC2 zinc finger domain* (residues 1127-1177; Pfam: PF02928).

However, all three isoforms share the same 301 amino acids at the amino terminus. Located within the common amino terminus are three nuclear localization signals, a domain that interacts with core PRC2 proteins and nucleosomes, an RNA binding domain, as well as a key functional domain that allosterically enhances PRC2 activity (Son et al., 2013). All three JARID2-mCherry fusion proteins exhibit similar subcellular localization when transfected into primary ovarian cells. This result is consistent with the observation that all isoforms of JARID2 contain nuclear localization signals in the amino terminus. This finding also suggests JARID2 isoforms may retain

similar capacity to interact with chromatin. However, the isoforms exhibit functional differences from each other in terms of their target genes. We speculate that canonical JARID2-1231 retains its ability to recruit PRC2 to specific loci and facilitate the *de novo* deposition of H3K27 methylation. Li et al. found that the C5HC2 zinc finger, alone or with other domains, is responsible for the DNA-binding properties of JARID2 (G. Li et al., 2010). Even in mammalian cell lines, there are many questions about how JARID2 recruits PRC2 to specific loci. Our findings demonstrate the shorter isoforms are capable of regulating gene expression patterns and suggest that DNA binding via the C5HC2 zinc finger is not necessary for JARID2 function.

PRC2 methyltransferase requires allosteric activation of the SET domain of EZH2 by the binding of EED with H3K27me3 (Jiao & Liu, 2015; Margueron et al., 2009). Interestingly, in the absence of H3K27me3, JARID2 has been shown to mimic the tail of Histone H3 and act as a substrate for PRC2 in which case K116 of JARID2 is methylated and facilitates the *de novo* deposition of H3K27 methylation (Kasinath et al., 2018; Sanulli et al., 2015). JARID2 modulation of PRC2 activity depends on this interaction as well as an EZH2-binding domain (aa 229-349) and a nucleosome binding domain (aa 350-450) (Son, Shen, Margueron, & Reinberg, 2013) that are found in the amino terminus of all three snapping turtle isoforms.

Given that the physical interaction between PRC2 and JARID2 depends on these common domains, our data suggests that JARID2-1077 is likely capable of associating with PRC2 but it may fail to recruit PRC2 to a subset of JARID2 genomic targets due to the absence of the C-terminal zinc finger domain. Whether or not JARID2-318 is capable of associating with PRC2 is not known for certain. However, JARID2-318 retains the signals required for nuclear localization but appears to be missing part of the domain required for full stimulation of EZH2 catalytic activity. It is also missing the entire domain required for nucleosome binding. Interestingly, the

JARID2-318 isoform contains 17 unique carboxy terminal residues originating from intronic sequence. The structure and function of these amino acids is unknown, but they could contribute to the observed differences in expression patterns among the JARID2 isoforms.

Temperature-dependent expression and splicing of JARID2 provides an interesting perspective on the regulation of gene expression and sets the stage for future functional studies of the roles of PRC2 and JARID2 as transcriptional regulators in TSD.

Subcellular Localization of JARID2 Isoforms

We show here that the cloned JARID2 isoforms fused to mCherry are transcribed from the expression vector and proteins are likely folding in a stable conformation. Interestingly, we also observed that all three isoforms were found in the nucleus and appear to co-localize with the nuclear DNA stain (i.e., chromatin). Although heterogeneity of cell type and variation in the phase of the cell cycle in transfected cells impedes finer-scale inferences about sub-nuclear localization, we can speculate that all three isoforms are properly folded, localized in the nucleus, and are likely interacting with chromatin and/or PRC2 that is associated with chromatin.

Validation of *in vitro* Model of TSD

These results indicate that primary cell cultures maintain expression patterns consistent with gonads in vivo and that dissociated cells are capable of producing a transcriptional response to temperature. The processes involved in cell fate determination and cellular differentiation in the bipotential gonad in species with TSD is not as clear as in mammals, but it is likely that the same basic events occur. In mammals, progenitor cells first become specified as supporting or steroidogenic cells. Supporting cells then diverge into Sertoli or granulosa cells, while steroidogenic cells become Leydig or theca cells at a later stage of development (Stevant et al., 2019; Stevant et al., 2018). What remains unclear is the order of these events and the stimuli that

trigger them in TSD species. We hypothesize that the cell-cell interactions required for signaling between the different cell types of the bipotential gonad are disrupted and, as a result, the cells do not exhibit transcriptional patterns that are characteristic of differentiated supporting cells. The transcriptional patterns observed in this experiment are likely a composite snapshot of gene expression in cells at varying stages of specification, rather than differentiated lineages of supporting and steroidogenic cells. These cells retain their thermosensitivity *ex vivo* but lack cell-cell interactions that may contribute to determination of cell fate and cellular differentiation.

One difficulty in TSD research is that many genes exhibit thermosensitive expression patterns, and many genes are involved in the differentiation of sex-specific cell lineages. This makes it difficult to identify the signals and genes that trigger cell fate determination in the bipotential gonad, which is comprised of multiple lineages. In this *in vivo* model, transcriptional profiles are consistent with those of undifferentiated progenitor cells and they largely retain the machinery required to respond to temperature, though they lack the cell-cell interactions required to differentiate completely into Sertoli and granulosa cells or Leydig and theca cells.

JARID2-1231 Regulates Genes Involved in Differentiation

Hox genes have not previously been implicated in sex determination in mammals or reptiles and we have not identified them as thermosensitive in previous transcriptome-wide expression profiling experiments. The fact that several Hox genes are differentially expressed following JARID2-1231 overexpression suggests a conserved role of canonical isoform of JARID2 in the regulation of the Hox gene clusters (G. Li et al., 2010; Oksuz et al., 2018; Rinn et al., 2007). Other interesting transcriptional differences were observed. Genes involved in the regulation of the cell cycle, cell migration, and epithelial-to-mesenchymal transition (EMT) were overrepresented among differentially expressed genes. We also observed interactions between

JARID2 and temperature where JARID2 blocked temperature effects on some genes, while others exhibit synergism between temperature and JARID2. We hypothesize that JARID2-1231 overexpression may delay or partially inhibit differentiation of bipotential gonad cells, which would likely be associated with increased proliferation and cell migration. An interesting observation was that the majority of effects of JARDI2-1231 overexpression were temperature-dependent. This suggests that JARID2-1231 may be interacting with other temperature-dependent factors, such as ncRNAs. Elucidating the function of JARID2-1231 in TSD will require more work, but these findings do point to a conserved function of JARID2 in the regulation of gene expression and a possible novel function in the temperature-dependent regulation of gene expression.

Overexpression of JARID2-1077 Represses Transcription of Genes Involved in Steroidogenesis

Overexpression of JARID2-1077 caused more transcriptional changes than JARID2-1231 overexpression. Interestingly, the net effect of JARID2-1077 overexpression tended to be biased towards transcriptional repression, suggesting that the primary regulatory mechanism could be PRC2-mediated H3K27me3 deposition. Perhaps the most interesting finding was that JARID2-1077 strongly repressed several steroidogenic genes, including NR5A1, HSD3B1, FDXR, FDX2, and STAR. It is possible that JARID2-1077 is directly influencing expression of one or more of these genes through PRC2-mediated histone methylation. This hypothesis will require further investigation.

There is a lot of evidence that steroid hormone synthesis plays a key role in TSD: aromatase expression, estrogen synthesis, and estrogen signaling is involved in ovary determination at FPT (Crews, 1996; Ramsey & Crews, 2009). In mammals, estrogens do not play

a key role in ovary determination during embryogenesis, but they do help maintain ovarian identity after birth. Female mice that are genetically deficient in estrogen receptors a and b undergo a significant degree of postnatal sex reversal (Couse et al., 1999). In particular, adult females have ovaries that transdifferentiate seminiferous tubules, complete with Sertoli-like cells expressing SOX9 and AMH. Female mice that lack functional aromatase, the enzyme that converts androgens into estrogens, display the same phenotype. Ovaries in these females contain seminiferous tubules and Sertoli-like cells that express SOX9 (Britt et al., 2002). Interestingly, estrogen replacement blocks SOX9 expression, inhibits transdifferentiation of granulosa cells into Sertoli-like cells, and prevents development of seminiferous tubules in aromatase knockout females.

A large body of evidence in mammals shows that differentiation of steroidogenic cells requires signaling from differentiated supporting cells (Liu, Peng, Matzuk, & Yao, 2015; Stevant et al., 2019; Wijgerde et al., 2005; Yao, Whoriskey, & Capel, 2002). Additionally, steroidogenic cells in mammals do not display sexually dimorphic transcriptional patterns until well after the sex-determining period (Stevant et al., 2019). This is likely a key difference between reptiles with TSD and mammals with genotypic sex-determination. Though many of the same genes appear to be involved in sex determination, the timing of differentiation of sex-specific supporting cells versus sex-specific steroidogenic cells likely differs between mammals and non-mammalian vertebrates in which steroids play a role in sex determination. In support of this idea, thecal and granulosa cells differentiate earlier during gonad development in birds (Estermann et al., 2020), a group in which aromatase expression and estrogen synthesis are critical for ovary determination (Elbrecht & Smith, 1992).

Our results suggest that JARID2-1077 could play a key role in suppressing steroidogenesis and ovary determination. JARID2-1077 makes up a much larger proportion of JARID2 transcripts

at MPT in all assays we have conducted thus far (qPCR, Nanopore, *in vivo* RNA-seq, and *in vitro* RNA-seq). However, JARID2-1077 overexpression produced effects that were consistent between both thermal regimes, suggesting that JARID2-1077 functions the same way at MPT and FPT. Because JARID2-1077 expression is higher and it makes up a larger percentage of JARID2 transcripts at MPT, we hypothesize that JARID2-1077 plays a key role in inhibiting steroidogenesis within the steroidogenic lineage in the bipotential gonad. In turn, lower steroid levels permit testis determination.

Steroid signaling plays a key role in the determination of ovarian fate during TSD and aromatase has been identified as a key player in snapping turtles (Ramsey & Crews, 2009; Rhen et al., 2007). Interestingly, aromatase expression was not detectable in this study and we are left to speculate as to why. One possibility is that cell-cell signaling mechanisms required for induction of aromatase are disrupted upon dissociation of the gonad. A more likely possibility is that aromatase expression is extremely low, even within intact gonads (Rhen et al., 2007; Rhen and Schroeder, 2010).

