Regulation Of *Fshr* And *SF-1* In The Hypothalamus-Pituitary-Gonadal (HPG) Axis.

By

Jitu W. George

Submitted to the graduate degree program in Molecular and Integrative Physiology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dissertation Committee

_____

Leslie L. Heckert, PhD, Chair

_____

Vargheese M.Chennathukuzhi, PhD

_____

Lane K. Christenson, PhD

_____

Patrick E. Fields, Ph.D.

_____

Michael W. Wolfe, Ph.D.

Date Defended: July 10, 2013

The Dissertation Committee for Jitu W. George

certifies that this is the approved version of the following dissertation:

Regulation Of *Fshr* And *SF-1* In The Hypothalamus-Pituitary-Gonadal (HPG) Axis.

_____

Leslie L. Heckert, PhD, Chair

Date approved: July 17, 2013

Abstract

Mammalian reproduction is highly dependent on the delicate balance of signals within the hypothalamic-pituitary-gonadal (HPG) axis that maintains proper endocrine environment. One of the key signals is the pituitary glycoprotein follicle-stimulating hormone (FSH) which is critical for gonadal development and fertility. FSH acts via FSHR, a G protein-coupled receptor present on a few distinct cell types, prominently on the Sertoli cells of the testis and granulosa cells of the ovary Prior studies have identified several key regulatory elements required for *Fshr* transcription but a clear understanding of what controls its limited cell specific expression remains elusive. Comparative genomics identified a number of evolutionary conserved regions (ECR) distal to the proximal promoter, indicating these regions might harbor regulatory elements. One such distal regulatory element known to play a role in transcriptional regulation is the "multivalent" protein CCCTC-binding factor (CTCF). Computational analysis of ECRs identified multiple CTCF binding sites in the intergenic regions of the *Fshr* locus and depletion of CTCF in rat granulosa cells led to a two-fold increase in *Fshr* mRNA. These data indicate that CTCF either by itself or in conjunction with other protein complexes might play a role in transcriptional regulation of the *Fshr* gene.

Another important component of the HPG axis is the transcription factor, Steroidogenic factor-1 (SF-1) transcribed from the *Nr5a1* gene (also known as *Ftz-F1* or *Ad4bp*) which plays a pivotal role in adrenal and gonadal development and regulates genes at all levels of the axis. Comparative genomics identified two conserved sequences associated with *Ftz-F1*. One was found to encode a long non-coding RNA and the other a regulatory element important for SF-1 expression in the pituitary. A long non-coding RNA named *Fast* or *Ftz-F1* associated transcript was found transcribed in the opposite direction to *SF-1*, co-expressed, and co-regulated with *SF-*

1 and hormonally regulated. Knockdown and over-expression of the *Fast* transcript in MA-10, mouse Leydig cell line, did not alter the transcription or translation of SF-1. Long non-coding RNA have been focus of intense research and have been shown to play diverse role in various physiological and pathological processes. Considering their potential for transcriptional regulation, the identified long non-coding RNA, *Fast*, might play potentially important role in endocrine development and homeostasis.

Comparative sequence analysis identified an evolutionary conserved region, ECR3, approximately 4kb away from the transcriptional start site of SF-1. Transient transfection data revealed ECR3, upregulated SF-1 transcription in alpha T3 (gonadotrope cell line) and Y-1 (adrenal) cells, but downregulated SF-1 in MA-10 (Leydig cell line), MSC-1 (Sertoli cell line) and primary rat peritubular Myoid and Sertoli cells. This region was found to contain an E-box and was bound by upstream stimulatory factors 1 and 2 (USF1 and USF2) and a basic-helix-loop-helix (bHLH) protein E2A. In particular, co-transfection studies identified Inhibitors of DNA binding, Id2 and Id3 to downregulate SF-1 transcription.

In summary, this dissertation looks at the regulation of *Fshr* and SF-1 in the HPG axis and identifies the importance of distal regulatory sequences and non-coding transcripts and its role in gene regulation.

Acknowledgments

This dissertation is the result of a challenging journey, which would not have been possible without the help, guidance, and support of a great number of people. First, and foremost, I am greatly indebted to Dr. Leslie Heckert for accepting me as her graduate student. I am thankful for her undying patience and support and for teaching me to think, plan, and perform as a scientist. I have been fortunate to have learned a lot under her and really wish I could have learned more. I am truly blessed to have had you as my advisor.

I am grateful to all my committee members, Dr. Vargheese Chennathukuzhi, Dr. Lane Christenson, Dr. Michael Wolfe and Dr. Patrick Fields, for your input and guidance, has helped me develop as a scientist. Each of you made time from your busy schedule to counsel and advise me on my scientific research and career. No one could ask for a more dedicated and helpful committee and I thank you for that.

Special thanks go to Daren Rice for teaching me a number of molecular biology techniques. His thorough knowledge of molecular biology is truly admirable. During my time at the lab, he has helped me navigate through a number of technical problems that arose in molecular biology experiments and in my 13-year-old car. Not only was he a wonderful lab manager but also a great friend. I will truly miss the annual zoo outings that were made more enjoyable with the presence of his four wonderful kids, Addy, Zade, Levi, and Finn.

I also thank fellow lab members, Dr. Tatiana Karpova, Dr. Shixin Tao, and my fellow ex-graduate student, Dr. Valentine Agbor, from whom I have learned so much. It was a great pleasure to have worked with you all. I am also grateful to Dr. Ravichandiran Kumarasamy and Dr. Brian Hermann whose research work has helped form this thesis.

I owe a lot to Dr. Paul Cheney, who 5 years ago guided me to Leslie, when I was looking for a graduate student position. He has always been supportive of Physiology students and I thank him and the Molecular and Integrative Physiology Department for supporting my funding and research for over a year. I do realize that this has been very tough especially with the budget cuts and the current NIH funding and I am very appreciative of all the support.

I have been blessed to be part of the Indian community at the KUMC, which comprises of a great number of graduate students and post-doctorates, whom I have been fortunate to know. I am truly thankful to a number of my friends who brought back gifts and packages from India and for their companionship that helped compensate for the five and a half long years I have been away from home.

As I come to the end of my acknowledgments, I am thankful to my parents, for their love, support and all the sacrifices they have made throughout their life so that I could obtain a better education and life.

Finally my dear wife, Ani, I am thankful for your undying love, faith, patience, and extraordinary strength, which kept me going. I cannot even begin to comprehend how I would have made this journey without you.

Table of Contents

List of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Part 1: Reproduction and transcription regulation

Chapter 1: Introduction

The hypothalamus-pituitary-gonad axis and the gonadotropin hormones

In mammals, fertility depends on precise hormonal regulation of a functional feedback loop that involves the hypothalamus, pituitary, and gonads (HPG), responsible for production of gametes and hormones, and influences the overall reproduction system. The HPG axis was proposed in 1930 by Carl R. Moore and Dorothy Rice, who demonstrated that hormones produced in the testis and ovary feedback onto the pituitary reduced the production of gonadotropins (Moore and Price 1930, Moore and Price 1932). Shortly thereafter, Walter Hohlweg and Karl Junkmann deduced that there was an interdependent relationship between the gonads, the central nervous system, and the pituitary (Hohlweg and Junkmann 1932, Hohlweg 1975). The central thought was that hormones secreted by the gonads, via their effect on the central nervous system; control the secretion of gonadotropins from the pituitary, which in turn regulate hormone production in the gonads. Geoffrey Harris then demonstrated that the hypothalamus was the main control site, by inducing ovulation in rabbits by electric stimulation of either the hypothalamus or the pituitary, thus identifying the components of the HPG axis (Harris 1937).

The intricate relationship starts with the pulsatile secretion of the hypothalamic gonadotropin-releasing hormone (GnRH), leading to the secretion of bioactive gonadotropins, Follicle stimulating hormone (FSH) and luteinizing hormone (LH), which travel through the blood stream and act on specific cell surface receptors on the gonads, leading to hormone production. A complex feedback loop regulates the hormone levels, which is necessary to initiate and support gametogenesis, steriodogenesis, and ovulation. LH and FSH are integral parts of the neural and endocrine interchange between the hypothalamus, pituitary, and gonads that controls

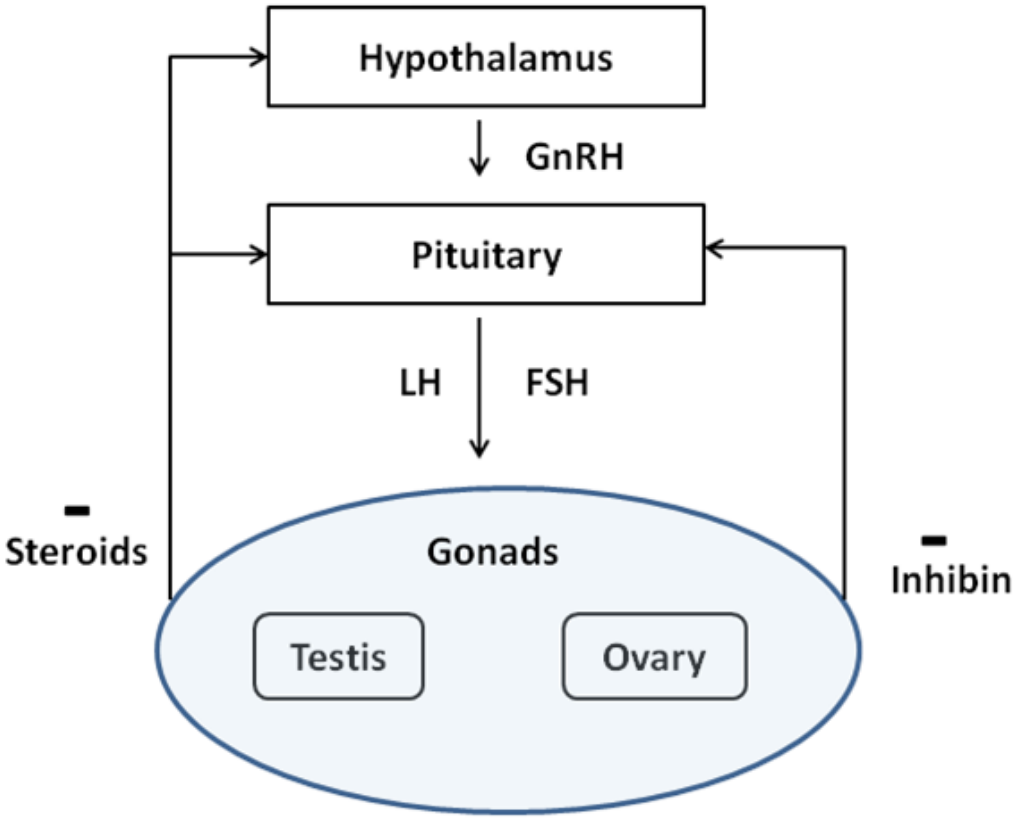


Figure 1: GnRH secreted from the hypothalamic neurons feeds into the anterior pituitary, where it binds its receptor GnRHR, leading to synthesis and secretion of gonadotropin hormones, LH and FSH. These gonadotropin hormones make their way through the blood to the cell surface receptors present on the gonadal cells. FSH binds to its receptor, FSHR present on the Sertoli cells of the testis and the granulosa cells in the ovary. LH binds to its receptor, LHCGR, on the Leydig cells of the testis and in granulosa, theca, luteal, and interstitial cells of the ovary. This leads to production of steroids and inhibin which feedback negatively on the hypothalamus and the pituitary leading to decrease in gonadotrope production.

steroid hormone synthesis and gamete production (reviewed in (Young 1995, Achermann and Jameson 1999, Plant 2008)). At the top of the network is GnRH, which, once released from the hypothalamus, binds specific receptors on pituitary gonadotrophs and induces the synthesis and secretion of LH and the FSH. These dimeric proteins formed by the heterodimerization of a common alpha subunit (α-glycoprotein hormone subunit) with the specific beta subunit (FSHβ or LHβ) confer biological specificity to the hormones. Once in circulation, FSH and LH, bind to their respective receptors on the gonad cells leading to regulation and expression of steroidogenic genes and production of steroid hormones (Figure 1). Expression of a number of steroid hormones in the HPG axis is tightly regulated in a temporal and development manner.

GnRH, GnRHr, and gonadotropins

Gonadotropin releasing hormone is an evolutionary conserved decapeptide that forms a bridge between the brain and the peripheral reproductive system. Since the first isolation of GnRH, in 1971, from pig and sheep brains, the GnRH family now encompasses 24 molecular isoforms identified in vertebrates and invertebrates (Amoss, Burgus et al. 1971, Matsuo, Baba et al. 1971, Gorbman and Sower 2003). In vertebrates, there are two major isoforms of GnRH; the hypothalamic GnRH1, which acts on the pituitary and extra-hypothalamic GnRH2, which has no known functional involvement to gonadotropin release (Gorbman and Sower 2003). Once released from the hypothalamus, GnRH1 travels via the median eminence to the pituitary where it binds the GnRH receptor, a member of the rhodopsin family of G protein-coupled receptor expressed on the pituitary gonadotropes (Tsutsumi, Zhou et al. 1992, Stojilkovic, Reinhart et al. 1994, Kaiser, Conn et al. 1997, Sealfon, Weinstein et al. 1997). GnRH binds to its receptor causing activation of G-proteins, including G_q and $G_{11,}$ subsequent activation of phospholipase Cβ, release of calcium, and increased activity of protein kinase C and calcium/calmodium kinase

(Stanislaus, Janovick et al. 1997, Liu, Austin et al. 2002, Haisenleder, Ferris et al. 2003). Pulsatile release of GnRH stimulates the expression of mitogen-activated protein kinase (MAPK) signaling cascades (MAPK1/3 [extracellular signal-regulated kinase, or ERK], MAPK8/9 [c-Jun N-terminal kinase, or JNK], and MAPK14 [p38]). While the role of p38 in transcriptional regulation of LHβ has been controversial, ERK1/2 and JNK are involved in the induction of early growth response-1 (Egr-1), a factor known to regulate the LHβ promoter (Yokoi, Ohmichi et al. 2000, Harris, Bonfil et al. 2002, Burger, Haisenleder et al. 2004, Yamada, Yamamoto et al. 2004). Female Egr-1 knockout mice were infertile and lacked expression of LHβ (Lee, Sadovsky et al. 1996, Topilko, Schneider-Maunoury et al. 1998). The proximal promoter region of GnRH contains binding site for Egr1, and transcription factors SF-1 and Ptx1. GnRH induces gene expression of Egr-1, leading to a synergistic interaction of Egr-1 with transcription factors, SF-1 and Ptx1 inducing LHβ transcription (Dorn, Ou et al. 1999, Tremblay and Drouin 1999).

GnRH also regulates FSH levels, as GnRH null mice reported 60% reduction in serum FSH levels when compared to normal mice (Mason, Hayflick et al. 1986). Furthermore, administration of a single pulse of GnRH led to a fourfold increase in FSHβ gene expression indicating that GnRH is an important regulator of the FSHβ gene (Burger, Dalkin et al. 2001). GnRH mediates FSHβ gene expression via the PKC and MAPK signaling pathways (Vasilyev, Lawson et al. 2002, Bonfil, Chuderland et al. 2004, Coss, Jacobs et al. 2004, Liu, Ruiz et al. 2005). GnRH also regulates FSHβ gene through the induction of activator protein-1 (AP-1), a heterodimeric transcription factor that consists of a variety of dimers of Fos and Jun isoforms (Wurmbach, Yuen et al. 2001, Kakar, Winters et al. 2003). Mice deficient of c-Fos had smaller ovaries with atretic follicles, similar to FSHβ knockout mice (Johnson, Spiegelman et al. 1992, Kumar, Wang et al. 1997). AP-1 also interacts with factors NF-Y and USF-1 to stimulate FSHβ

in response to GnRH (Coss, Jacobs et al. 2004, Ciccone, Lacza et al. 2008). Transcription factor cAMP response element-binding protein (CREB) binds to the CRE/AP-1 site in response to GnRH regulation of rat FSHβ (Ciccone, Lacza et al. 2008). CREB deficient mice did not have reduced FSH levels, indicating that either cAMP-responsive element modulator (de Kok, Merkx et al.) or activating transcription factor (ATF) family members may compensate for the deficiency or CREB activity on FSHβ activity is a species-specific phenomenon (Hummler, Cole et al. 1994, Ciccone, Lacza et al. 2008).

In this dissertation, I will explore the transcriptional regulation of two critical genes involved in the HPG axis. In Part I, I will discuss distal regulatory elements and the role of CCCTC-binding factor (CTCF) in the regulation of the Follicle Stimulating hormone receptor (*Fshr*) gene. In Part II, I will investigate distal regulatory regions and SF-1 transcription and define a long non-coding RNA transcribed at the SF-1 locus. These distinct regulatory elements will be discussed subsequently under "Transcriptional regulatory elements".

Transcriptional regulatory elements

Various aspects of human development hinge on proper spatial, temporal, and tissue-specific control of gene expression. As we mark 60 years since Watson and Crick proposed the double helical structure of DNA, questions remain on how thousands of genes are orchestrated in a controlled and precise manner (Watson and Crick 1953). Britten and Davidson hypothesized that discrete regulatory sequence not only played a role in regulating gene expression, but also played a major role in governing biological complexity in eukaryotes (Britten and Davidson 1969). This increased level of complexity is understandable with the fruition of the human genome project and identification of ~20,000-25,000 genes (Human Genome Sequencing 2004). It has become

abundantly clear that there are functional regions that extend from the promoters into the vast space of the human genome. Sequencing of human and other vertebrate genomes, coupled with high throughput in-vitro and in-vivo functional screens have facilitated the comprehensive identification of a number of regulatory sequences and helped compile the regulatory architecture required to establish the spatial and temporal expression of genes. Gene expression is regulated by the coordinated expression of distal regulatory elements (Figure 2) which can be crudely classified into a. Core and Proximal promoter, b. Enhancers, c. Silencers, d. Insulators, e. Locus Control regions, and f. non-coding RNA.

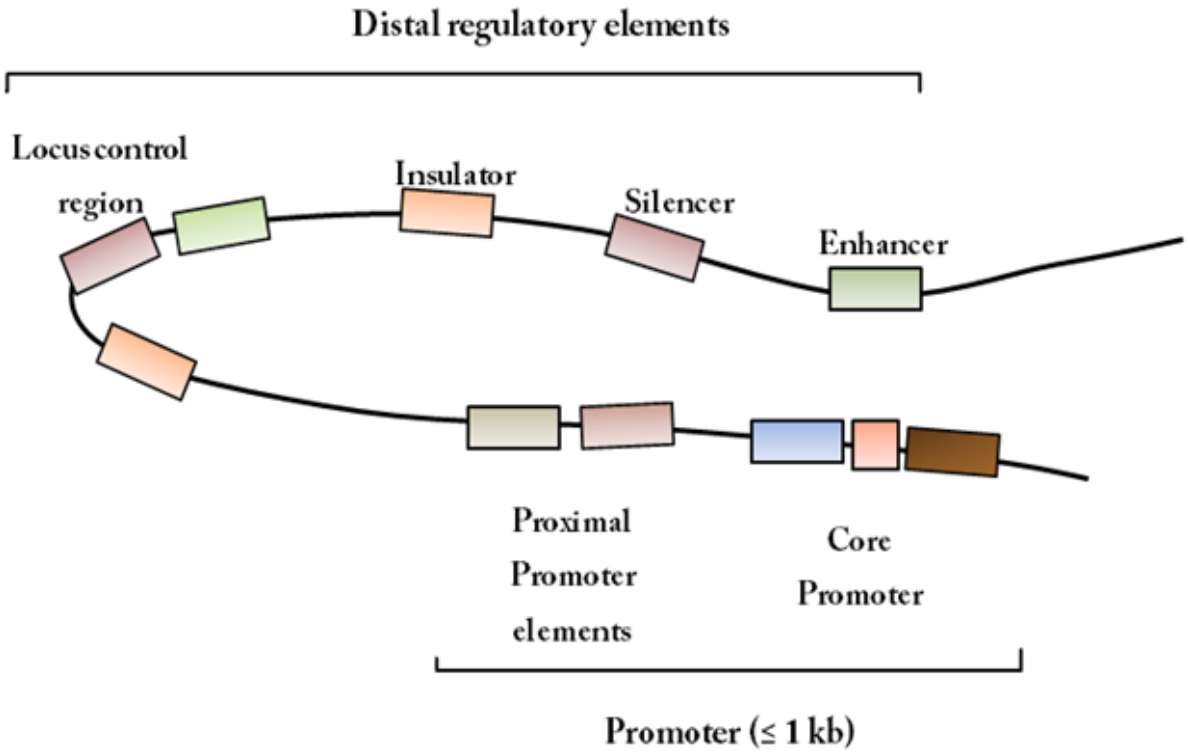


Figure 2: Schematic of classical gene regulator elements. The promoter is composed of a core promoter and proximal promoter elements and mediates basal transcriptional control. Upstream regulatory elements, enhancers and locus control regions, mediate positive transcription while silencers negatively affect transcription. Insulators, play a complex role by acting as enhancer-blockers and barrier insulators. These regulatory elements are thought to carry out their function by employing what is known as "DNA looping". Modified according to (Maston, Evans et al. 2006) .

a. Core and proximal promoters: The core promoter can be defined as the minimal stretch of DNA that serves as the docking site for the basic transcription machinery and encompasses the site and direction of the transcription start site (Smale and Kadonaga 2003). Accurate and efficient transcription of the core promoter requires RNA polymerase II (RNAP II) along with "basal" transcription factors that include transcription factor (TF) IIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH. Assembly of the preinitiation complex starts with the binding of TFIID followed by recruitment of TFIIA and TFIIB. This association leads to the addition of TFIIF bound with RNAP II and of further incorporation of TFIIE and TFIIH leading to the completion of the preinitiation complex and initiation of transcription. An alternate pathway of the preinitiation complex formation is the RNAPII holoenzyme pathway, which comprises of TFIIF, TFIIE, TFIIH and other heterogeneous proteins that are involved in chromatin remodeling, DNA repair, and mRNA processing. Irrespective of the manner in which the preinitiation complex is formed, TFIID is the first TF that binds to the core promoter element and has emerged as the central component of the transcription apparatus (Baumann, Pontiller et al. 2010).

In addition to TATA box, core promoters can consist of Initiator element (Inr), Downstream Promoter Element (DPE), Downstream Core Element (DCE), TFIIB-Recognition Element (BRE), and Motif Ten Element (MTE) (Maston, Evans et al. 2006). A statistical analysis of human core promoters revealed that Inr was the most common core promoter element, followed by DPE and BRE, and surprisingly, TATA being present in the lowest levels. Furthermore, a quarter of the analyzed promoters had neither TATA, Inr, DPE or BRE elements indicating presence of novel promoters (Gershenzon and Ioshikhes 2005). In line with this, recent reports have now identified a new core element labeled as ATG deserts (Lee, Howcroft et al. 2005). This

diverse content of the core promoter possibly play a functional role by contributing to gene specific regulation.

Proximal promoter can be defined as the region immediately upstream of the core promoter and contributes to its basal activity by the presence of multiple binding sites for activators. These transcriptional regulatory elements act synergistically and mutation of any of these sites leads to decreased transcriptional activity (Maston, Evans et al. 2006). While the core and proximal promoter comprise the very basic level of transcriptional regulation, elements that are more distal have now been identified to play a role in this exceeding intricate ballet of gene expression and include enhancers, silencers, insulators, locus control regions and non-coding RNA's which are discussed in the following sections.

b. Enhancers: Enhancers are typically small segments of DNA, typically few hundred base pairs in length, that can reside either in the intergenic regions, introns, or exons and can be located ten to hundred kilobases away from their target genes and operate as transcription factor binding sites, which work cooperatively to enhance transcription. They have the inherent capacity to act in a modular manner such that a single promoter can be acted upon by distinct enhancer elements in a spatial and temporal specific manner (Maston, Evans et al. 2006, Spitz and Furlong 2012). How do these enhancers carry out their function in a spatio-temporal manner? Analysis of transcription factor occupancy at different stages of development in both Drosophila and mammals have identified that precise timing of DNA occupancy is what controls the temporal nature of gene expression driving development (Sandmann, Jensen et al. 2006, Jakobsen, Braun et al. 2007, Cao, Yao et al. 2010, Lin, Jhunjhunwala et al. 2010). Temporal regulation was thought to correspond to relative affinity or number of transcription factor binding sites, affecting change in occupancy relative to concentration of transcription factor over time (Gaudet and

Mango 2002, Sandmann, Jensen et al. 2006). However, evidence now indicate that enhancers with context-dependant occupancy with differential motif enrichment are co-occupied with different transcription factors in a time-dependant manner and this co-occupancy is needed for recruitment of additional transcriptional factors (Zeitlinger, Simon et al. 2003, Jakobsen, Braun et al. 2007, Sandmann, Girardot et al. 2007, Lin, Jhunjhunwala et al. 2010, Wilczynski and Furlong 2010, Mullen, Orlando et al. 2011, Trompouki, Bowman et al. 2011, Yanez-Cuna, Dinh et al. 2012).

Based on the relative order, orientation, and spacing of transcription factors, three models have been proposed to explain enhancer activity. In the enhanceosome model, as exemplified by *IFN-β* and *TCRα*, recruited transcription factors form an ordered and specific position relative to each other, such that only one arrangement can lead to gene expression (Thanos and Maniatis 1995, Spicuglia, Payet et al. 2000, Merika and Thanos 2001). While this ordered positioning is thought to be essential for rapid response, most developmental enhancers are more flexible, with a subset of transcription factors binding co-operatively (Senger, Armstrong et al. 2004). This flexible positioning known as the "billboard model" in which there are fewer constraints on relative binding of transcription factors but requires a subset of factors to be active, such as the stripe 2 enhancer of *Drosophila even-skipped* (Arnosti, Barolo et al. 1996, Ludwig, Patel et al. 1998, Kulkarni and Arnosti 2003, Arnosti and Kulkarni 2005, Hare, Peterson et al. 2008). The third model of regulation, known as the "transcription factor collective", suggests that the same set of transcription factors can bind to multiple enhancers and do so in a cooperative manner. In a recent example when one transcription factor was removed, all other transcription factors failed to activate the enhancer *in vitro* indicating cooperative binding occurs *in vivo*, with all or a sub-

set of transcription factor binding to the DNA and other transcription factors recruited via protein-protein interaction (Junion, Spivakov et al. 2012).

Since enhancers often reside at significant distances, either upstream or downstream of their target promoters, mechanisms must exist for them to communicate over large genomic distances. Current data indicates that chromatin "looping" permits the interaction between the enhancer and the core promoter (Blackwood and Kadonaga 1998, Bulger and Groudine 1999, de Laat, Klous et al. 2008). An alternative model supports a diffusion or "tracking" mechanism, in which the enhancer scans along the chromatin until it arrives at the core promoter (Blackwood and Kadonaga 1998). With the advent of techniques such as chromosome conformation capture (3C), in which the interactions between two genomic regions are determined by cross-linking, restriction enzyme digestion, followed by intra molecular ligation and PCR analysis of the resulted ligated products, data has now accumulated in favor of the looping model (Cullen, Kladde et al. 1993, Dekker, Rippe et al. 2002, Miele and Dekker 2009).

c. Silencers: Silencers are sequence specific elements that silence or repress transcriptional activity of a gene and act typically in a distance and orientation independent manner (Brand, Breeden et al. 1985, Ogbourne and Antalis 1998). Silencers can be situated at the proximal promoter, as part of distal elements, or independently far away from its target gene in the intron or the 3'-untranlsated region (Maston, Evans et al. 2006).

Silencers repress promoter activity by binding of negative transcription factors called repressors and recruitment of negative cofactors, called corepressors (Privalsky 2004). Based on distance of action, two classes of silencer elements have been identified in *Drosophila,* short-range silencers that reside within ~100 bp of the target gene and repress its transcription, and long-range

silencers, which can affect multiple enhancers or promoters over a span of few kilobases. The current understanding is that this distance dependant repression is due to the difference in recruitment of different co-factors (Kulkarni and Arnosti 2005).

These silencer elements repress gene transcription by blocking the binding of a nearby activator or by competing for activator binding sites (Li, He et al. 2004, Harris, Mostecki et al. 2005). Other models that have been proposed to explain the repressor activity suggests that they may establish a repressive chromatin structure by recruiting histone modifications or chromatin stabilizing factors or by inhibiting assembly of the preinitiation complex (Srinivasan and Atchison 2004, Chen and Widom 2005).

d. Insulators: Insulators, also known as boundary elements are a family of DNA sequence elements typically ~0.5-3kb in length, that protect genes from transcriptional activity of its neighboring genes, in a position-dependent and orientation independent manner (Maston, Evans et al. 2006). Originally described in Drosophila, insulators have now been found in most eukaryotes, from yeast to human (Phillips and Corces 2009). Insulators that disrupt communication between enhancer and promoter are known as enhancer-blocking insulators and those that protect against heterochromatin silencing are known as barrier insulators (Sun and Elgin 1999, Valenzuela and Kamakaka 2006).

The first vertebrate insulator identified was a Dnase1 hypersensitive site (HS), 5' HS4, sequence of the chicken β-globin locus. The β-globin locus was an ideal starting ground due to the extensive work on local and distant regulatory elements and the fact that chicken erythrocyte nuclei were relatively free of proteases and nucleases. The folate receptor gene located upstream is separated from the β-globin locus by a 16-kb long region of condensed chromatin that bear

silenced heterochromatin marks of H3K9 and K27 methylation (Litt, Simpson et al. 2001, Litt, Simpson et al. 2001). This sudden transition of heterochromatin to euchromatin at the active *β*-globin locus suggested the presence of a boundary sequence that separated the domains. Dnase 1 hypersensitivity and H3 lysine 9 acetylation histone modification identified the HS4 insulator element (Litt, Simpson et al. 2001, Litt, Simpson et al. 2001, Ma, Heath et al. 2011). Further work identified a core ~250bp sequence of the HS4 insulator element with properties of both enhancer blocking and barrier activity (Chung, Whiteley et al. 1993, Pikaart, Recillas-Targa et al. 1998, Bell, West et al. 1999, Recillas-Targa, Pikaart et al. 2002). The transcription factor mediating this activity was identified as CCCTC-binding factor (CTCF), an eleven-zinc finger protein, both necessary and sufficient for the enhancer-blocking activity of the 5' HS4 (Bell, West et al. 1999). Another CTCF binding site was identified in the 3' end of the chicken *β*-globin locus (Bell, West et al. 1999). In parallel, CTCF binding sites were subsequently identified within the imprinted control region (ICR) of the mammalian H19/Igf2 locus (Bell and Felsenfeld 2000, Hark, Schoenherr et al. 2000, Szabo, Tang et al. 2000). To date, CTCF has been the only protein identified to show enhancer-blocking activity in vertebrates**.**

CTCF is a multivalent factor, originally identified, owing to its ability to bind to a variety of sequences as well as co regulatory proteins using multiple zinc fingers (Filippova, Fagerlie et al. 1996). This unique structural feature suggested that CTCF plays a varied role in genome regulation. Global knockout of CTCF led to embryonic lethality prior to implantation (Splinter, Heath et al. 2006, Heath, Ribeiro de Almeida et al. 2008). Depletion of CTCF in oocytes caused transcriptional misregulation of hundreds of genes (Wan, Pan et al. 2008). Taken together this data suggests that CTCF plays an important and varied role in gene regulation that include

traditional enhancer blocking, promoter activation/repression, hormone-responsive silencing, long-range chromatin interactions and genomic imprinting (Phillips and Corces 2009).

The *H19/Igf2*, one of the first identified imprinted loci, is a classical example wherein genetic information is stored in two copies, one from each parent, with expression of one parental copy while silencing the other (Bartolomei, Zemel et al. 1991, DeChiara, Robertson et al. 1991, Ferguson-Smith, Cattanach et al. 1991). Gene knockout studies have identified that Igf2 is crucial for placental and fetal growth, whereas the H19 gene encodes for a non-coding RNA, and retards fetal growth. The same enhancer activates H19 on the maternal allele and Igf2 on the paternal allele. In the paternal allele, the ICR region immediately upstream of the H19 gene is methylated preventing CTCF binding, abrogating the insulator activity, resulting in the functional communication between the Igf2 promoter and enhancer. In the maternal allele, ICR is unmethylated, which allows CTCF to bind to the ICR region and prevents the enhancers from accessing the Igf2 promoter (Herold, Bartkuhn et al. 2012). Proper expression of the parental allele is crucial, as mutant H19 gene is one of the causes of Beckwith-Wiedemann syndrome (BWS); a condition associated with fetal overgrowth and increased risk of tumor formation. BWS phenotype has been observed by increased DNA methylation at ICR resulting in biallelic Igf2 expression and reduced H19 activity and also in patients with micro-deletions at the CTCF binding site at the ICR locus, indicating that loss of CTCF binding at ICR either due to mutations or due to DNA methylation causes Igf2 activation and H19 inactivation (Sparago, Cerrato et al. 2004, Azzi, Rossignol et al. 2009). Conversely, loss of methylation at the paternal ICR, results in biallelic expression of the H19 gene and reduced expression of Igf2, is observed in patients diagnosed with Silver-Russell syndrome (SRS), a condition characterized by growth retardation during fetal and postnatal development (Herold, Bartkuhn et al. 2012).