We found that many genes downregulated by JARID2-1077 contained retinoic acid response elements (RAREs) within their proximal promoters. This is interesting because many RAREs tend to be sites of PRC2 recruitment and exposure to retinoic acid causes PRC2 to dissociate from chromatin (Gillespie & Gudas, 2007; Kashyap & Gudas, 2010). The aromatase (CYP19A1) promoter has been shown to be regulated by retinoic acid receptors and repressed by DMRT1, even in the presence of retinoic acid, which could allow Sertoli cells to produce the retinoic acid required for spermatogenesis (Matson et al., 2011; Minkina et al., 2014). If JARID2-1077 plays a role in recruiting PRC2 to RAREs, one would expect PRC2 occupancy at RAREs to

increase and transcription to decrease. Whether or not this happens is not clear and will require more work, specifically analyzing PRC2 occupancy and H3K27me3 levels.

JARID2-318 Regulates Expression of Genes Associated with Pluripotency and Wnt Signaling

Most JARID2 transcripts at FPT retain intron 6 and encode a truncated protein JARID2-318. This isoform is missing many of the functional domains of JARID2 but is still localized in the nucleus and may retain the ability to associate with PRC2. Interestingly, JARID2-318 overexpression had the largest effect on transcription among the three isoforms, but this could be due in part to smaller size of the plasmid causing an increase in transfection efficiency and copy number. Nonetheless, it is clear that JARID2-318 is capable of regulating gene transcription.

Overexpression of JARID2-318 caused upregulation of several genes associated with pluripotency, including SNAI2, and a decrease in several genes involved in differentiation, including BMP6. JARID2-318 also downregulated several members of the Wnt signaling pathway. Like other isoforms, temperature and JARID2-318 interacted to regulate many genes, notably the SMAD5/BMP receptor signaling pathway and the NFKB2 signaling pathway. Together these results suggest JARID2-318 may play a role in the maintenance of pluripotency. The mechanism by which JARID2-318 regulates these genes is unclear.

The lack of DNA-binding domains in JARID2-318 suggests this isoform may regulate gene expression via protein-protein interactions with PRC2. It is unclear whether such an interaction would directly alter PRC2 activity or whether the JARID2-318 isoform would compete with other JARID2 isoforms and block their stimulatory effect on PRC2. For instance, the repressive effect of the JARID2-1077 isoform on steroidogenic genes might be swamped out by much higher

expression of the JARID2-318 isoform at FPT. In support of this hypothesized mechanism, we do see a substantial number of genes that are positively regulated by overexpression of JARID2-318 which is unlike the bias towards repression with JARID2-1077. If the primary effect of JARID2-1077 is via PRC2-mediated H3K27 methylation, we would expect to see a stronger bias towards transcriptional repression. A major caveat is that we do not know which genes are direct JARID2 targets and which are indirect.

Another possible mechanism for JARID2-318 inhibition of PRC2-mediated repression is through inhibition of PRC2 recruitment to chromatin because JARID2-318 may be deficient in its ability to bind nucleosomes. This is an attractive possibility given that CYP19A1 and FOXL2 are highly temperature-sensitive and expressed at an extremely low level at MPT. The precise role of different JARID2 isoforms and PRC2 in the regulation of gene expression during TSD will require much more investigation.

CONCLUSION

We have identified three JARID2 isoforms that may be implicated in TSD and taken the first steps toward elucidating their unique and overlapping functions. Though the story of JARID2 in TSD is complex, it is clear that all three isoforms regulate transcription in embryonic gonadal cells *in vitro*, likely through PRC2-mediated repression and potentially other mechanisms. These results establish a promising foundation for further experiments to elucidate the molecular function of JARID2.

Table 1: Primers for qPCR and cloning

Experiment	Amplicon	Direction	Sequence
qPCR	JARID2-E1	Forward	TGAGTAAAGAAAGACCCAAGAGGAA
qPCR	JARID2-E2	Reverse	ACCACCCTCTCCTCTGACCAT
qPCR	JARID2-E6	Forward	AGATCGGCTCAGGACTTAAGGA
qPCR	JARID2-E7	Reverse	CATTTGCACCTGGAGATGACA
qPCR	JARID2-E6	Forward	GCTAACAACCACCACACTCTTCATA
qPCR	JARID2-I6	Reverse	CATGTGCCCAAATGAGAGTACAA
qPCR	JARID2-I6	Forward	TCCTTCTGCTTCCTCCACAA
qPCR	JARID2-E7	Forward	AGAGCCACTGCTGGGAAGAA
qPCR	JARID2-E7	Reverse	TTCACAGGAGGCATTTTCAGTTT
qPCR	JARID2-E15	Forward	GGAAAATGGACCGACTCTCA
qPCR	JARID2-E16	Reverse	ATAGCGGGCTGAAGAATGGA
qPCR	JARID2-E15	Forward	AAAATGGACCGACTCTCACTACAAT
qPCR	JARID2-I15	Reverse	GAAACACAGGGTACATGAAACAGACT
qPCR	JARID2-I15	Forward	TCTCCTGAGTGTAGCTTTGCAG
qPCR	JARID2-E16	Reverse	ATAGCGGGCTGAAGAATGGA
qPCR	JARID2-E18	Forward	CCAGCACATCACTATGCATCTGT
qPCR	JARID2-E18	Reverse	CAGGCAGTCTTTTTTTCCATCTC
Cloning	All JARID2	Forward	TCAGGTGTCGTGACGCTAGCTTTGGATACCAGAATGAGTAAAG
Cloning	JARID2-1231	Reverse	GGATCCCGGGCCCGCGGTACACTTGAAGCACTTTTGGATG
Cloning	JARID2-1077	Reverse	GGATCCCGGGCCCGCGGTACCCTGAGCTCTCCAAGAAGTG
Cloning	JARID2-318	Reverse	GGATCCCGGGCCCGCGGTACGCACTTTATGAGCACATGC

Table 2: Genes upregulated at FPT *in vivo* and *in vitro*

Gene ID	Gene	Product	log2FC	p-value	Adj. p-value
CS000009596	NFKB2	nuclear factor kappa B subunit 2	-0.77	1.21E-06	2.04E-04
CS000004172	SERPINH1	serpin peptidase inhibitor%2C clade H (heat shock protein 47)%2C member 1%2C (collagen binding protein 1)	-1.47	4.29E-07	9.34E-05
CS000012493	NA*	serpin peptidase inhibitor%2C clade H (heat shock protein 47)%2C member 1%2C (collagen binding protein 1)	-1.43	6.56E-07	1.34E-04
CS000012872	PLA2G6	phospholipase A2 group VI	-0.59	8.99E-05	4.53E-03
CS000005183	NA*	chromosome 12 orf 57	-0.65	8.28E-05	4.26E-03
CS000008789	TCN2	transcobalamin 2	-1.19	6.73E-10	8.81E-07
CS000022298	NA*	arylsulfatase A	-0.92	8.04E-11	1.45E-07
CS000008138	SERF2	small EDRK-rich factor 2	-1.04	2.23E-09	1.76E-06
CS000004171	NA*	ribosomal protein S3	-0.87	9.79E-05	4.72E-03
CS000019557	ILVBL	ilvB acetolactate synthase like	-0.63	9.79E-07	1.72E-04
CS000012912	PIM3	Pim-3 proto-oncogene%2C serine/threonine kinase	-0.64	2.32E-06	3.33E-04
CS000004700	GGA2	golgi associated%2C gamma adaptin ear containing%2C ARF binding protein 2	-0.63	2.29E-04	8.45E-03
CS000007864	KLF10	Kruppel like factor 10	-0.89	4.16E-04	1.28E-02
CS000011685	sall2	spalt like transcription factor 2	-0.85	2.43E-08	1.10E-05
CS000004432	RPL12	ribosomal protein L12	-0.59	1.89E-05	1.47E-03
CS000021495	CYP3A80	cytochrome P450%2C family 3%2C subfamily A%2C polypeptide 80	-0.82	1.22E-05	1.08E-03
CS000006436	CIRBP	cold inducible RNA binding protein	-0.43	9.68E-05	4.69E-03
CS000000139	glb1l	galactosidase beta 1 like	-1.15	3.80E-11	8.00E-08
CS000001540	CTSK	cathepsin K	-1.14	4.58E-06	5.62E-04
CS000020212	CX3CL1	C-X3-C motif chemokine ligand 1	-2.04	5.26E-05	NA
CS000025342	NA*	ATP binding cassette subfamily F member 3	-0.49	3.43E-05	2.30E-03
CS000025328	EEF2	eukaryotic translation elongation factor 2	-0.62	3.58E-06	4.71E-04
CS000007827	SEMA3B	semaphorin 3B	-1.04	4.16E-08	1.55E-05

Gene ID	Gene	Product	log2FC	p-value	Adj. p-value
CS000018177	COL26A1	collagen type XXVI alpha 1 chain	-1.11	6.37E-06	7.32E-04
CS000022208	NA*	low molecular weight neuronal intermediate filament-like	-0.90	7.25E-05	3.90E-03
CS000000127	Bcs1l	BCS1 homolog%2C ubiquinol-cytochrome c reductase complex chaperone	-0.71	7.30E-04	1.81E-02
CS000000860	TRAF4	TNF receptor associated factor 4	-0.98	1.25E-06	2.05E-04
CS000011246	SLC3A2	solute carrier family 3 member 2	-0.68	9.49E-05	4.63E-03
CS000013408	SLC8A2	solute carrier family 8 member A2	-1.49	2.19E-04	8.21E-03
CS000025045	ppdpf	pancreatic progenitor cell differentiation and proliferation factor	-0.84	4.64E-06	5.64E-04
CS000018491	TRIM35	tripartite motif containing 35	-0.52	4.97E-05	2.93E-03
CS000013633	AKR7A2	aldo-keto reductase family 7 member A2	-0.44	4.10E-03	5.75E-02
CS000005722	PLPPR2	phospholipid phosphatase related 2	-1.51	1.21E-03	NA
CS000006303	RPS28	ribosomal protein S28	-0.59	1.52E-04	6.31E-03
CS000005186	P3H3	prolyl 3-hydroxylase 3	-0.65	6.44E-07	1.34E-04
CS000017504	UNC5A	unc-5 netrin receptor A	-1.42	1.42E-05	1.20E-03
CS000001969	PDE4C	phosphodiesterase 4C	-0.75	1.24E-04	5.49E-03
CS000004391	ccdc83	coiled-coil domain containing 83	-1.11	8.79E-05	4.48E-03
CS000008417	TLE3	transducin like enhancer of split 3	-0.59	1.16E-05	1.04E-03
CS000006525	NOP10	NOP10 ribonucleoprotein	-0.73	5.76E-04	1.57E-02
CS000019873	GLTSCR2	glioma tumor suppressor candidate region gene 2	-0.59	1.94E-04	7.61E-03
CS000010756	PPP1R10	protein phosphatase 1 regulatory subunit 10	-0.43	1.31E-03	2.64E-02
CS000009675	PBXIP1	PBX homeobox interacting protein 1	-0.99	3.89E-08	1.51E-05
CS000011487	PHF1	PHD finger protein 1	-0.59	6.71E-04	1.73E-02
CS000024891	EDF1	endothelial differentiation related factor 1	-0.50	1.34E-03	2.68E-02
CS000013297	PCDH10	protocadherin 10	-1.51	5.14E-04	1.46E-02

Gene ID	Gene	Product	log2FC	p-value	Adj. p-value
CS000022997	NA*	transmembrane protein 176A-like	-0.77	2.13E-07	5.97E-05
CS000001204	CTF1	cardiotrophin 1	-0.73	5.96E-05	3.38E-03
CS000003895	NA*	immunoglobulin superfamily member 1-like	-1.51	5.63E-03	NA
CS000001227	C2	complement C2	-1.11	1.19E-05	1.06E-03
CS000007073	NA*	chromosome 15 orf 39	-0.53	6.16E-05	3.44E-03
CS000002371	KHK	ketohexokinase	-1.13	4.28E-04	1.31E-02
CS000008517	TCEA3	transcription elongation factor A3	-0.81	6.42E-06	7.32E-04
CS000014160	CLU	clusterin	-0.68	1.51E-08	7.33E-06
CS000018234	RPL3	ribosomal protein L3	-0.41	1.49E-03	2.88E-02
CS000005212	FAM167B	family with sequence similarity 167 member B	-1.56	2.56E-03	NA
CS000014737	CHCHD6	coiled-coil-helix-coiled-coil-helix domain containing 6	-0.70	2.46E-03	4.12E-02
CS000012900	NEU1	neuraminidase 1	-0.68	1.60E-08	7.49E-06
CS000004733	EIF4A2	eukaryotic translation initiation factor 4A2	-0.59	3.78E-03	5.42E-02
CS000019123	NA*	flap structure-specific endonuclease 1	-0.85	6.67E-06	7.34E-04
CS000024087	NA*	collagen%2C type XIV%2C alpha 1	-1.09	1.41E-03	2.76E-02
CS000025371	NA*	Heat Shock Protein 90 Beta Family Member 1-like	-0.43	4.03E-06	5.20E-04
CS000013073	CCNI	cyclin I	-0.62	6.02E-04	1.60E-02
CS000022683	APBB3	amyloid beta precursor protein binding family B member 3	-0.83	8.83E-04	2.03E-02
CS000024902	FSTL1	follistatin like 1	-0.58	2.03E-04	7.81E-03
CS000009688	SDHC	succinate dehydrogenase complex subunit C	-0.41	4.13E-03	5.77E-02
CS000010306	RPL35	ribosomal protein L35	-0.38	6.79E-04	1.74E-02
CS000002311	MMP19	matrix metallopeptidase 19	-0.80	7.96E-04	1.90E-02
CS000001777	FBXO18	F-box protein%2C helicase%2C 18	-0.48	4.87E-03	6.26E-02
CS000016608	ripk3	receptor interacting serine/threonine kinase 3	-0.50	7.89E-05	4.12E-03
CS000011278	GPR137	G protein-coupled receptor 137	-0.59	4.21E-04	1.29E-02
CS000017494	FAXDC2	fatty acid hydroxylase domain containing 2	-0.72	3.87E-03	5.51E-02

Gene ID	Gene	Product	log2FC	p-value	Adj. p-value
CS000000887	Rilp	Rab interacting lysosomal protein	-0.54	1.23E-03	2.52E-02
CS000007505	CACNA1S	calcium voltage-gated channel subunit alpha1 S	-0.95	1.05E-03	2.26E-02
CS000018849	NA*	globoside alpha-1%2C3-N-acetylgalactosaminyltransferase 1	-1.00	4.58E-03	6.04E-02
CS000001908	WIPF3	WAS/WASL interacting protein family member 3	-1.13	1.31E-05	1.12E-03
CS000016831	STMN3	stathmin 3	-1.06	1.57E-03	2.99E-02
CS000002865	MFAP4	microfibrillar associated protein 4	-1.61	6.22E-05	3.45E-03
CS000011966	BCO2	beta-carotene oxygenase 2	-0.62	8.32E-03	8.64E-02
CS000025205	NA*	collagen%2C type VI%2C alpha 2	-1.41	5.85E-04	1.57E-02
CS000025061	COL1A1	collagen type I alpha 1 chain	-1.08	3.69E-05	2.44E-03
CS000003712	SPARC	secreted protein acidic and cysteine rich	-0.88	1.06E-04	4.93E-03
CS000025386	COL1A2	collagen type I alpha 2 chain	-1.20	2.76E-03	4.49E-02
CS000011375	NA*	reticulon 2 (Z-band associated protein)	-1.86	1.14E-03	2.37E-02
CS000019019	NRK	Nik related kinase	-0.71	6.87E-04	1.74E-02
CS000015423	EEF1D	eukaryotic translation elongation factor 1 delta	-0.34	2.25E-03	3.91E-02
CS000013630	TMCO4	transmembrane and coiled-coil domains 4	-0.64	3.14E-03	4.87E-02
CS000001278	NA*	ribosomal protein S15a	-0.48	8.06E-04	1.91E-02
CS000021156	RPL10A	ribosomal protein L10a	-0.44	9.44E-04	2.12E-02
CS000015866	NA*	eukaryotic translation elongation factor 1 beta 2 L homeolog	-0.46	9.48E-03	9.30E-02
CS000013847	CTSB	cathepsin B	-0.54	5.31E-05	3.05E-03
CS000025195	EIF3F	eukaryotic translation initiation factor 3 subunit F	-0.35	3.97E-03	5.62E-02
CS000010442	NA*	collagen type V alpha 1 chain	-0.56	4.66E-03	6.10E-02
CS000004164	AGAP3	ArfGAP with GTPase domain%2C ankyrin repeat and PH domain 3	-0.51	1.64E-03	3.08E-02
CS000005297	NA*	elongation factor 1-alpha 1-like	-1.07	4.81E-04	1.41E-02
CS000004280	LRRC15	leucine rich repeat containing 15	-0.76	3.29E-04	1.08E-02

Gene ID	Gene	Product	log2FC	p-value	Adj. p-value
CS000005087	CARTPT	cocaine- and amphetamine-regulated transcript protein	-1.85	1.43E-03	NA
CS000005979	KEAP1	kelch like ECH associated protein 1	-0.52	7.39E-03	8.07E-02
CS000025395	NA*	zinc finger protein 251	-0.54	4.32E-03	5.90E-02
CS000003753	NSD1	nuclear receptor binding SET domain protein 1	-0.26	9.65E-03	9.37E-02
CS000019510	TSC22D3	TSC22 domain family member 3	-0.46	6.10E-03	7.17E-02
CS000013443	RPLP1	ribosomal protein lateral stalk subunit P1	-0.39	1.69E-03	3.15E-02
CS000001419	NA*	inositol-trisphosphate 3-kinase B-like	-0.66	8.16E-04	1.92E-02
CS000011226	SNX32	sorting nexin 32	-0.57	1.50E-03	2.89E-02
CS000004853	COL3A1	collagen type III alpha 1 chain	-1.28	1.34E-03	2.68E-02
CS000018200	NA*	eukaryotic translation elongation factor 1 beta 2 L homeolog	-0.55	4.69E-03	6.11E-02
CS000010746	RPLP0	ribosomal protein lateral stalk subunit P0	-0.37	4.73E-03	6.15E-02
CS000012398	LGALS12	galectin 12	-0.76	1.26E-06	2.05E-04
CS000025312	MSMB	microseminoprotein beta	-2.16	9.50E-05	4.63E-03
CS000003599	RPL26L1	ribosomal protein L26 like 1	-0.41	2.42E-03	4.08E-02
CS000011236	EFEMP2	EGF containing fibulin like extracellular matrix protein 2	-0.44	3.41E-03	5.13E-02
CS000018034	RPS8	ribosomal protein S8	-0.33	3.27E-03	4.98E-02
CS000025151	JARID2	jumonji and AT-rich interaction domain containing 2	-1.30	5.48E-17	6.93E-13
CS000019960	PLOD1	procollagen-lysine%2C2-oxoglutarate 5-dioxygenase 1	-0.47	3.76E-03	5.41E-02
CS000001089	BOC	BOC cell adhesion associated%2C oncogene regulated	-0.36	8.33E-03	8.65E-02
CS000002523	PXDNL	peroxidasin like	-1.27	5.03E-03	6.40E-02
CS000017139	RPL7A	ribosomal protein L7a	-0.35	4.63E-03	6.08E-02
CS000007450	MNT	MAX network transcriptional repressor	-0.53	9.18E-04	2.08E-02

Gene ID	Gene	Product	log2FC	p-value	Adj. p-value
CS000020596	SORCS2	sortilin related VPS10 domain containing receptor 2	-0.93	8.26E-03	8.63E-02
CS000022179	NACA	nascent polypeptide-associated complex subunit alpha	-0.37	1.80E-03	3.31E-02
CS000005263	NA*	PR/SET domain 1	-1.58	2.16E-03	NA
CS000018747	DGKA	diacylglycerol kinase alpha	-0.65	3.19E-03	4.92E-02
CS000008590	PTRF	polymerase I and transcript release factor	-0.53	3.08E-03	4.81E-02
CS000022768	RACK1	receptor for activated C kinase 1	-0.42	2.85E-03	4.58E-02
CS000003254	TTLL3	tubulin tyrosine ligase like 3	-0.75	1.09E-03	2.32E-02
CS000025382	PABPC1	polyadenylate-binding protein 1	-0.43	4.33E-03	5.90E-02