With the use of techniques such as fluorescence in situ hybridization (FISH) and chromosome capture conformation (3C), a greater understanding has now emerged on the importance of nuclear architecture (Fraser and Bickmore 2007, Miele and Dekker 2009). For the mouse *H19/Igf2* locus, silencing of the *Igf2* occurs by the formation of a tightly coiled loop by the interaction of CTCF bound at the ICR and the upstream differentially methylated region (DMR) and a downstream matrix attachment region (MAR), which prevent the interaction between the maternal-specific enhancer and the promoter (Yoon, Jeong et al. 2007, Li, Hu et al. 2008). The overlapping distribution patterns of genome-wide analysis of CTCF and cohesin as well as functional tests have now identified cohesin to be an important mediator for CTCF dependent loop function, with cohesin conferring the linking function and CTCF responsible for the sequence-specific binding (Parelho, Hadjur et al. 2008, Wendt, Yoshida et al. 2008, Hadjur, Williams et al. 2009, Nativio, Wendt et al. 2009). In fact, the cohesin component SA2 directly contacts CTCF, with the ring-forming components recruited by SA2 (Xiao, Wallace et al. 2011). Other players involved in the looping and insulation functions of CTCF include RNA helicase p68 and the non-coding RNA steroid receptor RNA activator (SRA), which is thought to help stabilize the CTCF-cohesin interaction (Yao, Brick et al. 2010).

e. Locus control regions: Locus control regions (LCR) are groups of regulatory sequences that regulate transcriptional programming of either an entire locus or gene cluster. They are operationally defined as elements that are able to enhance tissue-specific physiological expression of linked genes in a position-independent and copy-number dependent manner(Li, Peterson et al. 2002). LCRs are typically composed of multiple regulatory elements such as enhancers, silencers, insulators, and matrix attachment regions that are bound by transcription factors, which differentially affect gene expression. These collective activities define a LCR and

regulate spatio-temporal gene expression (Maston, Evans et al. 2006). Similar to enhancers and silencers, LCRs can regulate gene transcription from a distance. While typically located upstream, LCRs are reported to be present in introns, downstream of genes and even in introns of neighboring genes (Lang, Mamalaki et al. 1991, Aronow, Silbiger et al. 1992, Adlam and Siu 2003, Lee, Fields et al. 2003).

The first LCR region to be identified and the best studied to date, is the mammalian β-globin gene (Grosveld, van Assendelft et al. 1987). Transgenic mice containing 1.5kb of the mammalian β-globin promoter region failed to express β-globin gene to the extent seen in vivo, indicating that major regulatory regions required for proper expression were missing (Magram, Chada et al. 1985, Townes, Lingrel et al. 1985, Kollias, Wrighton et al. 1986). Further evidence from cases of β-thalassemia and later from transgenic mouse studies indicated that a LCR region upstream of the promoter was required for proper expression of the β-globin gene in vivo (Kioussis, Vanin et al. 1983, Grosveld, van Assendelft et al. 1987, Driscoll, Dobkin et al. 1989) .

Although a number of models have been proposed to understand how LCRs function over long distance, a series of recent studies showed that in the case of the β-globin LCR, DNA looping is the primary mechanism employed by LCRs to make long-range physical contacts that lead to active chromatin formation and transcription of the β-globin gene (Liu, Austin et al. 2002, Tolhuis, Palstra et al. 2002, Bank 2006). The looping mechanism is also currently established for the Th2 LCR (Spilianakis, Lalioti et al. 2005).

f. non-coding RNA: The advent of high throughput assays and next generation sequencing has expanded the catalog of functional non-coding RNA. Among the growing classes of functional RNA, non-coding RNAs can be classified as short non-coding RNA, which are less than 200

base pairs (bps) in length, and the long-non coding RNAs (lncRNA) which are larger than 200 bps in length (Gibb, Brown et al. 2011).

Small non-coding RNA are functionally sub-divided into housekeeping non-coding RNA and regulatory non-coding RNA(Ponting, Oliver et al. 2009). The housekeeping non-coding RNA consists of ribosomal, transfer, small nuclear and small nucleolar RNAs. While a plethora of regulatory non-coding RNA have been identified (see (Esteller 2011) for review), the three main classes of small regulatory non-coding RNAs that have been well studied are the small interfering RNA (siRNA), micro RNA (miRNA) , and the piwi-interacting RNA (piRNA).

Small interfering RNA's are derived from long double stranded RNA (dsRNA) precursors that are either endogenous or exogenous in origin (Mello and Conte 2004). Dicer, a ribonuclease III enzyme, processes the dsRNA, leading to the formation of a short double stranded intermediate RNA. One strand, labeled as the guide strand is loaded onto the multi protein RNA-silencing complex (Driscoll, Dobkin et al.), leading to the posttranscriptional repression by endonucleolytic cleavage of the mRNA (McCaffrey, Meuse et al. 2002, Vagin, Klenov et al. 2004, Siomi, Sato et al. 2011). Small interfering RNA's are also know to induce transcriptional gene silencing by inducing heterochromatin formation (Carthew and Sontheimer 2009).

micro RNA are a family of endogenous, short, non-coding RNA that consist of 21-25 nucleotides that regulate gene expression at the post-transcriptional level (Du and Zamore 2005, Cannell, Kong et al. 2008, Suh, Baehner et al. 2010). RNA polymerase II transcribes miRNA into 2-4kb long single stranded RNA which are capped and polyadenylated (Bartel 2004, Bartel 2005). These primary transcripts called the pri-mRNA form a stem-loop structure with an imperfectly paired ~33bp stem, a terminal loop and flanking segment. Drosha, a RNase III enzyme, bound to

co-factor DiGeorge Syndrome Critical region 8 (DGCR8) processes the pri-miRNA in the nucleus by excising the stem-loop to create the pre-miRNA. The pre-miRNA is exported to the cytoplasm by exportin-5 together with RAN-GTPase. Once exported into the cytoplasm, Dicer processes the pre-miRNA, in concert with Tar RNA-binding protein to form the ~22 nucleotide mature duplex miRNA (Liu, Fortin et al. 2008, Kim, Hwang do et al. 2009). One strand of the miRNA incorporates onto the RISC to function as mature miRNA and guide the RISC to target the specific mRNA (Lund and Dahlberg 2006, Chua, Armugam et al. 2009). The miRNA binds to the 3'-UTR "seed sequence" of the target mRNA transcript. If perfect or near-perfect base pairing occurs between the seed sequence and the miRNA, the mRNA transcript is degraded, inhibiting translation (Jackson and Standart 2007). However, multiple partial complementarity between miRNA and target mRNA can lead to decreased mRNA and protein levels (Pillai 2005).

miRNA are also implicated in regulation of gene expression by increasing translation in some biological systems (Vasudevan, Tong et al. 2007). Since the discovery of miRNA in *Caenorhabditis elegans*, miRNA have been shown to have diverse expression and act as master regulators of a number of fundamental biological progresses, such as proliferation, metabolism, aging, and cell death (Pasquinelli and Ruvkun 2002, He and Hannon 2004, Brennecke, Aravin et al. 2007, Melton, Judson et al. 2010). In the female reproductive tract, miRNAs have been shown to be essential for proper development and function, with loss of *Dicer1* resulting in female infertility (Nothnick 2012).

P-element induced wimpy testis (Piwi)-interacting RNA (piRNA), discovered in 2006, are small non-coding RNAs, ~26-30 nucleotides in length. PiRNAs were isolated by sequencing the small non-coding RNAs pulled down by immunoprecipitation of the Piwi protein in mammalian testes (Aravin, Gaidatzis et al. 2006, Girard, Sachidanandam et al. 2006, Grivna, Pyhtila et al. 2006,

Lau, Seto et al. 2006). Unlike the miRNA and the siRNA, these piRNAs are generated in a RnaseIII independent manner (Houwing, Kamminga et al. 2007). They also differ from miRNA and siRNA in origin, length, and structure. piRNAs are generated from nascent transcripts that arise from the piRNA clusters, intergenic repetitive elements in the genome, which are loaded onto the PIWI proteins. On maturation, the 3' ends of the piRNAs are 2′-*O*-methylated by Hen1/Pime, leading to the stability of the piRNAs in vivo (Horwich, Li et al. 2007, Billi, Alessi et al. 2012). The ping-pong pathway amplifies primary piRNAs, wherein the antisense piRNA are associated with the Piwi subfamily of Argonaute proteins including Mili and Miwi, leading to the production of sense piRNAs, which then associate with Argonaute protein, Ago 3, and target anti-sense transcripts (Figure 2) (Siomi, Sato et al. 2011).

While the main role of piRNAs been associated with curbing the silencing of transposable elements in the germ cell line, recent studies have also implicated them in regulating memory storage in the brain (Siomi, Sato et al. 2011, Rajasethupathy, Antonov et al. 2012). In *Drosophila melanogaster*, piRNAs are transmitted maternally and are able to mount an effective silencing response in the progeny (Brennecke, Malone et al. 2008).

Noncoding RNAs that are greater than 200 nucleotides in length and do not have any coding potential are defined as long noncoding RNA (lncRNA). The completion of the Human Genome project and advancements in RNA sequencing, microarray technology, and cDNA cloning led to the conclusion that the number of genes coding for RNA were far greater than protein-coding genes. It was now clear that a vast majority of the genome was transcribed, however, the function of these noncoding RNAs was elusive, leading to the belief that these were products of translational noise (Shoemaker, Schadt et al. 2001, Birney, Stamatoyannopoulos et al. 2007, Ponjavic, Ponting et al. 2007, Struhl 2007, Clark, Amaral et al. 2011). Furthermore, RNA

polymerase II is known to initiate transcription spuriously in more than 90% of the cases, and transcription events can "spill over" or ripple to the neighboring genes leading to leaky expression (Ebisuya, Yamamoto et al. 2008).

Identifying functional lncRNA from the pervasive transcription of the genome is a daunting task and tantamount to finding a needle in a haystack. The massive parallel sequencing and understanding of different chromatin modifications identified clear signatures of polymerase II binding, namely Histone H3 lysine 4 trimethylation (H3K4me3) at the gene promoters, and histone H3 lysine 36 trimethylation (K4-K36 chromatin domain) at the transcribed region. Coupled with conserved patterns, coding patterns, and anatomical properties, great progress is now being made in identifying lncRNA (Mikkelsen, Ku et al. 2007, Marson, Levine et al. 2008, Guttman, Amit et al. 2009, Khalil, Guttman et al. 2009).

Based on their geographical location in the genome, relative to nearby protein-coding gene lncRNA can be grouped in: a. stand-alone lncRNA: b. natural antisense transcripts, c. pseudogenes, d. long intronic ncRNA and e. divergent transcripts, promoter-associated and enhancer RNA (Figure 3) (Kung, Colognori et al. 2013).

a. Stand-alone lncRNA: Also referred to as large intergenic noncoding RNAs (lincRNA), are located in regions that do not overlap protein-coding genes (Guttman, Amit et al. 2009, Cabili, Trapnell et al. 2011, Ulitsky, Shkumatava et al. 2011). Largely identified through active chromatin signatures of H3Kme3 at the promoter and H3K36me3 along the transcribed region, these noncoding RNA are usually 1kb in length, are transcribed by RNA pol II, are polyadenylated, and spliced. Examples of this noncoding RNA are *Xist*, *H19*, *HOTAIR*, and

MALAT1 (Brannan, Dees et al. 1990, Brockdorff, Ashworth et al. 1992, Brown, Hendrich et al. 1992, Ji, Diederichs et al. 2003, Rinn, Kertesz et al. 2007).

b. Natural antisense transcript: As the name suggests, these transcripts occur opposite the sense DNA strand and are enriched at either the 5' promoter or the 3' terminator end of the sense transcript. Rarely are the lncRNA belonging to this class spliced or polyadenylated when compared to the stand-alone lncRNA (Kung, Colognori et al. 2013). Well-documented transcripts of this class are *Kncq1ot1* and *Air* (Lyle, Watanabe et al. 2000, Kanduri, Thakur et al. 2006).

c. Pseudogenes: The term "pseudogene" coined in 1977, refers to genes that have lost their ability to code for functional proteins due to nonsense, frameshift and other mutations (Jacq, Miller et al. 1977, Pink, Wicks et al. 2011). Estimated to be numerically equal to protein-coding genes, these pseudogenes are by majority, transcriptionally silent. However, there are pseudogenes that are transcriptionally active and exhibit high levels of sequence conservation. These expressed pseudogenes maybe on their way to be completely silenced or are possibly resurrected for a functional purpose (Pink, Wicks et al. 2011). Long noncoding RNA produced from pseudogenes have added another interesting layer for gene

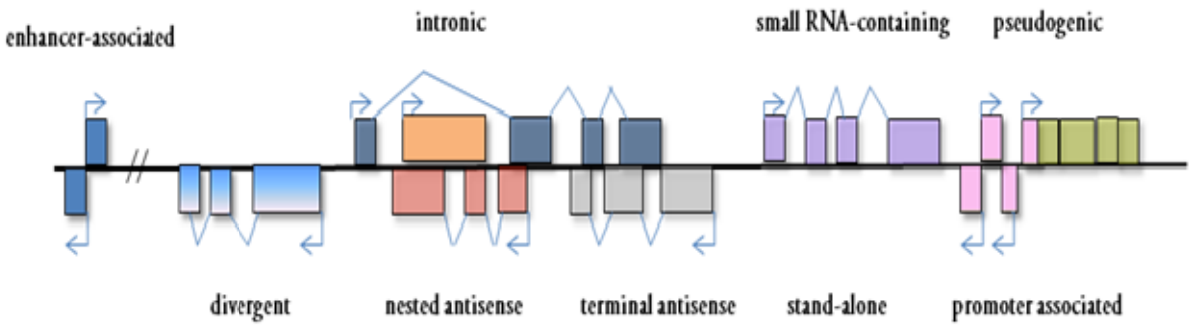


Figure 3: A summary of genomic location of lncRNA: LncRNA can be transcribed from enhancer, associated with the promoter or could be stand-alone transcripts. They are also intronic, or can be antisense to other genes, completely anti-sense (nested) or partially (terminal). They can also arise from pseudogenes or could host small RNA. Adapted from (Kung, Colognori et al. 2013)

regulation. A recent example is the PTENpg1, a *PTEN* pseudogene that regulates *PTEN* by encoding for three different lncRNA; two functional antisense (as) RNA and one sense PTENpg1. The sense PTENpg1 regulates *PTEN* by acting as a sponge for several miRNA that target *PTEN*. The PTENpg1 asRNA_{α} is a trans acting transcript that binds to the promoter of PTEN, and inhibits transcription by recruiting epigenetic repressor complexes. The PTENpg1 as RNA_{β}, stabilizes the sense PTENpg1 by binding to the 5'end (Johnsson, Ackley et al. 2013).

d. Long Intronic RNA: RNAs with regulatory roles have long been known to reside in intronic regions, indicating that intronic regions maybe more stable than previously thought. While a substantial fraction of the long intronic RNA correspond to the antisense transcript, a recent survey revealed noncoding RNA to have a strong preference to be associated with the transcribing sense strand (Nakaya, Amaral et al. 2007, Valen, Preker et al. 2011). Many of these are implicated in a myriad of biological functions and have been observed to respond to stimuli or misregulated in cancer (Guil, Soler et al. 2012). An example of long intronic RNA is COLDAIR, transcribed from the sense strand, it has been implicated in plant vernalization by recruiting PRC2 complex and inducing epigenetic silencing (Heo and Sung 2011).

e. divergent transcripts, promoter-associated and enhancer RNA: As a result of Pol II pausing a number of transcripts, such as transcription start site-associated (TSSa-) RNA, upstream antisense (ua) RNA or promoter upstream transcripts (PROMPTS), are formed at the transcription start site, both in the sense and anti-sense direction (Buratowski 2008, Core and Lis 2008). These transcripts are usually capped, polyadenylated, low in abundance, and degraded rapidly by exosomes. The current thought is that the act of transcription of these heterogeneous transcripts helps maintain the open chromatin status.

Another class of regulatory long noncoding RNA is the short, bidirectional transcripts formed at enhancer regions and hence called enhancer (e) RNA. These transcripts are postulated to function by activating the promoter or by keeping the chromatin in an open state, or by recruiting and interacting with enhancer-associated proteins (Orom, Derrien et al. 2010).

The ubiquitous expression of lncRNA indicates that these transcripts engage in diverse mechanisms to regulate gene expression. Over the past 15 years, multiple studies have shown that lncRNA are regulated during development, epigenetic regulation, imprinting, and are also associated with human disease (Kung, Colognori et al. 2013). Broadly, lncRNA can be classified according to their mode of action into four archetypes, a. signals, b. decoys, c. scaffold, and d. guide (Wang and Chang 2011). While this classification will help in demarcating their mode of regulation, an lncRNA may utilize multiple modes to bring about their function indicating that these archetypes are not mutually exclusive.

a. Signals: LncRNA can act as signaling molecules as they show cell-specific expression and respond to various stimuli. While some lncRNA belonging to this category have regulatory function, even as by-products, the act of transcription of lncRNA in itself indicates that the chromatin is in an active state. The advantage of using RNA as a medium is that the cell can skip protein translation and as a result, the regulatory functions can be performed quickly (Wang and Chang 2011). One example where lncRNA can act as signaling molecules is during imprinting. Kcnq1ot1 and Air transcribed at the *Kcnq1* and *Igf2r* locus respectively, mediate transcriptional silencing by recruiting chromatin modifying complexes. *Kcnq1ot1*, a 90kb lncRNA directs silencing of a cluster of genes at the *Kcnq1* domain, in the paternal allele. Kcnq1ot1 recruits histone methyltransferases G9a and PRC2, leading to bidirectional silencing of genes in the Kcnq1 domain (Pandey, Mondal et al. 2008). Similarly, the lncRNA *Air*, transcribed from the

second intron of the *Igf2r* gene, represses several genes at the paternal chromosome in a tissue specific manner. The *Air* lncRNA mediates silencing by recruiting G9a, a histone methyltransferase (Nagano, Mitchell et al. 2008).