*Gene symbol or product name not defined

Table 3: Genes upregulated at MPT *in vivo* and *in vitro*

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000019025	HMGB3	high mobility group box 3	0.89	2.06E-09	1.74E-06
CS000008170	GCA	grancalcin	0.78	2.73E-08	1.17E-05
CS000020229	HEATR3	HEAT repeat containing 3	0.75	8.55E-08	2.84E-05
CS000017806	PSMA5	proteasome subunit alpha 5	0.72	2.28E-05	1.65E-03
CS000000445	ADPGK	ADP dependent glucokinase	0.49	1.15E-05	1.04E-03
CS000006157	PSMC1	proteasome 26S subunit%2C ATPase 1	0.52	1.87E-07	5.51E-05
CS000025408	PSMA3	proteasome subunit alpha 3	0.59	8.26E-06	8.22E-04
CS000010993	PSMD12	proteasome 26S subunit%2C non-ATPase 12	0.65	1.39E-09	1.46E-06
CS000017548	FLVCR1	feline leukemia virus subgroup C cellular receptor 1	0.86	1.36E-06	2.10E-04
CS000009378	TSN	translin	0.67	1.50E-08	7.33E-06
CS000005964	VRK1	vaccinia related kinase 1	1.03	1.16E-08	6.50E-06
CS000023950	FUBP1	far upstream element binding protein 1	0.54	1.03E-08	6.50E-06
CS000022177	PTGES3	prostaglandin E synthase 3	0.71	4.20E-12	2.65E-08
CS000017371	PBRM1	polybromo 1	0.55	2.12E-07	5.97E-05
CS000025411	NA*	high mobility group box 1	0.87	1.08E-11	4.54E-08
CS000005920	TSSC4	tumor suppressing subtransferable candidate 4	0.43	6.47E-04	1.68E-02
CS000001991	AP1M1	adaptor related protein complex 1 mu 1 subunit	0.64	1.44E-06	2.20E-04
CS000025467	PSMA4	proteasome subunit alpha type-4	0.54	1.04E-04	4.92E-03
CS000012933	DLD	dihydrolipoamide dehydrogenase	0.47	4.17E-05	2.62E-03
CS000008806	COPS5	COP9 signalosome subunit 5	0.49	8.16E-05	4.21E-03

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000003143	CHTF8	chromosome transmission fidelity factor 8	1.34	5.36E-04	1.49E-02
CS000005748	RBBP7	RB binding protein 7%2C chromatin remodeling factor	0.54	1.50E-05	1.24E-03
CS000025318	NA*	proteasome subunit alpha 2	0.53	3.91E-05	2.52E-03
CS000006807	SMC3	structural maintenance of chromosomes 3	0.44	8.81E-06	8.57E-04
CS000008290	ACTL6A	actin like 6A	0.64	4.34E-05	2.67E-03
CS000002223	TRA2B	transformer 2 beta homolog	0.75	7.04E-04	1.77E-02
CS000022883	UCHL5	ubiquitin C-terminal hydrolase L5	0.62	6.79E-05	3.70E-03
CS000013717	HNRNPD	heterogeneous nuclear ribonucleoprotein D	0.47	2.56E-05	1.82E-03
CS000006500	RIOK1	RIO kinase 1	0.60	9.22E-07	1.67E-04
CS000003677	HNRNPAB	heterogeneous nuclear ribonucleoprotein A/B	0.64	7.76E-07	1.49E-04
CS000012259	TCP1	t-complex 1	0.62	1.46E-08	7.33E-06
CS000001171	HNRNPU	heterogeneous nuclear ribonucleoprotein U	0.41	1.99E-05	1.52E-03
CS000014244	OPA1	OPA1%2C mitochondrial dynamin like GTPase	0.41	3.41E-04	1.11E-02
CS000002720	RANBP1	RAN binding protein 1	0.92	2.89E-11	8.00E-08
CS000000341	CDK5RAP2	CDK5 regulatory subunit associated protein 2	0.75	4.02E-07	9.08E-05
CS000006284	POLD3	DNA polymerase delta 3%2C accessory subunit	0.67	1.02E-04	4.84E-03
CS000007101	CENPF	centromere protein F	0.86	6.96E-06	7.40E-04
CS000006734	IFRD1	interferon related developmental regulator 1	0.43	4.44E-05	2.71E-03
CS000000505	CCAR1	cell division cycle and apoptosis regulator 1	0.39	4.36E-05	2.67E-03
CS000024754	MIS18BP1	MIS18 binding protein 1	1.06	6.79E-06	7.37E-04

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000025383	YWHAZ	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta	0.47	8.98E-10	1.03E-06
CS000009401	MAP3K7	mitogen-activated protein kinase kinase kinase 7	0.40	5.24E-04	1.46E-02
CS000004541	IMPA1	inositol monophosphatase 1	0.65	4.31E-05	2.67E-03
CS000011006	ATAD5	ATPase family%2C AAA domain containing 5	0.97	1.52E-09	1.48E-06
CS000006786	SFPQ	splicing factor proline and glutamine rich	0.55	2.15E-04	8.07E-03
CS000005278	HMGN2	high mobility group nucleosomal binding domain 2	0.76	4.14E-06	5.28E-04
CS000005207	KHDRBS1	KH RNA binding domain containing%2C signal transduction associated 1	0.53	2.74E-07	6.94E-05
CS000017078	RABGAP1	RAB GTPase activating protein 1	0.34	1.09E-03	2.32E-02
CS000016180	TPP2	tripeptidyl peptidase 2	0.49	3.44E-05	2.30E-03
CS000020659	MPHOSPH6	M-phase phosphoprotein 6	0.57	1.58E-04	6.53E-03
CS000008760	CENPE	centromere protein E	1.31	3.44E-06	4.58E-04
CS000012370	HNRNPM	heterogeneous nuclear ribonucleoprotein M	0.57	1.93E-04	7.57E-03
CS000002594	DPM1	dolichyl-phosphate mannosyltransferase subunit 1%2C catalytic	0.53	2.85E-03	4.58E-02
CS000008906	MMADHC	methylmalonic aciduria and homocystinuria%2C cblD type	0.48	1.75E-04	7.06E-03
CS000020800	SMC6	structural maintenance of chromosomes 6	0.55	4.64E-04	1.39E-02
CS000011589	VDAC2	voltage dependent anion channel 2	0.36	5.21E-04	1.46E-02

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000009277	NUP43	nucleoporin 43	0.64	1.45E-05	1.21E-03
CS000004597	EIF4G2	eukaryotic translation initiation factor 4 gamma 2	0.38	2.66E-05	1.88E-03
CS000015176	GMPS	guanine monophosphate synthase	0.64	2.81E-06	3.82E-04
CS000002842	CWF19L1	CWF19-like 1%2C cell cycle control	0.58	9.00E-04	2.05E-02
CS000007472	DDX52	DExD-box helicase 52	0.58	7.63E-04	1.85E-02
CS000010356	DCAF7	DDB1 and CUL4 associated factor 7	0.38	2.63E-03	4.33E-02
CS000023273	DCTD	dCMP deaminase	0.81	4.22E-05	2.64E-03
CS000007279	NUP133	nucleoporin 133	0.47	6.00E-05	3.39E-03
CS000008307	EIF5A2	eukaryotic translation initiation factor 5A2	0.67	2.20E-05	1.61E-03
CS000014860	TDG	thymine DNA glycosylase	0.58	3.87E-05	2.52E-03
CS000021519	TXNDC5	thioredoxin domain containing 5	0.44	7.78E-06	7.87E-04
CS000005144	RIOK2	RIO kinase 2	0.52	1.86E-04	7.37E-03
CS000005331	LARP7	La ribonucleoprotein domain family member 7	0.67	1.08E-04	5.01E-03
CS000004058	MRS2	MRS2%2C magnesium transporter	0.53	7.45E-04	1.84E-02
CS000025206	NA*	proteasome (prosome%2C macropain) subunit%2C alpha type 6	0.41	4.54E-03	6.03E-02
CS000009028	THOC1	THO complex 1	0.56	1.98E-04	7.67E-03
CS000009859	SRSF2	serine and arginine rich splicing factor 2	0.63	1.31E-03	2.64E-02
CS000006185	NA*	protein Spindly-like	0.78	3.73E-03	5.39E-02
CS000002153	NOP14	NOP14 nucleolar protein	0.46	5.24E-04	1.46E-02
CS000009766	DUSP11	dual specificity phosphatase 11	0.49	1.88E-03	3.41E-02

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000019999	AIDA	axin interactor%2C dorsalization associated	0.44	1.01E-03	2.22E-02
CS000024895	STRAP	serine/threonine kinase receptor associated protein	0.38	7.79E-04	1.87E-02
CS000010787	NUP93	nucleoporin 93	0.49	7.17E-04	1.79E-02
CS000019431	CENPK	centromere protein K	1.91	1.89E-05	1.47E-03
CS000011802	NA	chromosome 1 orf 112	0.57	9.05E-03	9.08E-02
CS000006182	PRDX1	peroxiredoxin 1	0.67	7.29E-05	3.90E-03
CS000013122	KIF20B	kinesin family member 20B	1.28	1.14E-07	3.69E-05
CS000011529	CDR2	cerebellar degeneration-related protein 2%2C 62kDa	0.67	3.08E-03	4.81E-02
CS000008403	AAGAB	alpha- and gamma-adaptin binding protein	0.36	4.09E-03	5.75E-02
CS000010528	BCCIP	BRCA2 and CDKN1A interacting protein	0.45	4.93E-04	1.43E-02
CS000005931	TNNI2	troponin I2%2C fast skeletal type	2.04	1.35E-03	2.68E-02
CS000013871	PLA2G7	phospholipase A2 group VII	0.78	4.55E-03	6.04E-02
CS000014262	PSMD1	proteasome 26S subunit%2C non-ATPase 1	0.43	4.60E-03	6.06E-02
CS000019414	depdc1b	DEP domain containing 1B	1.65	3.11E-05	2.12E-03
CS000025048	PSMA7	proteasome subunit alpha 7	0.41	4.41E-04	1.33E-02
CS000014230	CNOT6L	CCR4-NOT transcription complex subunit 6 like	0.47	1.12E-03	2.36E-02
CS000003633	RAD50	RAD50 double strand break repair protein	0.53	5.81E-04	1.57E-02
CS000013455	TMX1	thioredoxin related transmembrane protein 1	0.42	2.37E-04	8.65E-03
CS000015346	DEPDC1	DEP domain containing 1	1.30	4.10E-05	2.60E-03

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000009895	SLC25A19	solute carrier family 25 member 19	0.50	4.03E-03	5.68E-02
CS000023966	BCL10	B-cell CLL/lymphoma 10	0.69	7.54E-04	1.85E-02
CS000024849	CALM2	calmodulin 2	0.52	8.55E-06	8.44E-04
CS000015905	FANCL	Fanconi anemia complementation group L	0.73	3.70E-04	1.18E-02
CS000003861	PSMD6	proteasome 26S subunit%2C non-ATPase 6	0.48	1.20E-04	5.37E-03
CS000004892	BZW1	basic leucine zipper and W2 domains 1	0.35	1.31E-03	2.64E-02
CS000019794	VRK3	vaccinia related kinase 3	0.63	8.96E-03	9.03E-02
CS000004367	ZUFSP	zinc finger with UFM1 specific peptidase domain	0.70	6.82E-04	1.74E-02
CS000009233	ARFGAP3	ADP ribosylation factor GTPase activating protein 3	0.41	8.92E-04	2.04E-02
CS000007775	CENPP	centromere protein P	1.44	6.82E-05	3.70E-03
CS000004131	RNMT	RNA guanine-7 methyltransferase	0.51	1.35E-03	2.68E-02
CS000005158	WDR36	WD repeat domain 36	0.61	3.75E-09	2.73E-06
CS000003054	GPI	glucose-6-phosphate isomerase	0.35	1.67E-04	6.82E-03
CS000009030	YES1	YES proto-oncogene 1%2C Src family tyrosine kinase	0.46	2.60E-03	4.31E-02
CS000009276	PPIL4	peptidylprolyl isomerase like 4	0.56	3.45E-03	5.15E-02
CS000009972	GPATCH11	G-patch domain containing 11	0.64	2.11E-03	3.70E-02
CS000002599	NA*	TMEM189-UBE2V1 readthrough	0.40	3.25E-04	1.07E-02
CS000020269	UBE3A	ubiquitin protein ligase E3A	0.40	1.31E-03	2.64E-02
CS000025115	DEK	DEK proto-oncogene	0.63	1.31E-06	2.05E-04
CS000017772	API5	apoptosis inhibitor 5	0.38	4.77E-04	1.41E-02
CS000008948	NA*	kynurenine/alpha-aminoadipate aminotransferase%2C mitochondrial	0.72	1.92E-03	3.47E-02

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000000210	ZMPSTE24	zinc metallopeptidase STE24	0.77	1.12E-03	2.36E-02
CS000007520	SRSF3	serine and arginine rich splicing factor 3	0.50	7.87E-06	7.90E-04
CS000020286	PPP2CB	protein phosphatase 2 catalytic subunit beta	0.37	3.44E-04	1.12E-02
CS000016023	NA	NA	1.19	4.53E-03	6.03E-02
CS000008188	STK39	serine/threonine kinase 39	0.58	1.18E-04	5.31E-03
CS000008044	mcm4.S	minichromosome maintenance complex component 4 S homeolog	0.81	2.44E-04	8.83E-03
CS000005905	DDX11	DEAD/H-box helicase 11	0.94	8.65E-03	8.86E-02
CS000008261	NUP35	nucleoporin 35	0.53	3.62E-03	5.30E-02
CS000005071	BTBD1	BTB domain containing 1	0.41	6.59E-03	7.51E-02
CS000018930	DNA2	DNA replication helicase/nuclease 2	1.18	1.27E-05	1.10E-03
CS000009804	PPP2R2A	protein phosphatase 2 regulatory subunit Balpha	0.49	5.60E-04	1.53E-02
CS000025463	HNRNPK	heterogeneous nuclear ribonucleoprotein K	0.29	2.10E-04	8.04E-03
CS000023764	NOL10	nucleolar protein 10	0.47	6.17E-03	7.22E-02
CS000004009	SYNCRIP	synaptotagmin binding cytoplasmic RNA interacting protein	0.51	1.56E-07	4.70E-05
CS000001154	LYAR	Ly1 antibody reactive	0.62	6.63E-06	7.34E-04
CS000003885	PPP4R2	protein phosphatase 4 regulatory subunit 2	0.51	4.30E-04	1.31E-02
CS000002260	KIAA1143	KIAA1143	0.68	2.06E-03	3.66E-02
CS000021460	SKA3	spindle and kinetochore associated complex subunit 3	0.87	3.29E-05	2.22E-03
CS000008448	HACD3	3-hydroxyacyl-CoA dehydratase 3	0.42	2.86E-03	4.58E-02
CS000004767	PUM3	pumilio RNA binding family member 3	0.51	3.90E-04	1.22E-02

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000000624	SLC25A24	solute carrier family 25 member 24	0.37	3.69E-03	5.37E-02
CS000009494	FAM122A	family with sequence similarity 122A	0.37	6.28E-03	7.29E-02
CS000017265	NA*	ferritin light chain-like	0.49	2.22E-04	8.28E-03
CS000000598	WDR62	WD repeat domain 62	1.22	2.15E-05	1.59E-03
CS000002709	RAN	RAN%2C member RAS oncogene family	0.56	1.28E-04	5.58E-03
CS000005502	AKT1	AKT serine/threonine kinase 1	0.45	6.48E-04	1.68E-02
CS000008212	HAT1	histone acetyltransferase 1	0.63	2.15E-04	8.07E-03
CS000008703	CDC45	cell division cycle 45	1.38	3.97E-05	2.54E-03
CS000023592	SSB	Sjogren syndrome antigen B	0.60	6.35E-04	1.67E-02
CS000013060	GRSF1	G-rich RNA sequence binding factor 1	0.41	4.78E-04	1.41E-02
CS000014303	dis3	DIS3 homolog%2C exosome endoribonuclease and 3'-5' exoribonuclease	0.53	5.52E-04	1.52E-02
CS000012145	CELF1	CUGBP%2C Elav-like family member 1	0.43	1.34E-03	2.68E-02
CS000005487	EXO1	exonuclease 1	1.59	4.20E-04	1.29E-02
CS000013316	KIAA0368	KIAA0368	0.26	4.87E-03	6.26E-02
CS000009495	HPRT1	hypoxanthine phosphoribosyltransferase 1	0.47	1.31E-03	2.64E-02
CS000018171	MDH2	malate dehydrogenase 2	0.47	1.02E-03	2.22E-02
CS000006040	MCM8	minichromosome maintenance 8 homologous recombination repair factor	0.72	3.89E-04	1.22E-02
CS000005713	GET4	golgi to ER traffic protein 4	0.62	1.22E-03	2.51E-02
CS000010443	WDR5	WD repeat domain 5	0.43	3.65E-03	5.32E-02
CS000016221	SEPT10	septin 10	0.47	6.22E-04	1.64E-02
CS000004927	NA*	NADH ubiquinone oxidoreductase core subunit S1	0.50	1.51E-03	2.89E-02

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000023675	CTR9	CTR9 homolog%2C Paf1/RNA polymerase II complex component	0.39	3.02E-04	1.01E-02
CS000018740	PSMA1	proteasome subunit alpha 1	0.31	3.65E-03	5.32E-02
CS000009392	BORA	bora%2C aurora kinase A activator	0.95	1.44E-07	4.53E-05
CS000011391	NA*	ATP synthase subunit O%2C mitochondrial	0.45	4.16E-03	5.79E-02
CS000011812	TXNL1	thioredoxin like 1	0.45	4.55E-03	6.04E-02
CS000011005	NA*	NA*	1.18	1.31E-03	2.64E-02
CS000004866	MYO1B	myosin IB	0.44	2.30E-03	3.95E-02
CS000007922	CHEK1	checkpoint kinase 1	0.66	1.59E-03	3.01E-02
CS000016809	MAPRE1	microtubule associated protein RP/EB family member 1	0.40	2.29E-04	8.45E-03
CS000003414	BARD1	BRCA1 associated RING domain 1	0.70	5.55E-04	1.52E-02
CS000007755	NA*	haloacid dehalogenase-like hydrolase domain-containing 5	0.84	1.47E-04	6.16E-03
CS000011009	ZNF207	zinc finger protein 207	0.33	6.42E-03	7.38E-02
CS000016810	DNMT3B	DNA methyltransferase 3 beta	0.63	9.09E-04	2.07E-02
CS000002171	TACC3	transforming acidic coiled-coil containing protein 3	0.87	1.48E-06	2.22E-04
CS000009784	NFU1	NFU1 iron-sulfur cluster scaffold	0.45	8.17E-04	1.92E-02
CS000006259	ATG4C	autophagy related 4C cysteine peptidase	0.56	8.04E-03	8.49E-02
CS000002771	RSC1A1	regulator of solute carriers 1	0.40	8.07E-03	8.50E-02
CS000009493	FAM122B	family with sequence similarity 122B	0.38	1.23E-03	2.52E-02
CS000002216	DBR1	debranching RNA lariats 1	0.41	9.59E-03	9.33E-02
CS000008538	KLHDC2	kelch domain containing 2	0.36	5.99E-03	7.11E-02

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000003654	DDX46	DEAD-box helicase 46	0.40	7.73E-04	1.86E-02
CS000001383	iars2	isoleucyl-tRNA synthetase 2%2C mitochondrial	0.34	3.08E-03	4.81E-02
CS000009335	DARS	aspartyl-tRNA synthetase	0.32	5.41E-03	6.74E-02
CS000007980	ANLN	anillin actin binding protein	1.35	4.10E-03	5.75E-02
CS000008165	PSMD14	proteasome 26S subunit%2C non-ATPase 14	0.48	3.65E-03	5.32E-02
CS000023220	mta3	metastasis associated 1 family member 3	0.42	8.72E-03	8.88E-02
CS000016841	MRGBP	MRG domain binding protein	0.56	7.18E-04	1.79E-02
CS000011143	THUMPD3	THUMP domain containing 3	0.63	1.27E-04	5.58E-03
CS000012861	UBE2N	ubiquitin conjugating enzyme E2 N	0.40	3.03E-03	4.75E-02
CS000014612	CKAP2	cytoskeleton associated protein 2	1.09	7.15E-06	7.48E-04
CS000011491	mcph1	microcephalin 1	0.99	7.66E-05	4.05E-03
CS000008344	SMC4	structural maintenance of chromosomes 4	0.94	2.31E-03	3.95E-02
CS000013889	PARPBP	PARP1 binding protein	1.01	6.47E-03	7.44E-02
CS000020092	NA*	hippocampus abundant transcript-like protein 1	0.38	2.88E-03	4.61E-02
CS000012542	TICRR	TOPBP1 interacting checkpoint and replication regulator	1.03	1.14E-03	2.38E-02
CS000005017	NOL9	nucleolar protein 9	0.43	2.91E-03	4.64E-02
CS000003992	TTK	TTK protein kinase	0.92	1.14E-05	1.04E-03
CS000002656	KIF18A	kinesin family member 18A	0.80	2.07E-03	3.67E-02
CS000020038	NA*	mitochondrial intermediate peptidase	0.43	5.46E-03	6.76E-02
CS000004394	CCDC90B	coiled-coil domain containing 90B	0.45	5.27E-04	1.47E-02
CS000004403	SHCBP1	SHC binding and spindle associated 1	2.10	5.75E-03	6.96E-02