X inactivation, a process in which one of the two chromosomes in the female is silenced, such that only one of the X chromosome is expressed is now known to be controlled by a cluster of lncRNA termed X-inactivation center. The 17kb X inactive specific transcript (*Xist*), is expressed from the inactive X. It coats the X chromosome and forms a "Xist cloud" leading to the recruitment of the PRC2 complex (Zhao, Sun et al. 2008) . *Xist* is regulated by an antisense lncRNA called *Tsix*, which reverses the action of *Xist* by repressing the silencing caused by *Xist* and *Tsix* accomplishes this by several different mechanisms. Recruitment of the polycomb complex to the 5' end of *Xist* is accomplished by a 1.6kb ncRNA called RepA. *Tsix* blocks loading of this complex to *Xist*, inhibiting induction of the lncRNA, *Xist* (Zhao, Sun et al. 2008). Transcription of *Tsix* also causes silencing of *Xist* activity by recruiting DNA (cytosine-5)-methyltransferase 3A (Dnmt3a) enzyme leading to the formation of repressive histone modifications inhibiting transcription of *Xist* (Sado, Hoki et al. 2005). More recently, it has been shown that the very act of transcription of *Tsix* can lead to suppression of *Xist* as insertion of poly A cassette leading to the truncation of *Tsix,* releases the repressive environment at the Xist promoter (Ohhata, Hoki et al. 2008). Another lncRNA *Jpx*, regulated in trans activates *Xist* on the inactive X chromosome (Lee, Davidow et al. 1999, Chureau, Prissette et al. 2002, Johnston, Newall et al. 2002, Tian, Sun et al. 2010).

b. Decoys: In this mechanism, lncRNA act as molecular decoys by binding and pirating a protein away from its intended target and restricting the functional outcome. In this case, lncRNA act as a repressor of function and its action can be negated by knocking down the lncRNA, allowing

the protein to bind and leads to gain of function (Wang and Chang 2011). The human dihydrofolate reductase (*DHFR*) gene is an example of RNA dependant repression. The minor upstream promoter of the *DHFR* gene initiates transcription of the DHFR lncRNA, inhibiting the formation of the Pre-initiation complex (PIC) at the major complex by forming a stable ncRNA-DNA complex and by directly interacting with the general transcription factor IIB (TFIIB). Knockdown of the lncRNA, removed the repressor effect of the lncRNA, leading to high occupancy of the TFIIB at the major promoter (Martianov, Ramadass et al. 2007).

The "molecular sink" mechanism of these lncRNA is not restricted to proteins but also encompasses miRNAs and splicing factors. The tumor suppressor gene PTENP1 functions as a decoy and sequesters miRNA that would affect the transcriptional regulation of the PTEN gene. The 3' UTR sequence of the PTENP1 RNA is similar to that of the PTEN gene, allowing the miRNA to bind to the PTENP1 RNA, and in the process, allowing the PTEN gene to be transcribed and translated (Poliseno, Salmena et al. 2010).

Metastasis-associated lung adenocarcinoma transcript (MALAT1) is a nuclear lncRNA abundantly present in nuclear speckles. This lncRNA binds to several serine/argenine (SR) splicing factors and sequesters them into nuclear speckles. Depletion of MALAT1 leads to altered splicing pattern for pre-mRNA (Tripathi, Ellis et al. 2010). MALAT1 regulates splicing factors in hippocampal neurons and is important for synapse formation (Bernard, Prasanth et al. 2010). Thus, lncRNA can act as decoys and sequester away regulatory factors both in the cytoplasmic and nuclear domains.

c. Guides: The third mechanism of lncRNA is as a guide leading to the proper localization of specific complexes to their target regions. This can occur both in *cis* (on neighboring genes) or in

trans (distantly located genes) and can include both repressive and activating complexes. The underlining concept is that these groups of lncRNA convey regulatory information and control gene expression. The functional role of these lncRNA can be predicted by knocking down the lncRNA leading to the loss of proper localization of the effector molecule or loss of function of the effector, or both.

One of the best-studied "guide" lncRNA is at the X-inactivation center (XIC), which controls silencing of one of the X chromosomes. Coating of Xist on the X chromosome marks it for repression. This silencing is initiated by the recruitment of the PRC2 complex by RepA, a 1.5kb lncRNA, which originates from the 5' end of Xist (Wutz, Rasmussen et al. 2002, Sun, Deaton et al. 2006). A similar mechanism is employed by the lncRNA Air, which recruits G9a leading to H3K9 methylation and silencing (Nagano, Mitchell et al. 2008). In the plant lncRNA, COLDAIR, guides the PRC2 complex to the floral repressor gene FLC, during vernalization, leading to its repression by trimethylation of H3K27 (Heo and Sung 2011).

LncRNA act both in *cis* as well as in *trans* as seen in the case of the Hox lncRNA HOTAIR. HOTAIR plays a critical role in cancer cells, as depletion of this lncRNA reduces invasiveness of cells that express high levels of PRC2 (Gupta, Shah et al. 2010). Long intergenic non-coding RNA-p21 induces gene expression changes across multiple sites in the genome. Ectopic expression of LincRNA-p21 leads to apoptosis by bypassing the upstream p53 (Huarte, Guttman et al. 2010).

d. Scaffolds: Originally thought to be the function of proteins, recent evidence now point to the possibility that lncRNA may play similar roles (Good, Zalatan et al. 2011). Highly intricate and complex, the lncRNA binds to multiple effectors at the same time leading to either gene

activation or repression. Knockdown of these lncRNA could change the spatial occupancy of the effector leading to dismantling of the complex, loss of phenotype, or both.

The scaffold mechanism can be seen in telomerase activity, a fundamental process that maintains genomic stability by adding back telomere DNA repeats lost from chromosome ends. This catalytic activity requires the association of TERT, a catalytic protein subunit and a telomerase RNA (TERC). TERC provides a template, contributes to binding of TERT and its catalytic activity, and plays a major role in stability of the complex(Lustig 2004). In the disease condition of dyskeratosis congenita, mutations affect the conformation of TERC leading to loss of RNA scaffold structure underlining the functional importance of TERC (Chen and Greider 2004).

Recent work has identified that a 300-nucleotide fragment present at the 5' end of the lncRNA HOTAIR, is responsible for the binding of the PRC2 complex (Tsai, Manor et al. 2010). In addition, the 700 nucleotides of the 3'end of HOTAIR was found to interact with Lysine specific demethylase (LSD1), RE1 silencing transcription factor (REST), and its co-repressor protein, CoREST that demthylates H3 on K4 leading to gene repression (Tsai, Manor et al. 2010). These findings identify HOTAIR as a scaffold and bridge between the PRC2 and the LSD1/CoREST/REST complex leading to suppression of gene expression.

While these regulatory mechanisms encompass the majority of known functional lncRNAs, the growing numbers of lncRNA are certain to reveal more mechanistic and functional aspects of lncRNA. While the argument remains as to if these merely represent transcriptional noise, support is increasing for their functional significance and thus adds another level of complexity to the mechanisms of gene regulation. The varied sub-cellular expression, nuclear or cytoplasmic, cell-specific expression patterns, and varying levels of conservation among the

lncRNA add chaos to the understanding of these transcripts. Another intriguing possibility is that lncRNAs may be tools for evolution to tinker in generating mechanisims for survival and advancement of the species(Kung, Colognori et al. 2013). One thing is for sure, we have barely scratched the surface of the lncRNA world and as more are discovered, their versatility as regulators of gene expression will be revealed.

The writing in this chapter has been previously published in Biology of Reproduction.

Current concepts Of follicle stimulating hormone receptor gene regulation.
George JW, Dille EA, Heckert LL.
Biol Reprod. 2011 Jan;84(1):7---17.

The gonadotropin hormones

Luteinizing hormone (LH) and follicle-stimulating hormone (FSH), are integral parts of the neural and endocrine interchange between the hypothalamus, pituitary, and gonads that controls steroid hormone synthesis and gamete production (reviewed in (Plant 2008)). At the top of the network is gonadotropin releasing hormone (GnRH), which, once released from the hypothalamus, binds receptors on pituitary gonadotrophs and induces the synthesis and secretion of LH and the FSH. Once in circulation, LH and FSH finalize the communication by binding their receptors and transmitting signals to the gonads. These signals are at the hub of the regulatory network, relaying neuronal signals from the hypothalamus to the gonads and inducing feedback signals returned to the hypothalamus and pituitary. The receptors for FSH and LH, FSHR and LHR (LHCGR), reside on the surface of somatic cells in the gonads and are members of the Rhodopsin receptor family of G-protein coupled receptors, but unlike the other members, LHR and FSHR have extended NH2-terminal extracellular domains with numerous leucine-rich repeats that assist ligand specificity (Braun, Schofield et al. 1991, Dias, Cohen et al. 2002, Vassart, Pardo et al. 2004, Bogerd 2007, Lagerstrom and Schioth 2008). FSH binding elicits several diverse signaling events, but the most characterized is that initiated by adenylyl cyclase, followed by induction of cAMP, PKA activation, and protein phosphorylation (Heindel, Rothenberg et al. 1975, Dorrington and Armstrong 1979). FSH binding is also associated with increased intracellular calcium, activation of mitogen activated protein kinase (MAPK), and stimulation of inositol triphosphate (IP3) (Flores, Veldhuis et al. 1990, Tena-Sempere, Manna et al. 1999, Seger, Hanoch et al. 2001).

Since FSH acts exclusively through FSHR, mechanisms controlling receptor expression determines the FSH-responsive cell population and influences their sensitivity to hormone. Thus,

FSHR expression determines both the targets and extent of FSH action, ultimately directing hormone response to granulosa cells in the ovary and Sertoli cells in the testis (Heindel, Rothenberg et al. 1975). In the ovarian granulosa cells, temporal changes in FSH signaling regulate a number of transcriptional, metabolic, and hormonal activities that are important for the proliferation and differentiation events required for follicular growth and oocyte maturation (Abou-Issa and Reichert 1977, Peluso and Steger 1978, Grasso and Reichert 1990, Dunkel, Tilly et al. 1994, Rannikki, Zhang et al. 1995, Sairam, Jiang et al. 1996, Kumar, Wang et al. 1997, Simoni, Gromoll et al. 1997, Simoni, Gromoll et al. 1997, Huhtaniemi and Themmen 2005). In testicular Sertoli cells, the actions of FSH change with testis development (reviewed in (Kishi, Minegishi et al. 1998, Meachem, Ruwanpura et al. 2005)). Initially, during the perinatal period, FSH induces Sertoli cell proliferation and establishes the final Sertoli cell number that will ultimately determine spermatogenic output, while later in development FSH stimulates Sertoli cell transcriptional and metabolic activities, which contribute to the hormonal and nutritional environment necessary for germ cell survival and development (Orth 1984, Orth, Gunsalus et al. 1988, Russell and Griswold 1993, Boitani, Stefanini et al. 1995, Meachem, McLachlan et al. 1996, Shetty, Marathe et al. 1996, Ruwanpura, McLachlan et al. 2008). In both males and females, FSH induces hormonal signals that return to the pituitary and hypothalamus, as part of the feedback mechanism upholding the endocrine balance in the reproductive axis (Benson, Sorrentino et al. 1969, Moguilevsky, Libertun et al. 1970, Yen and Tsai 1971, Shahmanesh, Sedigh et al. 1980, Schwartz 1982, McNeilly, Souza et al. 2002).

FSHR expression

Expression of FSHR, both protein and mRNA, is remarkably limited with respect to its cellular profile, with Sertoli and granulosa cells by far the predominant expressing cell types

(Ketelslegers and Catt 1974, Orth and Christensen 1977, Heckert and Griswold 1991). *FSHR/Fshr* transcripts are first observed in embryonic gonads, around embryonic day 14.5 in males and 20.5 in females (Dankbar, Brinkworth et al. 1995, Rannikki, Zhang et al. 1995). These initial transcripts are incomplete and represent only the extracellular portion of the receptor, with full-length mRNA expressed several days later (Richards, Ireland et al. 1976, Sokka and Huhtaniemi 1990, Rannikki, Zhang et al. 1995). In the rodent ovary, FSHR expression coincides with primary follicle formation and follicular development through the pre-antral stage, with initial full-length transcripts and hormone binding observed shortly after birth (around postnatal day 3) and continuing to increase through day 21 (Sokka and Huhtaniemi 1990, Dunkel, Tilly et al. 1994, O'Shaughnessy, Marsh et al. 1994, Rannikki, Zhang et al. 1995, Drummond, Dyson et al. 1996). In the testis, full-length FSHR mRNA initiates during fetal development (around embryonic day 16.5 in the rat) and expression is maintained throughout development and in the adult testis (Steinberger, Thanki et al. 1974, Nimrod, Erickson et al. 1976, Sprengel, Braun et al. 1990, Heckert and Griswold 1991, O'Shaughnessy, Marsh et al. 1994, Dankbar, Brinkworth et al. 1995, Rannikki, Zhang et al. 1995). Once the spermatogenic cycle is initiated, FSHR levels change with the cycle, with levels highest at stages X-II and lowest at VI-VII (Heckert and Griswold 1991, Kliesch, Penttila et al. 1992, Heckert and Griswold 1993, Rannikko, Penttila et al. 1996). Several signals that regulate ovarian and testicular physiology also influence FSHR expression. In the ovary, FSHR is regulated by a combination of transcriptional and posttranscriptional mechanisms induced by FSH and activin and indirectly by follistatin through its influence on activin (Knecht, Ranta et al. 1983, Woodruff, D'Agostino et al. 1988, Sanford and Batten 1989, Nakatani, Shimasaki et al. 1991, Themmen, Blok et al. 1991, Nakamura, Minegishi et al. 1993, Sites, Patterson et al. 1994,

Minegishi, Tano et al. 1995, Tano, Minegishi et al. 1995, Tano, Minegishi et al. 1997). In the testis, FSHR is primarily regulated by its ligand, which decreases its expression through several mechanisms, including membrane receptor internalization, mRNA stability, and transcriptional regulation (Jahnsen, Gordeladze et al. 1980, Fletcher and Reichert 1984, Shimizu, Tsutsui et al. 1987, Themmen, Blok et al. 1991, Monaco, Foulkes et al. 1995, Viswanathan, Wood et al. 2009).

While FSHR expression is considered gonad-specific and restricted to Sertoli cells in the testis and granulosa cells in the ovary, there are a few notable reports of its presence elsewhere; in particular, uterus, prostate, bone, and the ovarian surface epithelia (Zheng, Magid et al. 1996, Ben-Josef, Yang et al. 1999, Mariani, Salvatori et al. 2006, Sun, Peng et al. 2006). In prostate and ovarian epithelial cells, FSH signaling is implicated in cell proliferation and tumor invasiveness in precancerous and malignant cells, and thus reports of receptor expression are often within the same cell context (Ben-Josef, Yang et al. 1999, Choi, Choi et al. 2004, Ji, Liu et al. 2004, Zhang, Chen et al. 2009). In bone, there is strong evidence linking FSH and FSHR to hypogonadal bone loss in women but there is still uncertainty as to the site of FSH action (Sun, Peng et al. 2006, Prior 2007, Zaidi, Blair et al. 2007, Ritter, Thuering et al. 2008, Robinson, Tourkova et al. 2010). Consequently, it has been difficult to assess from the literature the degree to which normally differentiated cells, other than Sertoli and granulosa cells, actively transcribe *FSHR/Fshr*. However, its specificity can be greatly appreciated using publically available, high-throughput expression data to examine distribution of *FSHR/Fshr* mRNA. As an example, data derived from more than 30 tissues, using massively parallel signature sequencing, (FSHR query with GEO Series accession number GSE1581 http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE1581 and dataset record GDS868)

showed *FSHR/Fshr* mRNA present only in the testis and ovary; a remarkable finding given the high sensitivity of the detection method (Edgar, Domrachev et al. 2002, Barrett, Troup et al. 2007, Barrett, Troup et al. 2009). What is the mechanism responsible for this remarkable cell-specificity? Currently, it is unknown but as reviewed below, the evidence indicates that *FSHR/Fshr* expression is directed by transcription factors that function through elements located at significant distances from the gene itself.