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000013480	TMPO	thymopoietin	0.47	6.87E-04	1.74E-02
CS000024002	TSNAX	translin associated factor X	0.45	5.50E-03	6.77E-02
CS000018426	GSPT1	G1 to S phase transition 1	0.34	3.13E-03	4.87E-02
CS000008065	GNPDA2	glucosamine-6-phosphate deaminase 2	0.54	6.69E-03	7.58E-02
CS000000564	HNRNPL	heterogeneous nuclear ribonucleoprotein L	0.49	6.01E-03	7.11E-02
CS000009438	MCTS1	MCTS1%2C re-initiation and release factor	0.45	8.06E-04	1.91E-02
CS000000195	SLC20A2	solute carrier family 20 member 2	0.50	5.43E-04	1.51E-02
CS000010685	YWHAB	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein beta	0.32	6.44E-04	1.68E-02
CS000017690	NA*	DLG associated protein 5	1.21	5.49E-04	1.52E-02
CS000016391	PSME3	proteasome activator subunit 3	0.50	4.32E-03	5.89E-02
CS000012748	GAS2L3	growth arrest specific 2 like 3	0.56	6.72E-03	7.59E-02
CS000002305	RFC3	replication factor C subunit 3	0.63	2.34E-03	3.99E-02
CS000025460	HMGB2	high mobility group box 2	0.48	1.03E-03	2.24E-02
CS000003473	H2AFZ	H2A histone family member Z	0.72	1.62E-03	3.05E-02
CS000000674	PUS7	pseudouridylate synthase 7 (putative)	0.72	3.85E-04	1.22E-02
CS000006806	SMNDC1	survival motor neuron domain containing 1	0.49	1.44E-03	2.79E-02
CS000018631	SLK	STE20 like kinase	0.42	6.63E-04	1.71E-02
CS000007603	lemd1	LEM domain containing 1	1.10	4.69E-03	6.11E-02
CS000025207	BRMS1L	breast cancer metastasis-suppressor 1-like	0.66	7.62E-03	8.23E-02

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000022400	ARSK	arylsulfatase family member K	0.66	1.23E-03	2.52E-02
CS000012758	ZC3H13	zinc finger CCCH-type containing 13	0.34	4.82E-03	6.23E-02
CS000004561	COMMD2	COMM domain containing 2	0.42	8.47E-03	8.75E-02
CS000006591	NA*	NA*	0.51	8.79E-03	8.92E-02
CS000009408	MMS22L	MMS22 like%2C DNA repair protein	0.76	4.87E-04	1.42E-02
CS000008776	MTFR2	mitochondrial fission regulator 2	0.84	1.69E-04	6.88E-03
CS000011998	SLC35A3	solute carrier family 35 member A3	0.52	6.76E-03	7.62E-02
CS000000412	PGM2	phosphoglucomutase 2	0.33	2.71E-03	4.45E-02
CS000007490	RPS6KB1	ribosomal protein S6 kinase B1	0.39	7.91E-04	1.90E-02
CS000004268	gnai3	G protein subunit alpha i3	0.30	3.87E-03	5.51E-02
CS000005588	NA*	tetratricopeptide repeat domain 39A	0.71	3.66E-04	1.18E-02
CS000008876	PAK1IP1	PAK1 interacting protein 1	0.56	8.76E-04	2.01E-02
CS000001800	NUF2	NUF2%2C NDC80 kinetochore complex component	0.97	3.79E-03	5.43E-02
CS000001707	GNL2	G protein nucleolar 2	0.46	7.23E-04	1.80E-02
CS000013130	lipA	lipase A%2C lysosomal acid type	0.37	4.48E-03	6.00E-02
CS000020721	NIN	ninein	0.40	7.46E-03	8.12E-02
CS000003570	HMMR	hyaluronan mediated motility receptor	1.10	7.61E-04	1.85E-02
CS000008301	ECT2	epithelial cell transforming 2	1.02	6.57E-04	1.70E-02
CS000020039	NA*	mitochondrial intermediate peptidase	0.48	9.10E-03	9.09E-02
CS000013319	SMC2	structural maintenance of chromosomes 2	0.91	3.85E-05	2.52E-03
CS000003108	GINS2	GINS complex subunit 2	0.71	7.17E-04	1.79E-02
CS000008352	GFM1	G elongation factor mitochondrial 1	0.35	7.58E-03	8.21E-02
CS000000370	SET	SET nuclear proto-oncogene	0.38	2.92E-03	4.64E-02

Gene ID	Gene	Product	Log2Fold Change	p-value	Adj. p-value
CS000005330	ZGRF1	zinc finger GRF-type containing 1	1.01	8.47E-03	8.75E-02
CS000019729	SOS1	SOS Ras/Rac guanine nucleotide exchange factor 1	0.32	8.28E-03	8.63E-02
CS000010296	FBXW2	F-box and WD repeat domain containing 2	0.42	1.97E-03	3.52E-02
CS000006872	BUB3	BUB3%2C mitotic checkpoint protein	0.52	2.34E-03	3.99E-02
CS000017689	DLGAP5	DLG associated protein 5	1.27	5.20E-04	1.46E-02
CS000008421	CTDSPL2	CTD small phosphatase like 2	0.44	9.39E-03	9.26E-02
CS000017977	DDX18	DEAD-box helicase 18	0.54	1.60E-05	1.30E-03
CS000009999	HAUS1	HAUS augmin like complex subunit 1	0.50	6.58E-03	7.51E-02
CS000005775	TRIP13	thyroid hormone receptor interactor 13	0.84	1.68E-03	3.14E-02
CS000005953	DDX24	DEAD-box helicase 24	0.47	6.81E-04	1.74E-02
CS000013221	DHX29	DExH-box helicase 29	0.38	4.65E-03	6.09E-02
CS000004226	NUP58	nucleoporin 58	0.57	9.31E-04	2.10E-02
CS000022817	NA*	protein regulator of cytokinesis 1	0.97	4.31E-03	5.89E-02
CS000007153	METAP2	methionyl aminopeptidase 2	0.36	7.99E-03	8.46E-02
CS000003399	COPS8	COP9 signalosome subunit 8	0.34	7.13E-03	7.86E-02
CS000008383	EIF2A	eukaryotic translation initiation factor 2A	0.38	9.51E-03	9.30E-02
CS000003859	THOC7	THO complex 7	0.47	1.05E-03	2.27E-02
CS000000456	RPAP3	RNA polymerase II associated protein 3	0.44	2.06E-03	3.66E-02

*Gene symbol or product name not defined

REFERENCES

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114-2120. doi:10.1093/bioinformatics/btu170

Bott, R. C., Clopton, D. T., & Cupp, A. S. (2008). A proposed role for VEGF isoforms in sex-specific vasculature development in the gonad. *Reprod Domest Anim, 43 Suppl 2*, 310-316. doi:10.1111/j.1439-0531.2008.01179.x

Britt, K. L., Kerr, J., O'Donnell, L., Jones, M. E., Drummond, A. E., Davis, S. R., . . . Findlay, J. K. (2002). Estrogen regulates development of the somatic cell phenotype in the eutherian ovary. *FASEB J, 16*(11), 1389-1397. doi:10.1096/fj.01-0992com

Chuang, L. S., Ito, K., & Ito, Y. (2013). RUNX family: Regulation and diversification of roles through interacting proteins. *Int J Cancer, 132*(6), 1260-1271. doi:10.1002/ijc.27964

Couse, J. F., Hewitt, S. C., Bunch, D. O., Sar, M., Walker, V. R., Davis, B. J., & Korach, K. S. (1999). Postnatal sex reversal of the ovaries in mice lacking estrogen receptors alpha and beta. *Science, 286*(5448), 2328-2331. doi:10.1126/science.286.5448.2328

Crews, D. (1996). Temperature-dependent sex determination: the interplay of steroid hormones and temperature. *Zoolog Sci, 13*(1), 1-13. doi:10.2108/zsj.13.1

Cupp, A. S., Kim, G., & Skinner, M. K. (1999). Expression and action of transforming growth factor beta (TGFbeta1, TGFbeta2, and TGFbeta3) during embryonic rat testis development. *Biol Reprod, 60*(6), 1304-1313. doi:10.1095/biolreprod60.6.1304

Das, D., Singh, S. K., Bierstedt, J., Erickson, A., Galli, G. L. J., Crossley, D. A., & Rhen, T. (2020). Draft Genome of the Common Snapping Turtle, Chelydra serpentina, a Model for Phenotypic Plasticity in Reptiles. *G3 (Bethesda)*. doi:10.1534/g3.120.401440

Deveson, I. W., Holleley, C. E., Blackburn, J., Marshall Graves, J. A., Mattick, J. S., Waters, P. D., & Georges, A. (2017). Differential intron retention in Jumonji chromatin modifier genes is implicated in reptile temperature-dependent sex determination. *Sci Adv, 3*(6), e1700731. doi:10.1126/sciadv.1700731

Duffy, J. B., & Gergen, J. P. (1991). The Drosophila segmentation gene runt acts as a position-specific numerator element necessary for the uniform expression of the sex-determining gene Sex-lethal. *Genes Dev, 5*(12A), 2176-2187. doi:10.1101/gad.5.12a.2176

Ehrlund, A., Jonsson, P., Vedin, L. L., Williams, C., Gustafsson, J. A., & Treuter, E. (2012). Knockdown of SF-1 and RNF31 affects components of steroidogenesis, TGFbeta, and Wnt/beta-catenin signaling in adrenocortical carcinoma cells. *PLoS One, 7*(3), e32080. doi:10.1371/journal.pone.0032080