Transcription of *FSHR/Fshr*

FSHR/Fshr transcription contributes to receptor levels and directs its cell-specificity, indicating components of the underlying mechanisms are important for both FSH responsiveness and target cell identity. Our current understanding of *FSHR/Fshr* transcription is derived largely from studies on the rat, murine, ovine, and human *FSHR/Fshr* genes, which focused on characterization of the 5' flanking region. The resulting data provided significant insight, revealing both similarities and differences between promoters of the four species. Since most promoter characteristics were detailed in a previous review, the discussion here will be limited to its prominent features (Heckert 2005). The accumulated information represents primarily transient transfection and DNA/protein binding results from Sertoli and granulosa cells. These studies identified regulatory elements and their associated binding proteins within promoters represented by various species and lengths and revealed both similarities and differences in promoter function (Huhtaniemi, Eskola et al. 1992, Gromoll, Dankbar et al. 1994, Goetz, Lloyd et al. 1996, Heckert, Daggett et al. 1998, Heckert, Sawadogo et al. 2000, Heckert 2001, Kim and Griswold 2001, Xing and Sairam 2001, Xing, Danilovich et al. 2002, Xing and Sairam 2002, Xing and Sairam 2002). Promoter sequence comparisons between several species showed significant conservation of approximately 1000bp 5' to the translational start, a reference point

used to avoid uncertainty in transcriptional start sites for some species (Huhtaniemi, Eskola et al. 1992, Gromoll, Dankbar et al. 1994, Sairam and Subbarayan 1997, Tena-Sempere, Manna et al. 1999, Xing and Sairam 2001, Heckert and Griswold 2002, Heckert 2005). While variations in methodology and promoter context make it hard to assess relevance to regulatory sequences identified in a single species or study, results representing the rat, sheep, and human promoters identified a common E box element (Goetz, Lloyd et al. 1996, Heckert, Daggett et al. 1998, Findlay and Drummond 1999, Xing and Sairam 2001, Xing, Danilovich et al. 2002, Xing and Sairam 2002, Xing and Sairam 2002, Putowski, Schillings et al. 2004). The element contributes significantly to promoter activity and provides a common mechanistic theme that features the E box and its cognate binding factors upstream stimulatory factor 1 (USF1) and 2 (USF2) (Figure 4A). Other promoter sequences implicated in *FSHR/Fshr* regulation include binding sites for steroidogenic factor-1 (SF-1), SMAD3 (mothers against decapentaplegic Homolog 3), E2F (transcription factor E2F), GATA-1 (GATA-binding factor 1), and ETS proteins (Heckert 2001, Kim and Griswold 2001, Levallet, Koskimies et al. 2001, Gong and McGee 2009, Brune, Adams et al. 2010).

Located just 5' to the transcriptional start sites, the E box and its binding proteins have, by far, received the greatest attention with respect to *FSHR/Fshr* transcription. USF1 and USF2 are members of the helix-loop-helix family that form both homo- and heterodimers (USF1, USF2, USF1/2, respectivity) and considerable evidence, from studies in rodents, document their role directing *Fshr* promoter activity via the E box (Shieh, Sparkes et al. 1993, Goetz, Lloyd et al. 1996, Viollet, Lefrancois-Martinez et al. 1996, Heckert, Daggett et al. 1998, Findlay and Drummond 1999, Heckert, Sawadogo et al. 2000, Griswold and Kim 2001, Heckert 2001, Rodriguez, Girones et al. 2003). In both testis and ovary, *in vitro* (electrophoretic mobility shift

assay) and *in vivo* (chromatin immunoprecipitation, ChIP) studies showed USF1, USF2 and USF1/2 bind the promoter, but dimer composition differed between males and females, with USF1/2 and USF1 favored in the testis and USF1, USF2, and USF1/2 equally matched in the ovary (Figure 1B) (Hermann, Hornbaker et al. 2008). Furthermore, *Fshr* mRNA expression and *in vivo* promoter binding evaluated in *Usf1*- and *Usf2*-null mice, revealed differences between testis and ovary in their response to loss of either USF protein (Hermann, Hornbaker et al. 2008). Thus, in testis, *Fshr* expression was unchanged by either *Usf1* or *Usf2* deletion and showed compensatory increases in promoter-bound USF homodimers (Hermann, Hornbaker et al. 2008, Viswanathan, Wood et al. 2009). In contrast, ovarian *Fshr* expression declined in both and compensatory change in homodimer binding was not indicated. Additional studies on the *Fshr* promoter in Sertoli cells, showed USF binding to the E box increases during differentiation and decreases, together with promoter activity, upon FSH treatment (Scobey, Fix et al. 2004, Viswanathan, Wood et al. 2009, Wood and Walker 2009). The predicted mediator of these FSH-induced changes is the inhibitor of DNA binding/differentiation protein, ID2, which increases with FSH treatment and inhibits both E box binding and promoter activity (Figure 4B).

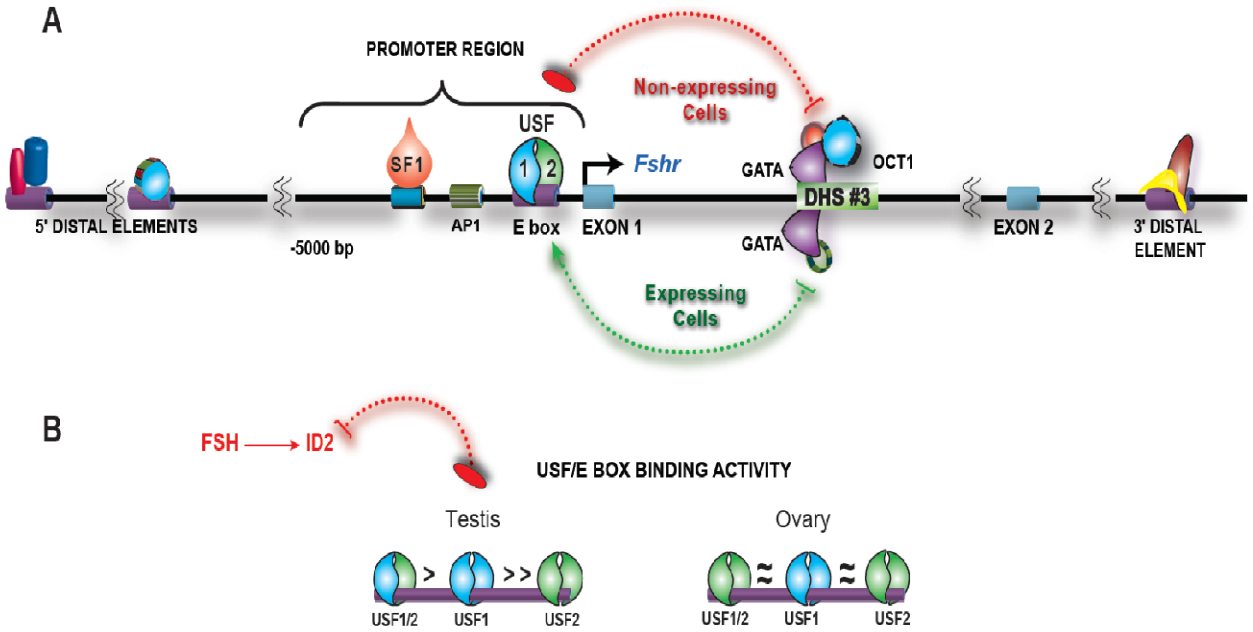


FIGURE 4: Summary of *FSHR/Fshr* transcriptional regulation. **A)** Identified elements within the *FSHR/Fshr* promoter region are depicted (rectangles) and their identified binding proteins indicated as shapes above the element: SF1 (Steroidogenic factor 1), USF (Upstream stimulatory factor), GATA binding protein (GATA), and Octamer-binding protein 1 (OCT1, POU2F1). DHS#3, noted in the first intron, is bound by a GATA-containing complex, which contains the OCT1 in non-expressing cell but not FSHR-expressing Sertoli cells. The presence of OCT1 within the DHS#3-bound GATA complex correlates with DHS#3 silencing activity observed in non-expressing cells. Also indicated are distal regulatory elements predicted by studies in transgenic mice, which failed to recapitulate *Fshr* expression from using its promoter region (up to 5.0Kbp in size) or a yeast artificial chromosome containing the entire gene plus considerable flanking sequence. **B)** Binding of USF1 and USF2 to the *Fshr* promoter occurs only in expressing (Sertoli cells and granulosa) cells. In Sertoli cells, dimer composition favors heterodimers and USF1 homodimers, while in granulosa cells, the relative composition of heterodimer and homodimer binding is similar. In Sertoli cells, FSH induces the inhibitory protein ID2, which reduces USF binding and *Fshr* expression

Despite numerous studies on the *FSHR/Fshr* promoter, no mechanism has evolved to explain the gene's remarkable cell specific expression. This deficiency in promoter specificity was also demonstrated by studies that evaluated 16 distinct transgenic mouse lines for cell-specific expression of reporters directed by either 5.0kbp or 143bp (8 lines each) of rat *Fshr* promoter sequence, none of which showed Sertoli or granulosa cell expression (Heckert, Sawadogo et al. 2000). Similar findings were reported for 1.5kbp of the human promoter (Nordhoff, Gromoll et al. 2003). The recognition that sequences beyond the 5.0Kbp promoter were required for expression, together with emerging data on regulatory elements that act from distal positions, led to studies using yeast artificial chromosomes (YACs) as transgenes, in order to define the region required for *FSHR/Fshr* specificity. The absent transgene expression in Sertoli and granulosa cells of mice carrying a YAC with the entire rat *Fshr* gene, plus bordering sequences stretching more than 50kbp 5' and 30kbp 3', further supports involvement of distal regulation and suggests regulatory regions extend well beyond the gene itself (Hermann, Hornbaker et al. 2007).

With the evidence that proper expression of *FSHR/Fshr* requires contributions from regulatory elements located outside the promoter region, most likely at significant distances, the challenge became identifying these sites within a vast amount of potential sequence. Initially, with only a small amount of available sequence, this was tackled using conventional DNase I hypersensitivity mapping to identify regions of accessible chromatin and thus potential sites of regulatory importance (Hermann and Heckert 2005). This revealed four hypersensitive sites located within a 45kbp region surrounding the first exon, three of which showed significant sequence conservation, a predictive feature of important regulatory elements (Hermann and Heckert 2005). One of these, DHS#3, was located approximately 4kb downstream in intron 1 and showed much greater sensitivity to DNase I in non-expressing cells (myoid) than expressing

cells (Sertoli), suggesting an association with gene silencing (Figure 1). Further functional characterization, using transient transfections and *in vitro* and *in vivo* binding assays, revealed important elements that attenuated gene expression and their cognate binding proteins, OCT1, GATA4, and GATA1. The studies also implicated OCT1 in selective binding to this element in order to maintain its silent state in non-expressing cells (Hermann and Heckert 2005). While the approach proved to be a valid means to identify important regulatory elements, it was evident that scanning extensive regions of the genome without better knowledge of its DNA content or genomic landscape was impractical. This has since been remedied by the infusion of genomics data that brought, not only new sequence information, but also a wealth of insight on the *FSHR/Fshr* gene and it residing landscape.

The *FSHR* gene

The initial characterization of the *Fshr* gene in 1992, which revealed a pronounced structural similarity to the *LHCGR/Lhcgr* gene that suggested the two evolved through duplication of a common ancestral gene (Heckert, Daley et al. 1992). Similarities between *FSHR/Fshr* and *LHCGR/Lhcgr* and the genes for other G-protein coupled receptors (GPCR) also suggested the predecessor for *FSHR/Fshr* and *LHCGR/Lhcgr* was formed by combining a common GPCR ancestral gene. The ancestor for these genes presumably arose by encoding the characteristic transmembrane and intracellular domains, with multiple repeated exons derived from tandem duplications of a module for a leucine-rich motif, a featured attribute of the extracellular domain of glycoprotein hormone receptors (Heckert, Daley et al. 1992). Formation of the gonadotropin receptor genes through tandem duplication of an ancestral gene was further substantiated by recent genomics data that show they are arranged in tandem on chromosomes from nearly all annotated Tetrapoda genomes (Heckert, Daley et al. 1992, Montgomery, Tate et al. 1995,

Chauvigne, Tingaud-Sequeira et al. 2010). While the two ancestral descendents, *FSHR/Fshr* and *LHCGR/Lhcgr,* differ by one exon (10 for *FSHR/Fshr* and 11 for *LHCGR/Lhcgr*), the coding scheme is largely the same, with the carboxy-terminal, transmembrane-intracellular domain encoded by the last exon and the amino-terminal, extracellular domain by all preceding exons (Koo, Ji et al. 1991, Tsai-Morris, Buczko et al. 1991, Heckert, Daley et al. 1992). Exons that partition the receptors' extracellular domains also share a repeated structure that delineates seven of the leucine-rich motifs into exons 2-8 and two into exon 9 (Figure 5).

However, despite knowing the structure of *FSHR/Fshr* for nearly two decades, it was not until the various genome-sequencing projects greatly expanded the available sequence data that there was accurate knowledge of its size or chromosome habitat. Now, *FSHR/Fshr* chromosome locations and annotated sequences are reported for more than 40 vertebrate species through the University of California, Santa Cruz Genome Browser (http://genome.ucsc.edu/) and similar web sites. In human, rat, and mouse *FSHR/Fshr* are located on chromosomes 2, 6 and 17, respectively, and span roughly 200kb, a size much larger than originally predicted (Heckert, Daley et al. 1992, Huhtaniemi, Eskola et al. 1992, Rousseau-Merck, Atger et al. 1993, Gromoll, Ried et al. 1994, Rhead, Karolchik et al. 2010) (Figure 2). This wealth of sequence information also disclosed several defining features of *FSHR/Fshr* and its surrounding neighborhood that indicate its associated regulatory environment is strongly influenced by evolutionary constraints that retain regulatory sequences directing distally-located genes (Hermann and Heckert 2005, Hermann, Hornbaker et al. 2007). This includes the tandem placement of *FSHR/Fshr* between it closest 5' neighbor *NRXN1/Nrxn1*, which encodes the synaptic neuronal adhesion protein NEUREXIN 1, and its closest 3' neighbor *LHCGR/Lhcgr*, the gene encoding the luteinizing hormone receptor, to form a highly conserved syntenic block, a feature identified by several

studies to indicate evolutionary constraints intended to preserve the relative positions between noncoding sequences and their target genes (Hermann and Heckert 2005, Engstrom, Ho Sui et al. 2007, Kikuta, Laplante et al. 2007, Akalin, Fredman et al. 2009). The genomic relationship between *FSHR/Fshr* and *LHCGR/Lhcgr*, when examined together with their similar expression profiles and functions in gonadotropin signaling, raises the intriguing possibility that function of one or both genes depends on their relative positions. In considering potential mechanisms, a link between transcriptional regulation and the shared synteny can be readily conceived if *FSHR/Fshr* and/or *LHCGR/Lhcgr* expression depends on regulatory sequences that would be lost if the genes separated. This implies the sequences are either; 1) required for expression of one gene and reside within the locus defined by the other (gene plus regulatory domain) or 2) required by both genes (i.e. identical or overlapping sites) and function depends on their position relative to *FSHR/Fshr* and *LHCGR/Lhcgr*. The second mechanism can be expanded to include the use of sequences for concurrent gene regulation, which, for co-regulation of *FSHR/Fshr* and *LHCGR/Lhcgr* activity is limited to granulosa cells of growing follicles, the only cells that express both genes. So despite similar gonad-specific expression profiles, the receptors are largely confined to distinct cell types, where any simultaneous use of a regulatory sequence will require opposite transcriptional effects on the two genes. Thus, most gonadotropin-responsive cells do not co-express *FSHR/Fshr* and *LHCGR/Lhcgr* and FSH response is restricted to Sertoli and granulosa cells of the testis and ovary, respectively, and LH response is confined to testicular Leydig cells and ovarian theca, maturing granulosa, and luteal cells (Camp, Rahal et al. 1991, Peng, Hsueh et al. 1991, Kaminski, Gawronska et al. 2000, Zhang, Shi et al. 2001, Dickinson, Stewart et al. 2009).

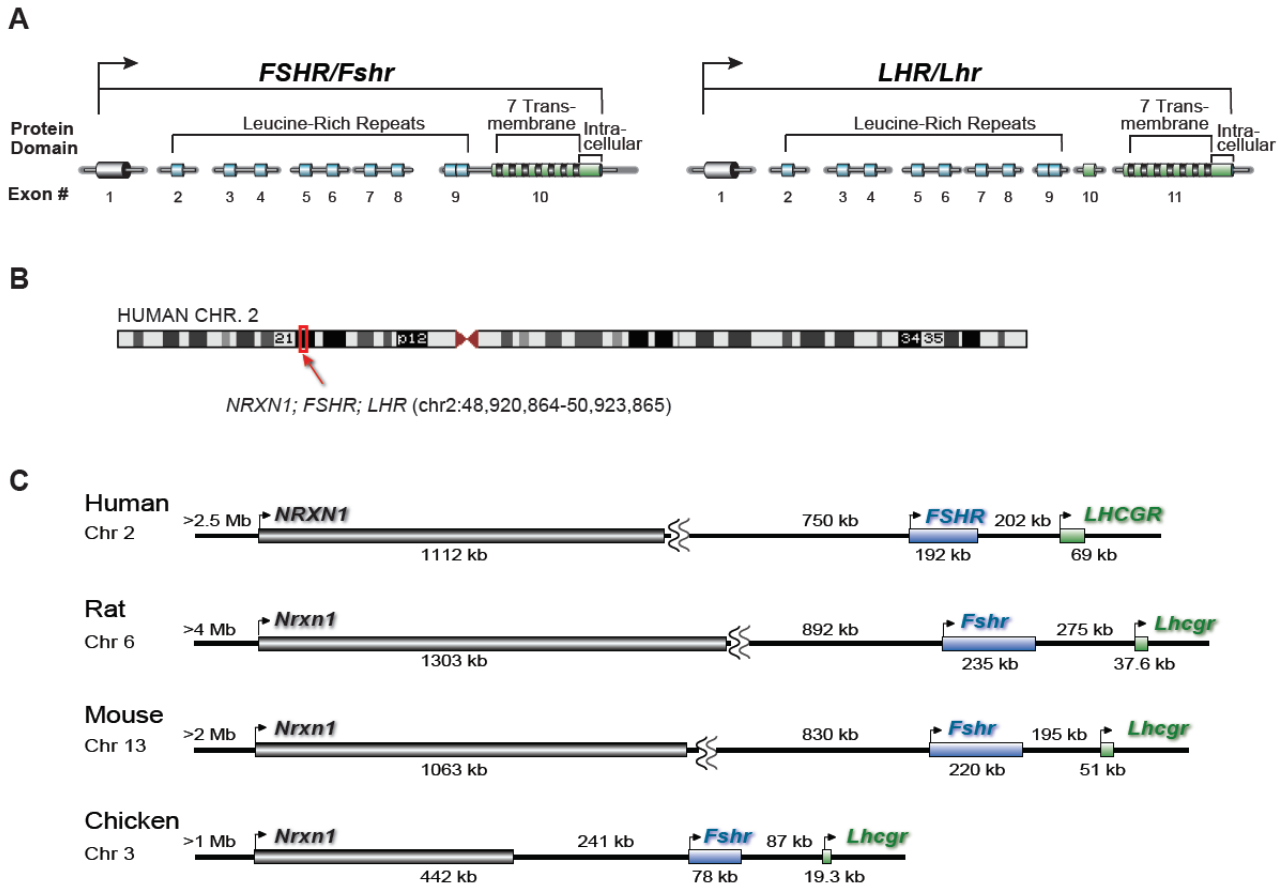


FIGURE 5: Organization of *FSHR/Fshr* and its genomic environment. A) Structure of *FSHR/Fshr* and *LHCGR/Lhcgr*, depicting the exon distribution of the receptor's protein domains. Exons are indicated by rectangles, intervening regions by lines, and transcriptional direction by arrows. B) Position of the *NRXN1;FSHR;LH/CGR* syntenic region on human chromosome 2 (red box indicated by arrow). C) Syntenic region of *NRXN1/Nrxn1, FSHR/Fshr*, and *LHCGR/Lhcgr* from human, rat, mouse, and chicken, with sizes indicated for the genes and intervening regions. Genes are indicated by rectangles, intervening regions by lines, and transcriptional direction by arrows.

To date, there are no identified regulatory elements shared by *FSHR/Fshr* and *LHCGR/Lhcgr* or located within the other's defined locus. However, our knowledge of their transcription is represented almost entirely by promoter characteristics and, therefore, insufficient to conclude the elements do not exist. Regardless, promoter characteristics do suggest divergent regulation, as they differ significantly with respect to sequence, identified regulatory elements, and activity in transgenic mice, i.e. *LHCGR/Lhcgr*, but not *FSHR/Fshr*, promoter directs cell-specific expression *in vivo* (Heckert and Griswold 1991, Heckert, Daley et al. 1992, Tsai-Morris, Xie et al. 1993, Tsai-Morris, Geng et al. 1994, Tsai-Morris, Geng et al. 1995, Goetz, Lloyd et al. 1996, Heckert, Daggett et al. 1998, Hamalainen, Poutanen et al. 1999, Hamalainen, Poutanen et al. 2001, Kim and Griswold 2001, Apaja, Aatsinki et al. 2005, Hermann, Hornbaker et al. 2007). On the other hand, there are a few common promoter features worth noting. First is the core structure, which, for both, lacks a TATA box and has multiple transcription initiation sites within a region similar to an initiator element (Tsai-Morris, Buczko et al. 1991, Heckert, Daley et al. 1992, Dufau, Tsai-Morris et al. 1995, Goetz, Lloyd et al. 1996, Juven-Gershon, Hsu et al. 2008). Second is the ubiquitous nature of each promoter's main functional element(s), which depend largely on non-specific, widely expressed but distinct, transcription factors; *LHCGR/Lhcgr* with two GC-rich sequences that bind SP1 and SP3 and *FSHR/Fshr* with a single E box bound by USF1 and USF2 (Tsai-Morris, Geng et al. 1995, Goetz, Lloyd et al. 1996, Geng, Tsai-Morris et al. 1999, Heckert, Sawadogo et al. 2000). While these similarities may reflect the common ancestry and/or association via an undisclosed fundamental mechanism, the predominant promoter characteristics emphasize functional divergence of key regulatory features that highlight the promoter for *LHCGR/Lhcgr* and distal regulatory sites for *FSHR/Fshr*.

Another notable feature is the large intergenic distances between *FSHR/Fshr* and its neighbors, which span more than 750kbp on its 5‘side and 200kbp on its 3‘side (Hermann and Heckert 2005). The uninhabited region between *NRXN1/Nrxn1* and *FSHR/Fshr* is characteristic of a gene desert, or long genomic stretch devoid of protein-coding sequences or other obvious biological functions. Such regions, when located within a conserved syntenic block, are associated with areas of enhanced sequence conservation, suggesting the content was under evolutionary pressure to retain functional and contextual information of resident elements (Venter, Adams et al. 2001, Ovcharenko, Loots et al. 2005, Akalin, Fredman et al. 2009). While evidence shows not all gene deserts have measurable activity, numerous risk loci and regulatory elements are documented within these regions, confirming their importance to the genome (Lodder, Eussen et al. 2009, Xu, Tsumagari et al. 2009, Kiltie 2010). Comparative sequence analysis has been consistently used in genome-wide studies to evaluate evolutionary constraint as a means to understanding genome structure, biological function, and evolution (Waterston, Lindblad-Toh et al. 2002). The initial whole-genome sequence comparisons between mouse and human provided considerable insight and estimated 5% of the human genome was conserved over 70-100 million years (Lander, Linton et al. 2001, Venter, Adams et al. 2001, Dermitzakis, Reymond et al. 2002, Mural, Adams et al. 2002, Waterston, Lindblad-Toh et al. 2002, Pennacchio 2003). What was remarkable in this finding was that only 1/3 of the conserved sequences were located in coding regions. Thus, the genome’s non-coding sector represents the largest portion under evolutionary selection, which suggests there is considerable functional information within conserved non-coding sequences, aka evolutionary conserved regions (ECRs) (Waterston, Lindblad-Toh et al. 2002, Dermitzakis, Reymond et al. 2005).

Expansion of genome sequence data and species representation added significantly to the power of cross-species sequence comparison, the genomic landscape and sites of potential biological functions. These breakthroughs have led to identification of many non-coding sequences involved in gene regulation, including enhancers, insulators, silencers and matrix attachment regions (Glazko, Koonin et al. 2003, Koonin 2003, Nobrega, Ovcharenko et al. 2003, de la Calle-Mustienes, Feijoo et al. 2005, Woolfe, Goodson et al. 2005, Pennacchio, Ahituv et al. 2006, Prabhakar, Poulin et al. 2006, Visel, Prabhakar et al. 2008). A connection between sequence conservation and regulatory sequences was nicely illustrated in a recent study that used ChIP with massively parallel sequencing to map DNA sites linked to p300, a co-regulator for many transcription factors (Visel, Blow et al. 2009). Results from mouse embryonic tissues (forebrain, midbrain and limb) showed p300 highly enriched at sites containing conserved non-coding sequence and, of those evaluated, nearly all were functional. The study also demonstrated that most p300-associated sites/enhancers were located at least 10kbp from a potential target gene, which suggests distal elements are commonly involved in transcriptional control. Direct evidence for long-distance gene regulation is found in numerous studies that demonstrate transcriptional effects from distant sequences on specific target genes, such as those for β-globin, IFγ, SOX9, GATA3, FGF4, IL-10, and CD69 (Martinez-Jimenez, Gomez-Lechon et al. 2005, Bejerano, Lowe et al. 2006, Schoenborn, Dorschner et al. 2007, Vazquez, Laguna et al. 2009). The significance of non-coding regulatory sequences to human disorders has also gained recognition with reports linking them to diseases and developmental disorders, as noted in cases of X-linked deafness, preaxial polydactyl, campomelic dysplasia, sex reversal, postaxial polydactyly (de Kok, Merkx et al. 1995, Bishop, Whitworth et al. 2000, Jamieson, Perveen et al. 2002, Epstein 2009, Lodder, Eussen et al. 2009). In the case of preaxial polydacyl, the reported

mutation was located within a regulatory element approximately 1Mbp from its target gene that emphasizes not only the remarkable linear distance in which the transcriptional signal passes but also the importance of nuclear architecture in positioning regions that collaborate in the transcriptional signal (reviewed in (Schneider and Grosschedl 2007)).

Identification of new regulatory elements

The search for new regulatory elements is clearly facilitated by predictions based on sequence conservation. However, the evidence is also clear that sequence conservation does not detect all regulatory elements nor provide any assurance that the predicted sites are functional, and thus, most effective when used in conjunction with other corroborating techniques, particularly ones that can assess chromatin changes linked to transcriptional activity (Nobrega, Ovcharenko et al. 2003, Dermitzakis, Reymond et al. 2005, Rizzolio, Bione et al. 2008, D'Haene, Attanasio et al. 2009, Liska, Snajdr et al. 2009, Vazquez, Laguna et al. 2009, Visel, Zhu et al. 2010). In studies to identify regulatory sequences that direct *FSHR/Fshr* transcription, results from transgenic and promoter studies shifted the experimental focus away from the promoter to the region encompassing all of *FSHR/Fshr* and its adjoining intergenic regions, which required a new set of tools and resources that fortunately evolved from the collection of genome sequences and efforts to understand their content. Initial reports on the *FSHR/Fshr* conservation profile compared human and rat sequences by direct pairwise comparison and analysis of precompiled LAGAN alignments through the web-based VISTA genome browser (Hermann and Heckert 2005). This revealed over 150 conserved sites, which, when matched together with DNase I hypersensitivity data, was instrumental in the identification of an important silencing region in the first intron (discussed above). However, it was also evident that greater constraints were required to improve functional prediction and therefore, once more distant genomes (e.g. chicken) were

available and included in computations, ECRs having greater predictive power could be distinguished and seven of the most conserved were selected for functional testing by transient transfection (Hermann, Hornbaker et al. 2007). With continuous enhancements in genome data and resources that improve regulatory element prediction, the number of selected sites has grown to more than twenty, which includes the original seven ECRs. These sequences were identified using the human genome and the UCSC Genome Browser (http://genome.ucsc.edu/) to reveal highly conserved (vertebrate consensus and net alignment with chicken as the target), non-coding sequences with a significant regulatory potential score (7X regulatory potential; (Waterston, Lindblad-Toh et al. 2002, 2004, Gibbs, Weinstock et al. 2004, Havlak, Chen et al. 2004, Kolbe, Taylor et al. 2004, King, Taylor et al. 2007)). Figure 6 provides an example of the *FSHR* locus spanning from *LHCGR* to *NRXN1* that was modified from the UCSC Genome Browser results to show the top predicted sites (marked ECRs) for *FSHR/Fshr* regulatory sequences and the key features used in their selection.

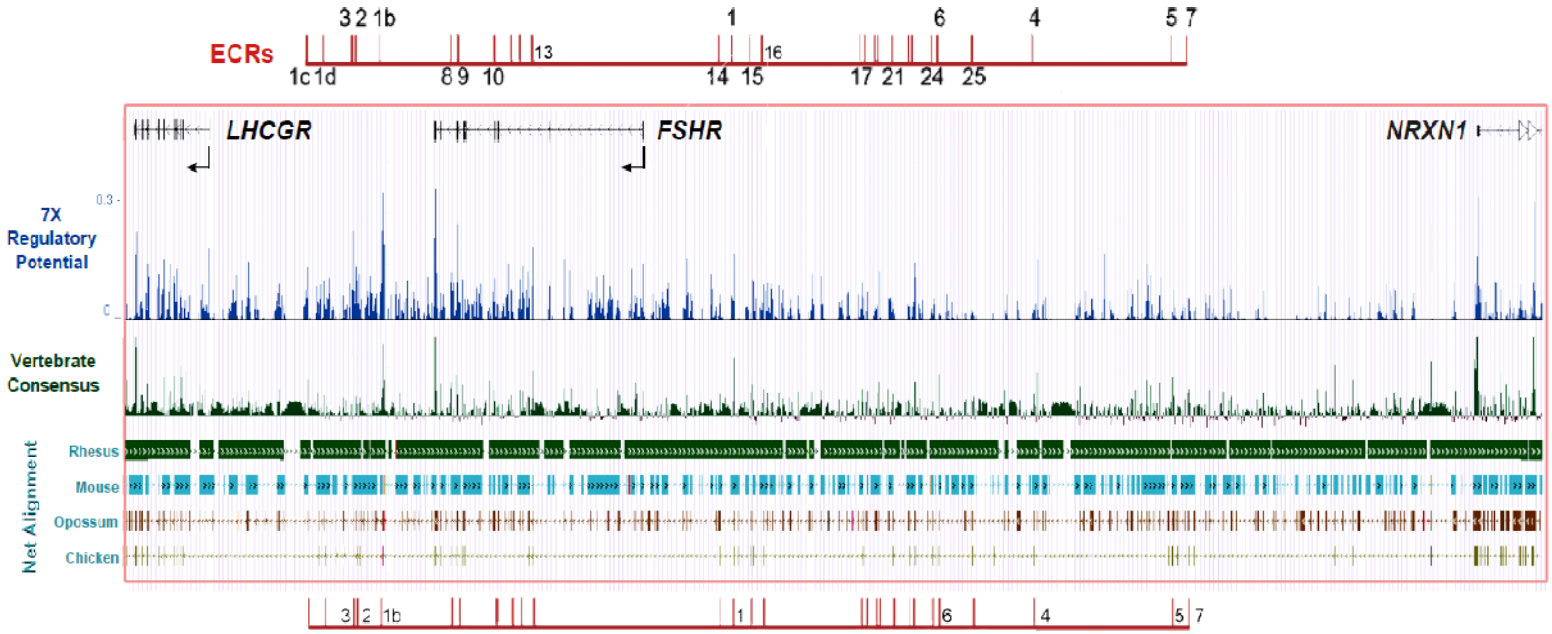


FIGURE 6: Annotated region of human *FSHR* and neighboring genes *NRXN1* and *LHCGR* from the UCSC Genome Browser (http://genome.ucsc.edu/). With human as the reference genome, noncoding evolutionary conserved regions (ECRs, top) were selected for their potential regulatory activity using sequence conservation, both vertebrate consensus and net alignment (chicken as the target), and 7X regulatory potential. Vertebrate consensus represents alignments of 44 vertebrate species generated using multiz and other tools and measurements of evolutionary conservation determined by phastCons and phyloP (Siepel, Bejerano et al. 2005, Pollard, Hubisz et al. 2010). The 7X regulatory potential track represents scores computed from human, chimpanzee, macaque, mouse, rat, dog, and cow by comparing short alignment pattern frequencies between known regulatory elements and neutral DNA (Waterston, Lindblad-Toh et al. 2002, Gibbs, Weinstock et al. 2004, Havlak, Chen et al. 2004, Kolbe, Taylor et al. 2004, King, Taylor et al. 2007). The net alignment tracks show the best mouse/human or chicken/human chains for every part of the human genome, with ungapped alignments represented by boxes and gaps by lines. Annotated sequences are from the chicken May 2006 (WUGSC 2.1/galGal3) (galGal3) assembly, mouse July 2007

Summary

The evidence to date shows that sequences directing *FSHR/Fshr* expression lie far from the start of transcription in a regulatory environment without defined boundaries, which complicates their detection using standard molecular approaches. While computational genomics has helped narrow the search, limitations due to false positives and undetected sequences caution its use without additional methods to substantiate the data (Giresi, Kim et al. 2007). Fortunately, many technologies have adapted to the influx of sequence data by developing high-throughput and genome-wide strategies. Two such strategies offer considerable promise for regulatory element identification and that the *FSHR/Fshr* transcriptional mechanism is within reach. Both strategies reveal chromatin signatures featured in regulatory sequences; one identifies sequences bound to modified histones linked to transcriptional activity by chromatin immunoprecipitation, the other identifies open regions of chromatin, similar to DNase I hypersensitivity, by formaldehyde associated identification of regulatory elements. (Wu, Smith et al. 2006, McGaughey, Stine et al. 2009). Implementation of such strategies together with comparative genomics will significantly enhance the probability of relevant sequence identification and the mechanistic understanding of the regulatory landscape. When combined with high-throughput strategies, such as DNase-Chip, high-density tiling arrays and next generation sequencing, to canvass the genome without the bias of conservation, additional insight is likely on mechanisms that employ non-conserved regulatory elements and possible contributions to species-specific regulatory features (Crawford, Davis et al. 2006, Roh, Wei et al. 2007, Chen, Lin et al. 2008).