Elbrecht, A., & Smith, R. G. (1992). Aromatase enzyme activity and sex determination in chickens. *Science, 255*(5043), 467-470. doi:10.1126/science.1734525

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol, 20*(1), 238. doi:10.1186/s13059-019-1832-y

Estermann, M. A., Williams, S., Hirst, C. E., Roly, Z. Y., Serralbo, O., Adhikari, D., . . . Smith, C. A. (2020). Insights into Gonadal Sex Differentiation Provided by Single-Cell Transcriptomics in the Chicken Embryo. *Cell Rep, 31*(1), 107491. doi:10.1016/j.celrep.2020.03.055

Ewert, M. A., Lang, J. W., & Nelson, C. E. (2005). Geographic variation in the pattern of temperature-dependent sex determination in the American snapping turtle (Chelydra serpentina). *Journal of Zoology*(265), 81-95. doi:10.1017/S0952836904006120

Feng, X. H., & Derynck, R. (2005). Specificity and versatility in tgf-beta signaling through Smads. *Annu Rev Cell Dev Biol, 21*, 659-693. doi:10.1146/annurev.cellbio.21.022404.142018

Fenouille, N., Tichet, M., Dufies, M., Pottier, A., Mogha, A., Soo, J. K., . . . Tartare-Deckert, S. (2012). The epithelial-mesenchymal transition (EMT) regulatory factor SLUG (SNAI2) is a downstream target of SPARC and AKT in promoting melanoma cell invasion. *PLoS One, 7*(7), e40378. doi:10.1371/journal.pone.0040378

Gillespie, R. F., & Gudas, L. J. (2007). Retinoid regulated association of transcriptional co-regulators and the polycomb group protein SUZ12 with the retinoic acid response elements of Hoxa1, RARbeta(2), and Cyp26A1 in F9 embryonal carcinoma cells. *J Mol Biol, 372*(2), 298-316. doi:10.1016/j.jmb.2007.06.079

Hartl, F. U. (1996). Molecular chaperones in cellular protein folding. *Nature, 381*(6583), 571-579. doi:10.1038/381571a0

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., . . . Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell, 38*(4), 576-589. doi:10.1016/j.molcel.2010.05.004

Honda, T., Yamamoto, H., Ishii, A., & Inui, M. (2010). PDZRN3 negatively regulates BMP-2-induced osteoblast differentiation through inhibition of Wnt signaling. *Mol Biol Cell, 21*(18), 3269-3277. doi:10.1091/mbc.E10-02-0117

Hu, Y., Xing, J., Chen, L., Guo, X., Du, Y., Zhao, C., . . . Sha, J. (2008). RGS22, a novel testis-specific regulator of G-protein signaling involved in human and mouse spermiogenesis along with GNA12/13 subunits. *Biol Reprod, 79*(6), 1021-1029. doi:10.1095/biolreprod.107.067504

Huang, C. C., & Yao, H. H. (2010). Diverse functions of Hedgehog signaling in formation and physiology of steroidogenic organs. *Mol Reprod Dev, 77*(6), 489-496. doi:10.1002/mrd.21174

Imamichi, Y., Mizutani, T., Ju, Y., Matsumura, T., Kawabe, S., Kanno, M., . . . Miyamoto, K. (2014). Transcriptional regulation of human ferredoxin reductase through an intronic enhancer in steroidogenic cells. *Biochim Biophys Acta, 1839*(1), 33-42. doi:10.1016/j.bbagrm.2013.11.005

Jiao, L., & Liu, X. (2015). Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science, 350*(6258), aac4383. doi:10.1126/science.aac4383

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., . . . Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics, 30*(9), 1236-1240. doi:10.1093/bioinformatics/btu031

Kashimada, K., Pelosi, E., Chen, H., Schlessinger, D., Wilhelm, D., & Koopman, P. (2011). FOXL2 and BMP2 act cooperatively to regulate follistatin gene expression during ovarian development. *Endocrinology, 152*(1), 272-280. doi:10.1210/en.2010-0636

Kashyap, V., & Gudas, L. J. (2010). Epigenetic regulatory mechanisms distinguish retinoic acid-mediated transcriptional responses in stem cells and fibroblasts. *J Biol Chem, 285*(19), 14534-14548. doi:10.1074/jbc.M110.115345

Kasinath, V., Faini, M., Poepsel, S., Reif, D., Feng, X. A., Stjepanovic, G., . . . Nogales, E. (2018). Structures of human PRC2 with its cofactors AEBP2 and JARID2. *Science, 359*(6378), 940-944. doi:10.1126/science.aar5700

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol, 37*(8), 907-915. doi:10.1038/s41587-019-0201-4

Klopfenstein, D. V., Zhang, L., Pedersen, B. S., Ramirez, F., Warwick Vesztrocy, A., Naldi, A., . . . Tang, H. (2018). GOATOOLS: A Python library for Gene Ontology analyses. *Sci Rep, 8*(1), 10872. doi:10.1038/s41598-018-28948-z

Kohno, S., Katsu, Y., Urushitani, H., Ohta, Y., Iguchi, T., & Guillette, L. J., Jr. (2010). Potential contributions of heat shock proteins to temperature-dependent sex determination in the American alligator. *Sex Dev, 4*(1-2), 73-87. doi:10.1159/000260374

Lai, Y. S., Chang, C. W., Pawlik, K. M., Zhou, D., Renfrow, M. B., & Townes, T. M. (2012). SRY (sex determining region Y)-box2 (Sox2)/poly ADP-ribose polymerase 1 (Parp1) complexes regulate pluripotency. *Proc Natl Acad Sci U S A, 109*(10), 3772-3777. doi:10.1073/pnas.1108595109

Landeira, D., & Fisher, A. G. (2011). Inactive yet indispensable: the tale of Jarid2. *Trends Cell Biol, 21*(2), 74-80. doi:10.1016/j.tcb.2010.10.004

Leers-Sucheta, S., Morohashi, K., Mason, J. I., & Melner, M. H. (1997). Synergistic activation of the human type II 3beta-hydroxysteroid dehydrogenase/delta5-delta4 isomerase promoter by the transcription factor steroidogenic factor-1/adrenal 4-binding protein and phorbol ester. *J Biol Chem, 272*(12), 7960-7967. doi:10.1074/jbc.272.12.7960

Li, G., Margueron, R., Ku, M., Chambon, P., Bernstein, B. E., & Reinberg, D. (2010). Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev, 24*(4), 368-380. doi:10.1101/gad.1886410

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics, 34*(18), 3094-3100. doi:10.1093/bioinformatics/bty191

Li, Y., Oh, H. J., & Lau, Y. F. (2006). The poly(ADP-ribose) polymerase 1 interacts with Sry and modulates its biological functions. *Mol Cell Endocrinol, 257-258*, 35-46. doi:10.1016/j.mce.2006.06.008

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics, 30*(7), 923-930. doi:10.1093/bioinformatics/btt656

Liu, C., Peng, J., Matzuk, M. M., & Yao, H. H. (2015). Lineage specification of ovarian theca cells requires multicellular interactions via oocyte and granulosa cells. *Nat Commun, 6*, 6934. doi:10.1038/ncomms7934

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol, 15*(12), 550. doi:10.1186/s13059-014-0550-8

Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., 3rd, . . . Gamblin, S. J. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature, 461*(7265), 762-767. doi:10.1038/nature08398

Margueron, R., & Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature, 469*(7330), 343-349. doi:10.1038/nature09784

Matson, C. K., Murphy, M. W., Sarver, A. L., Griswold, M. D., Bardwell, V. J., & Zarkower, D. (2011). DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature, 476*(7358), 101-104. doi:10.1038/nature10239

Matsuda, M., & Hirata, M. (2017). Phospholipase C-related but catalytically inactive proteins regulate ovarian follicle development. *J Biol Chem, 292*(20), 8369-8380. doi:10.1074/jbc.M116.759928

Memon, M. A., Anway, M. D., Covert, T. R., Uzumcu, M., & Skinner, M. K. (2008). Transforming growth factor beta (TGFbeta1, TGFbeta2 and TGFbeta3) null-mutant phenotypes in embryonic gonadal development. *Mol Cell Endocrinol, 294*(1-2), 70-80. doi:10.1016/j.mce.2008.08.017

Minkina, A., Matson, C. K., Lindeman, R. E., Ghyselinck, N. B., Bardwell, V. J., & Zarkower, D. (2014). DMRT1 protects male gonadal cells from retinoid-dependent sexual transdifferentiation. *Dev Cell, 29*(5), 511-520. doi:10.1016/j.devcel.2014.04.017

Mistry, D. S., Chen, Y., Wang, Y., Zhang, K., & Sen, G. L. (2014). SNAI2 controls the undifferentiated state of human epidermal progenitor cells. *Stem Cells, 32*(12), 3209-3218. doi:10.1002/stem.1809

Morrish, B. C., & Sinclair, A. H. (2002). Vertebrate sex determination: many means to an end. *Reproduction, 124*(4), 447-457. doi:10.1530/rep.0.1240447

Nakamoto, M., Fukasawa, M., Tanaka, S., Shimamori, K., Suzuki, A., Matsuda, M., . . . Shibata, N. (2012). Expression of 3beta-hydroxysteroid dehydrogenase (hsd3b), star and ad4bp/sf-1 during gonadal development in medaka (Oryzias latipes). *Gen Comp Endocrinol, 176*(2), 222-230. doi:10.1016/j.ygcen.2012.01.019

Nawaz, Z., Lonard, D. M., Smith, C. L., Lev-Lehman, E., Tsai, S. Y., Tsai, M. J., & O'Malley, B. W. (1999). The Angelman syndrome-associated protein, E6-AP, is a coactivator for the nuclear hormone receptor superfamily. *Mol Cell Biol, 19*(2), 1182-1189. doi:10.1128/mcb.19.2.1182

Nicol, B., Grimm, S. A., Chalmel, F., Lecluze, E., Pannetier, M., Pailhoux, E., . . . Yao, H. H. (2019). RUNX1 maintains the identity of the fetal ovary through an interplay with FOXL2. *Nat Commun, 10*(1), 5116. doi:10.1038/s41467-019-13060-1

Nieto, M. A. (2002). The snail superfamily of zinc-finger transcription factors. *Nat Rev Mol Cell Biol, 3*(3), 155-166. doi:10.1038/nrm757