Chapter 2

CTCF and transcription regulation of *Fshr* in rat granulosa cells

Abstract

The anterior pituitary in response to gonadotropin releasing hormone (GnRH) stimulates the release of follicle stimulating hormone (FSH). FSH targets the gonad by binding specifically to its cell-surface receptor, FSHR, present on Sertoli cells of the testis and granulosa cells of the ovary. FSHR hence forms a bridge for FSH action and plays a crucial role in mediating gonadal development and fertility. Prior studies investigating the transcriptional regulation of *Fshr* focused on the promoter region and identified a number of crucial elements required for the proper transcription of the gene. However, *in vitro* transfection studies and *in vivo* YAC transgenic mice soon identified that regions at considerable distance from the promoter are needed for correct spatio-temporal regulation. This thought is supported by numerous studies showing that many regulatory elements reside far from the target gene. Comparative genomics coupled with CTCF binding prediction tools identified multiple highly conserved regions both 5' and 3' to the *Fshr* gene that indicate binding sites for the versatile CCCTC-binding factor (CTCF). CTCF is a eleven-zinc finger protein known to play varied roles in genome regulation, including transcription, chromatin insulation, and high order chromatin structure. To identify the role of CTCF in *Fshr* regulation, granulosa cells harvested from estrogen treated female rats were depleted of CTCF by siRNA transfection, leading to a two-fold increase in *Fshr* mRNA expression while transcription of the nearby *Lhcgr* gene was unaffected. These data indicate that CTCF either by itself or in conjunction with other protein complexes might play a role in transcriptional regulation of the *Fshr* gene.

Introduction

Gonadotropin releasing hormone (GnRH), synthesized by the peptidergic neurons of the hypothalamus, binds to receptors present on the surface of the gonadotroph cells of the anterior pituitary gland, initiating the synthesis and secretion of Lutenizing hormone (LH) and Follicle stimulating hormone (FSH). FSH is a heterogeneous glycoprotein that recognizes and binds to FSHR, a G-protein coupled receptor present on the Sertoli cells of the testis, granulosa cells of the ovary, and osteoclasts of the bone (Richards and Midgley 1976, Heckert and Griswold 1991, Dankbar, Brinkworth et al. 1995, Sun, Peng et al. 2006). Binding to the receptor elicits a number of cellular activities, especially the activation of adenyl cyclase, increased cAMP levels, activation of PKA and phosphorylation of a number of transcriptional activators (Flores, Veldhuis et al. 1990, Tena-Sempere, Manna et al. 1999, Seger, Hanoch et al. 2001). Since FSH acts exclusively through FSHR, greater understanding of its regulation can provide insight on how the cells respond to the hormone, as changes in the receptor level will influence response. Furthermore, absence of FSH or its receptor is known to cause arrest in folliculogenesis and infertility(Huhtaniemi and Themmen 2005). Over the past two decade studies conducted in rat, mouse, porcine, and human have identified a number of regulatory elements and proteins that are crucial for the proper transcription of *Fshr* (Huhtaniemi, Eskola et al. 1992, Gromoll, Dankbar et al. 1994, Goetz, Lloyd et al. 1996, Sairam and Subbarayan 1997, Simoni, Gromoll et al. 1997, Heckert, Daggett et al. 1998, Kim and Griswold 2001). However, transgenic studies directed by either 5.0kbp or 143bp of rat *Fshr* promoter sequence were unable to replicate Sertoli and granulosa cell specific expression (Heckert, Sawadogo et al. 2000). Furthermore, transgenic mice carrying yeast artificial chromosomes containing 413-kbp region of the rat *Fshr* gene and its encompassing regions were found to be insufficient to direct proper spatial and temporal

expression (Hermann, Hornbaker et al. 2007). These data suggest that distal *cis*-acting elements present outside this region were required for proper specific expression.

Evolutionary conserved regions of DNA indicate that they are under pressure to retain functional significance. These highly conserved non-coding regions imply that they might play a role in gene regulation (Bejerano, Pheasant et al. 2004, Shin, Priest et al. 2005, Pennacchio, Ahituv et al. 2006). They are often home to regulatory elements involved in spatial and temporal gene regulation (Maston, Evans et al. 2006). Among these regulatory elements is the CCCTC-binding factor (CTCF), a highly conserved and versatile regulatory factor that plays a varied role in genome regulation. CTCF can act as enhancer blocking, gene activator/repressor, hormone-responsive silencing, long-range chromatin interactions, X-chromosome inactivation, genomic imprinting and in regulating chromatin architecture (Phillips and Corces 2009, Ohlsson, Bartkuhn et al. 2010). Recently microarray data analysis of CTCF depleted mice identified *Fshr* was among the genes that were up regulated (Wan, Pan et al. 2008). The current study explores the possible CTCF binding sites in the highly conserved regions identified by computational analysis and looks at the regulation of *Fshr* in rat granulosa cells depleted of CTCF.

Materials and Methods

Animal Use.

All experiments using animals were approved by the Institutional Animal Care and Use Committee of the University of Kansas Medical Center and performed in accordance with National Institute of Health guide for the care and use of laboratory animals.

Rat granulosa Cell culture

23 day old Immature Sprague Dawley female rats (Harlan, Indianapolis, IN) were injected subcutaneously with 1.5mg 17β-estradiol once daily at 24 ,25, and 26 days of age. Ovaries were isolated on day 27 and granulosa cells harvested as previously described (Alliston, Maiyar et al. 1997). Briefly, ovaries were harvested, cleaned of fat pads and placed into the M199 collection media containing M199 media (Sigma M2520), 10mM HEPES and 0.2% BSA. The collection media was then slowly aspirated and the ovaries were incubated in 1ml of Sucrose solution media (M199 collection media, 1.8mM EGTA and 0.5M Sucrose) for 15 minutes at 37˚C. The ovaries were rinsed three times with M199 collection media and poked gently with a 22-gauge needle to isolate the granulosa cells. The cells were centrifuged, counted and 100,000 cells were plated in each well of a six well plate coated with fibronectin and grown with Dulbecco modified Eagle medium/F12 (10% fetal bovine serum, 1% gentamicin).

siRNA transfection

Various amounts of CTCF siRNA and scrambled siRNA control (siGenome, rat CTCF, NM_031824, Dharmacon) were transfected into each well with 5µl of Lipofectamine RNAimax (Invitrogen). The transfection media was aspirated 6 hours later and replaced with rat granulosa media and the cells were harvested 48 hours post transfection.

RNA isolation and quantitative real time PCR

Total RNA was isolated using Trizol according to manufactures protocol (Life Technology). Total RNA (100ng) was reverse transcribed using iScript cDNA synthesis kit (Bio-Rad) according to manufactures protocol. 1ul of the of the cDNA was assayed using the SYBR Green PCR master mix (Applied Biosystems) and primers for rat CTCF (F:TGCCAGTGTAGAAGTCAGCAAATT; R: TGTATGTGTCCCTGCTGGCATA); *Fshr* (F: GCCAAGACAGCAAGGTGACA; R: GAGCACAAACCTCAGTTCAATGG); *Lhcgr* (F: GGTCGCCACGCTGACCTA; R: TCTGTTCTTTCTTCGGCAAATTC) and normalized to the internal control L7 (F: GCTAGGATGGCGAGGAAAGC; R: TGATACCTCGGATTCTGATGACA). Real time was carried out in a 7900HT Sequence Detection real-time analyzer (Applied Biosystems) as described (Gifford, Racicot et al. 2007). Samples and negative control (Hatano, Takayama et al.) were run in triplicate and gene expression was quantified using the delta delta C(T) method in comparison with L7.

Results

Computational identification of conserved regions and CTCF binding sites

To identify distal conserved regions, *Fshr* locus and its flanking areas encompassing a region of 1.4 Mb was analyzed for conservation using the UCSC genome browser (http://genome.ucsc.edu/cgi---bin/hgGateway) and associated software (King, Taylor et al. 2005, King, Taylor et al. 2007). Potential regulatory elements were identified by using human *FSHR* as the base genome and Gallus gallus (chicken) as the most distant. In addition to alignment, 7X regulatory potential was employed to increase predictive measure of regulatory regions. 7X regulatory potential displays potential predictive scores by aligning the genome of human, chimpanzee, macaque, mouse, rat, dog, and cow and comparing the frequencies of short alignment patterns between known regulatory elements and neutral DNA, allowing for the identification of new putative regulatory element(Kolbe, Taylor et al. 2004). This analysis led to the identification of 30 evolutionary conserved regions (ECR) of which seventeen were mapped upstream to *Fshr*, seven to the intergenic region, six downstream (Table 1).

These sequences were then analyzed in the CTCF database CTCFBSD 2.0 using the CTCFBS prediction tool (http://insulatordb.uthsc.edu/)(Bao, Zhou et al. 2008, Ziebarth, Bhattacharya et al. 2013). CTCF employs different combinations of its zinc finger proteins to bind to divergent regulatory DNA sequences. Recent studies have identified the core regulatory regions bound by CTCF and these motifs were assigned position weight matrices (PWM). Six such PWM were identified and the CTCFBS prediction tool scores these PWM in the query sequence. Motifs with an overall high positive score indicating favorable match were chosen as probable CTCF binding sites(Ziebarth, Bhattacharya et al. 2013). Of the 30 conserved regions, ECR1, ECR1d, ECR1f,

ECR2, ECR5, and ECR15 had overall positive scores greater than 3, indicating high probability of CTCF binding (Table 2).

Table 1: Genomic location of ECRs in *Rattus norvegicus* (Rat) genome.

ECR	Chromosome Location	Size (bp)
ECR 1	Chr6: 23036890---23037477	588
ECR 1c	Chr6: 22585376---22586088	713
ECR 1d	Chr6: 22597052---22597262	211
ECR 1f	Chr6: 22614761---22616669	1909
ECR 1g	Chr6: 22590321---22590502	182
ECR 2	Chr6: 22626739---22629962	3224
ECR 3	Chr6: 22596865---22597324	460
ECR 4	Chr6: 23287153---23287391	239
ECR 5	Chr6: 23423202---23423600	399
ECR 6	Chr6: 23219667---23220065	399
ECR 7	Chr6: 23437816---23438244	429
ECR 8	Chr6: 22771058---22771652	595
ECR 9	Chr6: 22776439---22777106	668
ECR 9b	Chr6: 22785969---22786300	332
ECR 10	Chr6: 22813283---22813805	523
ECR 11	Chr6: 22826398---22826718	321
ECR 12	Chr6: 22826285---22826757	473
ECR 13	Chr6: 22845802---22846436	635
ECR 14	Chr6: 23024900---23025427	528
ECR 15	Chr6: 23048424---23049218	795
ECR 16	Chr6: 23055601---23056494	894
ECR 17	Chr6: 23148636---23149418	783
ECR 18	Chr6: 23153766---23154371	606
ECR 19	Chr6: 23167660---23168222	563
ECR 20	Chr6: 23173974---23174670	697
ECR 21	Chr6: 23180610---23181272	663
ECR 22	Chr6: 23188053---23188745	693
ECR 23	Chr6: 23195435---23196040	606
ECR 24	Chr6: 23212514---23213040	527
ECR 25	Chr6: 23259302---23260347	1046

Table 2: ECRs predicted to have CTCF binding sites

Input Sequence	**Motif Sequence**	**Motif Start Location**	**Motif Length**	**Score**
ECR 1	CACCATCAGCTGCC	1104	14	7.5
ECR 1d	AAAACAATAATAGGGATGTG	165	20	22.9
ECR 1f	TGGCAAGCAGAGGGGAGTCT	44	20	13.7
ECR 2	TTCCATCTGCTGGA	98	14	15.0
ECR 5	AGCCATCTCCTGGC	69	14	14.3
ECR 15	TAGCCACAGGGTGGCATTC	139	19	16.0
ECR 19	GGAACAGCC	512	9	10.7

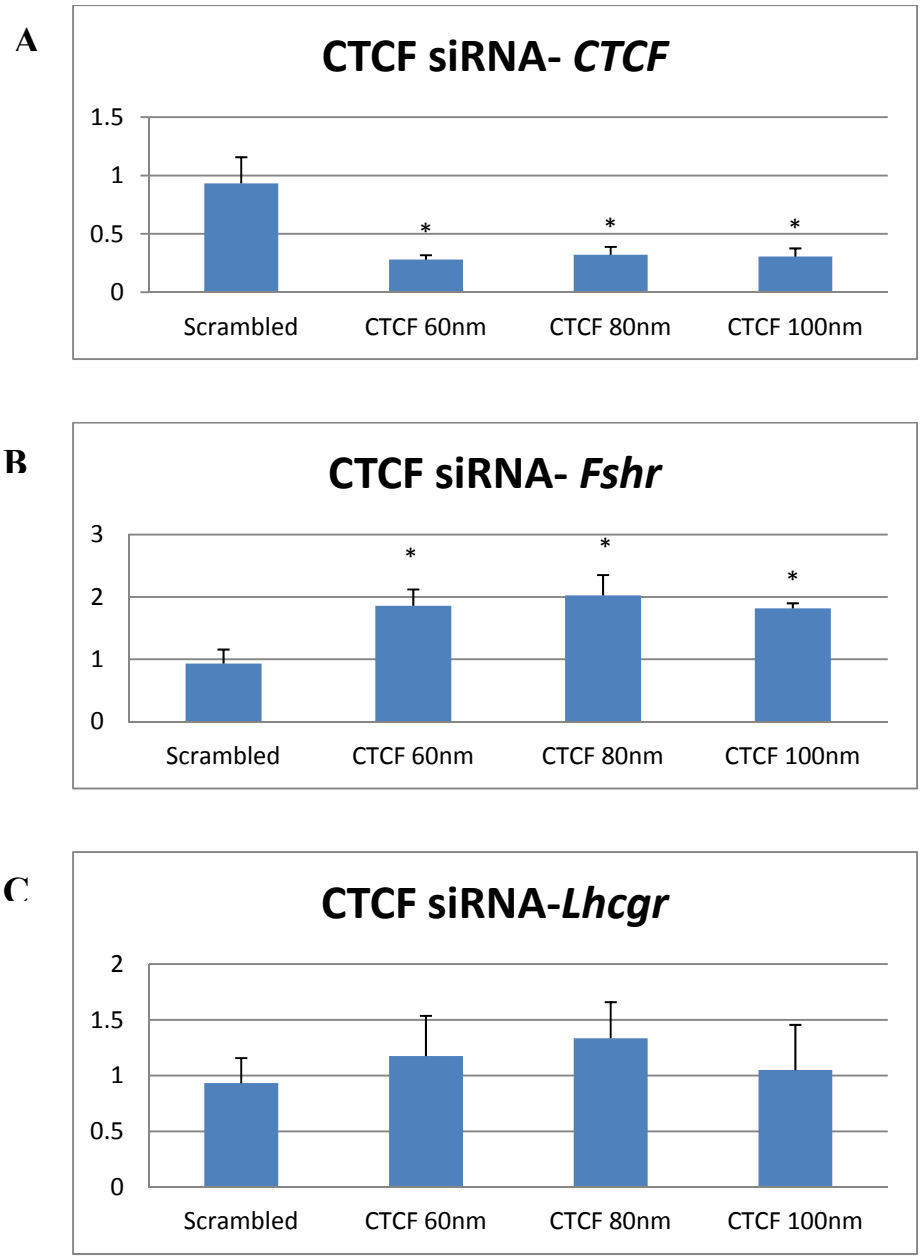


Figure 1: Expression of *Fshr* and *Lhcgr* in estrogen treated rat granulosa cells depleted of CTCF. Granulosa cells harvested from estrogen treated immature female rats were transfected with various amount of CTCF siRNA 60nm, 80nm, and 100nm. RNA was isolated 48 hours post transfection, reverse transcribed, and real time was done using primers against CTCF ; *Fshr*; *Lhcgr* and normalized to the internal control L7. CTCF was downregulated by approximately 70% (A), and *Fshr* levels was found to be increased roughly two-fold (B). No change was seen in *Lhcgr* levels (C). * $p < 0.05$

Knockdown of CTCF in rat granulosa cells increases transcription of Fshr

To test if CTCF plays a role in regulation of *Fshr*, rat granulosa cells were transiently transfected with CTCF siRNA or control siRNA and qPCR was performed using primers specific for each gene (Figure 1). CTCF was downregulated by approximately 70% and the results showed that granulosa cells in which CTCF was knocked down, expression of *Fshr* increased two-fold, while the levels of *Lhcgr* remained unchanged. This data indicate that CTCF can modulate the level of *Fshr* mRNA.