Oksuz, O., Narendra, V., Lee, C. H., Descostes, N., LeRoy, G., Raviram, R., . . . Reinberg, D. (2018). Capturing the Onset of PRC2-Mediated Repressive Domain Formation. *Mol Cell, 70*(6), 1149-1162 e1145. doi:10.1016/j.molcel.2018.05.023

Otsuka, F., Moore, R. K., & Shimasaki, S. (2001). Biological function and cellular mechanism of bone morphogenetic protein-6 in the ovary. *J Biol Chem, 276*(35), 32889-32895. doi:10.1074/jbc.M103212200

Pangas, S. A., Li, X., Umans, L., Zwijsen, A., Huylebroeck, D., Gutierrez, C., . . . Matzuk, M. M. (2008). Conditional deletion of Smad1 and Smad5 in somatic cells of male and female gonads leads to metastatic tumor development in mice. *Mol Cell Biol, 28*(1), 248-257. doi:10.1128/MCB.01404-07

Peng, J. C., Valouev, A., Swigut, T., Zhang, J., Zhao, Y., Sidow, A., & Wysocka, J. (2009). Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell, 139*(7), 1290-1302. doi:10.1016/j.cell.2009.12.002

Pratt, W. B. (1997). The role of the hsp90-based chaperone system in signal transduction by nuclear receptors and receptors signaling via MAP kinase. *Annu Rev Pharmacol Toxicol, 37*, 297-326. doi:10.1146/annurev.pharmtox.37.1.297

Radhakrishnan, S., Literman, R., Neuwald, J., Severin, A., & Valenzuela, N. (2017). Transcriptomic responses to environmental temperature by turtles with temperature-

dependent and genotypic sex determination assessed by RNAseq inform the genetic architecture of embryonic gonadal development. *PLoS One, 12*(3), e0172044. doi:10.1371/journal.pone.0172044

Ramsey, M., & Crews, D. (2009). Steroid signaling and temperature-dependent sex determination-Reviewing the evidence for early action of estrogen during ovarian determination in turtles. *Semin Cell Dev Biol, 20*(3), 283-292. doi:10.1016/j.semcdb.2008.10.004

Rhen, T., & Lang, J. W. (1998). Among-Family Variation for Environmental Sex Determination in Reptiles. *Evolution, 52*(5), 1514-1520. doi:10.1111/j.1558-5646.1998.tb02034.x

Rhen, T., Metzger, K., Schroeder, A., & Woodward, R. (2007). Expression of putative sex-determining genes during the thermosensitive period of gonad development in the snapping turtle, Chelydra serpentina. *Sex Dev, 1*(4), 255-270. doi:10.1159/000104775

Rhen, T., & Schroeder, A. (2010). Molecular mechanisms of sex determination in reptiles. *Sex Dev, 4*(1-2), 16-28. doi:10.1159/000282495

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., . . . Chang, H. Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell, 129*(7), 1311-1323. doi:10.1016/j.cell.2007.05.022

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res, 43*(7), e47. doi:10.1093/nar/gkv007

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics, 26*(1), 139-140. doi:10.1093/bioinformatics/btp616

Roush, D., & Rhen, T. (2018). Developmental plasticity in reptiles: Critical evaluation of the evidence for genetic and maternal effects on temperature-dependent sex determination. *J Exp Zool A Ecol Integr Physiol, 329*(6-7), 287-297. doi:10.1002/jez.2194

Rullinkov, G., Tamme, R., Sarapuu, A., Lauren, J., Sepp, M., Palm, K., & Timmusk, T. (2009). Neuralized-2: expression in human and rodents and interaction with Delta-like ligands. *Biochem Biophys Res Commun, 389*(3), 420-425. doi:10.1016/j.bbrc.2009.08.147

Sanulli, S., Justin, N., Teissandier, A., Ancelin, K., Portoso, M., Caron, M., . . . Margueron, R. (2015). Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation. *Mol Cell, 57*(5), 769-783. doi:10.1016/j.molcel.2014.12.020

Sasai, N., Kato, Y., Kimura, G., Takeuchi, T., & Yamaguchi, M. (2007). The Drosophila jumonji gene encodes a JmjC-containing nuclear protein that is required for metamorphosis. *FEBS J, 274*(23), 6139-6151. doi:10.1111/j.1742-4658.2007.06135.x

Sheftel, A. D., Stehling, O., Pierik, A. J., Elsasser, H. P., Muhlenhoff, U., Webert, H., . . . Lill, R. (2010). Humans possess two mitochondrial ferredoxins, Fdx1 and Fdx2, with distinct roles in steroidogenesis, heme, and Fe/S cluster biosynthesis. *Proc Natl Acad Sci U S A, 107*(26), 11775-11780. doi:10.1073/pnas.1004250107

Shoemaker, C., Ramsey, M., Queen, J., & Crews, D. (2007). Expression of Sox9, Mis, and Dmrt1 in the gonad of a species with temperature-dependent sex determination. *Dev Dyn, 236*(4), 1055-1063. doi:10.1002/dvdy.21096

Son, J., Shen, S. S., Margueron, R., & Reinberg, D. (2013). Nucleosome-binding activities within JARID2 and EZH1 regulate the function of PRC2 on chromatin. *Genes Dev, 27*(24), 2663-2677. doi:10.1101/gad.225888.113

Song, R., Koo, B. K., Yoon, K. J., Yoon, M. J., Yoo, K. W., Kim, H. T., . . . Kong, Y. Y. (2006). Neuralized-2 regulates a Notch ligand in cooperation with Mind bomb-1. *J Biol Chem, 281*(47), 36391-36400. doi:10.1074/jbc.M606601200

Stevant, I., Kuhne, F., Greenfield, A., Chaboissier, M. C., Dermitzakis, E. T., & Nef, S. (2019). Dissecting Cell Lineage Specification and Sex Fate Determination in Gonadal Somatic Cells Using Single-Cell Transcriptomics. *Cell Rep, 26*(12), 3272-3283 e3273. doi:10.1016/j.celrep.2019.02.069

Stevant, I., Neirijnck, Y., Borel, C., Escoffier, J., Smith, L. B., Antonarakis, S. E., . . . Nef, S. (2018). Deciphering Cell Lineage Specification during Male Sex Determination with Single-Cell RNA Sequencing. *Cell Rep, 22*(6), 1589-1599. doi:10.1016/j.celrep.2018.01.043

Sun, X., Chen, H., Deng, Z., Hu, B., Luo, H., Zeng, X., . . . Ma, L. (2015). The Warsaw breakage syndrome-related protein DDX11 is required for ribosomal RNA synthesis and embryonic development. *Hum Mol Genet, 24*(17), 4901-4915. doi:10.1093/hmg/ddv213

Tang, A. D., Soulette, C. M., van Baren, M. J., Hart, K., Hrabeta-Robinson, E., Wu, C. J., & Brooks, A. N. (2020). Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. *Nat Commun, 11*(1), 1438. doi:10.1038/s41467-020-15171-6

Tsutsumi, K., Matsuda, M., Kotani, M., Mizokami, A., Murakami, A., Takahashi, I., . . . Hirata, M. (2011). Involvement of PRIP, phospholipase C-related, but catalytically inactive protein, in bone formation. *J Biol Chem, 286*(35), 31032-31042. doi:10.1074/jbc.M111.235903

Vrenken, K. S., Vervoort, B. M. T., van Ingen Schenau, D. S., Derks, Y. H. W., van Emst, L., Grytsenko, P. G., . . . van Leeuwen, F. N. (2020). The transcriptional repressor SNAI2 impairs neuroblastoma differentiation and inhibits response to retinoic acid therapy. *Biochim Biophys Acta Mol Basis Dis, 1866*(3), 165644. doi:10.1016/j.bbadis.2019.165644

Wang, T., Danielson, P. D., Li, B. Y., Shah, P. C., Kim, S. D., & Donahoe, P. K. (1996). The p21(RAS) farnesyltransferase alpha subunit in TGF-beta and activin signaling. *Science, 271*(5252), 1120-1122. doi:10.1126/science.271.5252.1120

Warr, N., Carre, G. A., Siggers, P., Faleato, J. V., Brixey, R., Pope, M., . . . Greenfield, A. (2012). Gadd45gamma and Map3k4 interactions regulate mouse testis determination via p38 MAPK-mediated control of Sry expression. *Dev Cell, 23*(5), 1020-1031. doi:10.1016/j.devcel.2012.09.016

Wen, J., Zhu, H., Murakami, S., Leung, P. C., & MacCalman, C. D. (2006). Regulation of A Disintegrin And Metalloproteinase with ThromboSpondin repeats-1 expression in human endometrial stromal cells by gonadal steroids involves progestins, androgens, and estrogens. *J Clin Endocrinol Metab, 91*(12), 4825-4835. doi:10.1210/jc.2006-1567

Wijgerde, M., Ooms, M., Hoogerbrugge, J. W., & Grootegoed, J. A. (2005). Hedgehog signaling in mouse ovary: Indian hedgehog and desert hedgehog from granulosa cells induce target gene expression in developing theca cells. *Endocrinology, 146*(8), 3558-3566. doi:10.1210/en.2005-0311

Wong, J. J., Au, A. Y., Ritchie, W., & Rasko, J. E. (2016). Intron retention in mRNA: No longer nonsense: Known and putative roles of intron retention in normal and disease biology. *Bioessays, 38*(1), 41-49. doi:10.1002/bies.201500117

Wu, Y. Y., Peck, K., Chang, Y. L., Pan, S. H., Cheng, Y. F., Lin, J. C., . . . Yang, P. C. (2011). SCUBE3 is an endogenous TGF-beta receptor ligand and regulates the epithelial-

mesenchymal transition in lung cancer. *Oncogene, 30*(34), 3682-3693. doi:10.1038/onc.2011.85

Yao, H. H., Whoriskey, W., & Capel, B. (2002). Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. *Genes Dev, 16*(11), 1433-1440. doi:10.1101/gad.981202

Yatsu, R., Miyagawa, S., Kohno, S., Parrott, B. B., Yamaguchi, K., Ogino, Y., . . . Iguchi, T. (2016). RNA-seq analysis of the gonadal transcriptome during Alligator mississippiensis temperature-dependent sex determination and differentiation. *BMC Genomics, 17*, 77. doi:10.1186/s12864-016-2396-9