Discussion

Following the initial characterization of the *Fshr* gene, a number of studies have identified regulatory elements at the promoter and its proximal regions (Heckert, Sawadogo et al. 2000, Heckert and Griswold 2002). These studies while highlighting important factors required for proper transcription of the *Fshr* gene, were unable to explain the mechanism for the genes cell specific expression (Heckert, Sawadogo et al. 2000, Nordhoff, Gromoll et al. 2003). In a bid to identify regions required for the cell specific expression, efforts were made encompassing the *Fshr* promoter and its surrounding regions, this body of work summarized that *Fshr* belonged to the class of genes that employed distal regions for transcriptional control. Previous studies from our laboratory identified seven such distal ECR regions, which were shown to possess transcriptional activity in vitro (Hermann, Hornbaker et al. 2007). With the advent of the UCSC genome browser and inclusion of genome sequences from a number of species, ranging from human to chicken, we were able to identify thirty ECRs, which were highly conserved indicating that these sequences likely harbor regulatory elements. Among the distal regulatory elements identified to play a role in distal regulation is the ubiquitous, multivalent protein CTCF that binds

to multiple DNA sequences by using different combinations of the twelve zinc finger proteins (Filippova, Fagerlie et al. 1996). Earlier work in CTCF depleted transgenic mouse using microarrays a number of misregulated genes, among which *Fshr* was found to be upregulated (Wan, Pan et al. 2008). Building on this finding and identification of highly conserved regions, we utilized the CTCF binding site prediction tool. Among the 30 ECR regions tested, seven were found to have overall high positive scores indicative of probable CTCF binding. Among the ECR regions predicted to be bound by CTCF, ECR1d, ECR1f, and ECR2 are located 3' to the *Fshr* gene, while ECR1, ECR5, ECR15 and ECR 19 are 5' to the gene. The intergenic location of these predicted binding sites corresponds to a study done by Kim et al, who mapped global distribution of CTCF and reported 46% of CTCF binding sites to be intergenic, similar to where our predicted CTCF bound ECR regions are located (Kim, Abdullaev et al. 2007, Jothi, Cuddapah et al. 2008).

To analyze if CTCF does play a role in regulating *Fshr*, CTCF was knocked down in rat granulosa cells leading to relief of inhibition and increase of *Fshr* transcription by two-fold. This data indicates that CTCF binds within the *Fshr* locus, possibly at the predicted regions, inhibiting transcription, thus acting as a repressor. CTCF is a well-known repressor of gene transcription as documented in the chicken-*myc* gene and *hTERT* transcription (Filippova, Fagerlie et al. 1996, Renaud, Loukinov et al. 2005). It is possible that *in vivo*, CTCF activity keeps *Fshr* transcription level in check, optimizing binding of FSH to the receptor thus limiting ovarian response to gonadotropin stimulation. CTCF could mediate its repressor activity by homodimerizing as shown in Figure 2, isolating enhancers, and bound activators, physically blocking the interaction between these activator elements and the *Fshr* promoter. CTCF plays a major role in transcriptional activity, silencing, insulation and imprinting control (Phillips and Corces 2009,

Ohlsson, Bartkuhn et al. 2010). Concerning *Fshr*, CTCF possibly keeps *Fshr* regulation in check and thus indirectly regulating the delicate balance of folliculogenesis and oocyte development.

In summary, computational analysis identified putative CTCF binding sites within highly conserved regions. Knockdown of CTCF relieved repressor activity on *Fshr* in granulosa cells, thus establishing that CTCF could regulate transcription of *Fshr.* Future experiments designed to identify CTCF binding regions and the chromatin architecture both in the presence and absence of CTCF will reveal the role played by this protein in regulating transcription of *Fshr*.

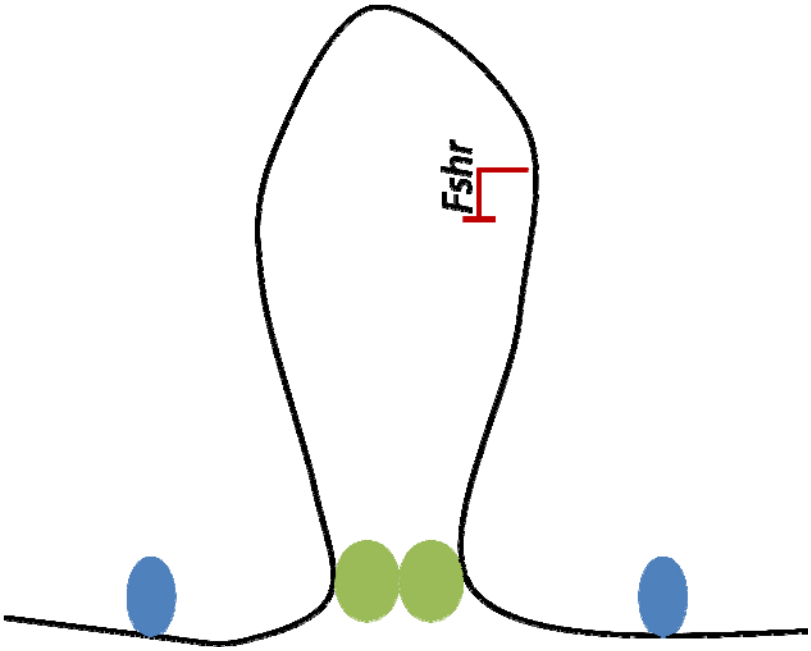


Figure 2: Possible mechanism of CTCF action on Fshr locus: CTCF (green ovals) homodimerize leading to chromatin looping and altering its conformation. This altered conformation inhibits enhancers (blue ovals) from interacting with *Fshr* (red solid line), leading to repressed transcriptional activity.

Part 2: Steroidogenic Factor 1

Chapter 3: Introduction

Identification of Steroidogenic Factor -1

In the early 1990's, analysis of the 5' flanking region of genes encoding steroid hydroxylases, identified a common AGGTCA DNA recognition motif that interacted with the same DNA-binding protein. This protein was designated as steroidogenic factor-1 (SF-1) or adrenal 4-binding protein (Ad4BP) (Rice, Mouw et al. 1991, Morohashi, Honda et al. 1992). The unique expression profile of SF-1 in the adrenal gland, gonad and the pituitary, and its regulation of distinct genes encoding steroid hydroxylases provided the first clue to its important role in directing cellular expression of steroidogenic enzymes (Rice, Mouw et al. 1991, Morohashi, Honda et al. 1992, Hatano, Takayama et al. 1994, Morohashi and Omura 1996, Sadovsky and Crawford 1998). Cloning of SF-1 cDNA from mouse and bovine adrenals revealed that it belongs to the nuclear hormone receptor family and shared homology to the Drosophila nuclear receptor fushi tarazu factor 1 (*FTz-F1*), hence, the gene encoding SF-1 was originally designated but is now referred to as *Nr5A1*, in compliance with Nuclear Receptors Nomenclature Committee of 1999 (Scott and Weiner 1984, Lala, Rice et al. 1992, Morohashi, Honda et al. 1992, Honda, Morohashi et al. 1993, 1999) .

Structure of SF-1

Like other nuclear hormone receptors, SF-1 harbors a DNA binding domain (DBD), a hinge region, a ligand-binding domain, and activation function 2 (AF-2) sequence activation domain (Figure 1). The DNA binding domain is the most conserved part of SF-1 and is comprised of two zinc-chelating modules, which coordinates binding to its DNA response element. The hinge region plays a role in homo-or heterodimerization of SF-1, and is a docking site for miscellaneous cofactors that affect transcriptional activity of SF-1 (Tan, Hall et al. 2002). The C-

terminal ligand-binding domain is composed of a ligand binding pocket, dimerization site, activation function 2 (AF-2) sequence and co-factor binding site. This highly conserved site mediates dimerization with other receptors and contains ligand-induced activation as well as ligand-reversed transcriptional silencing domains (Sadovsky and Crawford 1998). Unlike classic members of the nuclear receptor family, SF-1 lacks a ligand-independent activation function 1 sequence (AF-1) and depends on activation of AF-2 sequence for full transcriptional activity. Initially thought to function in absence of an exogenous ligand, X-ray crystallographic studies on bacterially expressed SF-1 has now identified phospholipids as ligands for SF-1 (Krylova, Sablin et al. 2005, Li, Choi et al. 2005, Wang, Zhang et al. 2005). Mass spectrometry studies on SF1 immunoprecipitated from cAMP-stimulated H295R adrenocortical cells identified endogenous ligands, sphingosine (SPH) and lyso-sphingomyelin (lysoSM) which bound to the receptor. These ligands bind to SF-1 under basal conditions with decreased binding of sphingolipids to the receptor under cAMP treatment (Urs, Dammer et al. 2006). Hence, SF-1 follows the basic structure of a classic nuclear receptor with the exception of the N-terminal AF-1 domain and is a constitutively active lipid binding protein.

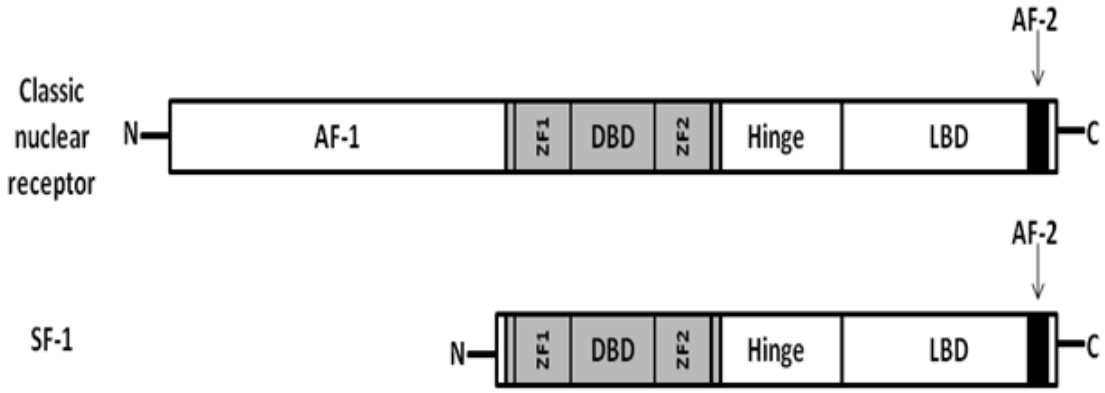


Figure 1: Structure of classic nuclear receptor and SF-1. The classic nuclear receptor has two activation function (AF) domains, a DNA-binding domain (DBD), a hinge region and a ligand-binding domain (LBD). In contrast, SF-1 lacks a functional AF-1 and transcriptional function rests on the conserved AF-2 (black box) located on the C-terminal. SF-1 has a DBD containing two Zn fingers (gray boxes) followed by a Hinge region and a ligand-binding domain.

SF-1 Expression

In developing embryos, *in situ* hybridization studies identified SF-1 expression as early as embryonic day 9.5 (E9.5) in the urogenital ridge, corresponding to the adrenogonadal primordium. During this time, the developing testes and the ovary are histologically indistinguishable and thus SF-1 represents the earliest marker of adrenal and gonadal differentiation. These groups of cells later give rise to two distinct populations of SF-1 expressing cells; a group of cells adjacent to the dorsal aorta representing the adrenocortical precursors and cells adjacent to the coelomic epithelium that give rise to the bipotential gonad (Hatano, Takayama et al. 1994).

SF-1 expression is first detected at E10-10.5 in the adrenal primordium preceding P450scc, which was not detected until E11, indicating that SF-1 is pivotal for the expression of steroid hydroxylases. The chromaffin cell precursors migrate to the adrenal primordium at E12.5-E13.5 and SF-1 expression is limited to the outer cortical cells wherein it is expressed throughout gestation and postnatal life. SF-1 expression in the urogenital ridge at E9 precedes the onset of Sry expression during which the testes and the ovaries are indistinguishable and are termed as bipotential gonads (Parker and Schimmer 1997). Expression of Sry triggers testes specific gene to initiate differentiation of the testis. SF-1 expression persists in the testes, in both the interstitial region where steroidogenic Leydig cells reside and in the testicular cords, which contain fetal Sertoli cells and primordial germ cells. In contrast, SF-1 expression diminishes in the ovaries, both in transcript and protein levels, and does not resume until just before birth, when it is first detected in the theca and granulosa cells, indicating that decrease in SF-1 expression facilitates ovarian development (Hatano, Takayama et al. 1994, Takayama, Sasano et al. 1995).

SF-1 transcripts were also detected in the fetal spleen, anterior pituitary gland, and hypothalamus (Hatano, Takayama et al. 1994, Ikeda, Shen et al. 1994, Hatano, Takakusu et al. 1996, Hanley, Ball et al. 1999, Morohashi, Tsuboi-Asai et al. 1999, Hanley, Rainey et al. 2001). In adults, SF-1 is highly expressed in major steroidogenic cells; the three zones of adrenal cortex, testicular Leydig cells, and ovarian theca cells and granulosa cells and luteal cells (Honda, Morohashi et al. 1993, Ikeda, Lala et al. 1993, Morohashi, Iida et al. 1994, Ramayya, Zhou et al. 1997, Morohashi 1999). It is also expressed in pituitary gonadotrophs (Barnhart and Mellon 1994), ventromedial hypothalamic neurons, a subset of hippocampal neurons that co-express steroidogenic acute regulatory protein and aromatase. and the endothelial linings of the venous sinuses and pulp veins of the spleen (Ramayya, Zhou et al. 1997, Morohashi, Tsuboi-Asai et al. 1999). Low levels of SF-1 transcripts were also detected in the placenta (Morohashi, Hatano et al. 1995). Taken together these findings indicate that SF-1 plays a major role in all components of the hypothalamus-pituitary-adrenal (HPA) and gonadal (HPG) axis.

Function of SF-1

Global knockout of SF-1

The expression pattern of SF-1 suggests that it acts at multiple levels of the HPG axis, and is required for adrenal and gonadal steriodogenesis. To understand the *in-vivo* role of SF-1, three separate laboratories utilized gene disruption models in embryonic stem cells to generate SF-1 knockout mice, all of which showed similar findings (Luo, Ikeda et al. 1995, Luo, Ikeda et al. 1995, Sadovsky, Crawford et al. 1995, Shinoda, Lei et al. 1995). Homozygous SF-1 knockout mice were born at expected Mendelian frequency of 1:4, establishing that SF-1 is not essential for prenatal survival. SF-1 null mice exhibited male to female sex reversal and adrenocortical

insufficiency due to complete absence of the adrenal glands and the gonads. Embryonic analysis of SF-1 null mice revealed an initial development of adrenogonadal primordia followed by apoptosis resulting in complete loss of adrenal glands and gonads. SF-1 is a known regulator of the Mullerian inhibiting substance (MIS), which is crucial for the regression of the Mullerian duct. These SF-1 null mice also lack testosterone resulting in regression of the Wolfian duct. The combination of these causes male-to-female sex reversal and these SF-1 knockout mice are born female irrespective of genetic sex. Interestingly, the gonadotropes of these SF-1 null mice showed impaired expression of a number of genes including LH-β, FSH-β, αGSU, and GnRHr, which are essential for reproduction. One of the surprising features was these knockout mice lacked the Ventral Medial Hypothalamus (VMH), a homeostatic relay center linked to metabolic and female reproductive behavior. Finally, these knockout mice also had defects in their splenic parenchyma. (Ingraham, Lala et al. 1994, Luo, Ikeda et al. 1994, Sadovsky, Crawford et al. 1995, Shinoda, Lei et al. 1995). These global knockout mice died shortly after birth owing to adrenocortical insufficiency since they were rescued by administration of exogenous steroids (Ikeda, Luo et al. 1995).

Tissue-specific knockout of SF-1

Combined with the expression pattern and global knockout of SF-1, mice further underlined the role of SF-1 in adrenal and gonadal development and function. However, these SF-1 knockout mice died soon after birth complicating the effort to identify the role of SF-1 at specific sites. To define the role of SF-1 at specific tissues, Cre/LoxP system was used to produce tissue specific knockouts. Pituitary specific deletion of SF-1 mice was carried out by cre recombinase expression directed to the anterior pituitary gland by the 5' flanking sequences of the α-subunit of glycoprotein hormones. These αGSU-Cre/loxP mice lacked SF-1 in the anterior pituitary but

displayed normal levels in other tissues. These mice displayed diminished levels of pituitary gonadotropins and exhibited severe gonadal hypoplasia. In males, testis showed some differentiation, however, the germ cells were severely decreased in number, and lacked mature sperm. The Leydig cells were low in number and devoid of steriodogenesis. In females, ovaries develop follicles through the antral stage but do produce preovulatory follicles and corpora lutea (Zhao, Bakke et al. 2001). The females had hypoplastic uteri, indicating a severe impairment on sex steroid production in pituitary specific SF-1 knockout mice. Administration of exogenous pregnant mares serum gonadotropins (PMSG) to these pituitary specific knockout mice induced sperm maturation and increased proliferation of Leydig cells. Female mice showed PMSG induced follicular maturation and formation of corpora lutea, establishing that gonads of these mice were functional. These findings indicate that SF-1 plays an important role in regulation of a number of genes involved in the HPG axis and is essential for normal pituitary gonadotrope function (Parker and Schimmer 1997, Schimmer and White 2010).

Anti-Müllerian hormone type 2 receptor-Cre (Amhr2-Cre) driven recombinase transgene was used to generate mice with SF-1 deletion specifically in the testicular Leydig cells and ovarian granulosa cells. These SF-1 adult male mice had hypoplastic testes and the testes failed to descend (Jeyasuria, Ikeda et al. 2004). These mice also had decreased expression of cholesterol side-chain cleavage enzyme (Cyp11a) and absence of steroidogenic acute regulatory (StAR) protein, which form the essential components of testosterone biosynthesis. Female mice with Amhr2 Cre specific deletion of SF-1 were infertile, with ovaries lacking corpora lutea and hemorrhagic cysts. Granulosa cell-specific SF-1 knockout mice had reduced estrogen levels, due to decreased expression of Cyp19a1 or reduced follicle numbers, thus providing conclusive

evidence that SF-1 was required for normal reproductive function (Jeyasuria, Ikeda et al. 2004, Pelusi, Ikeda et al. 2008).

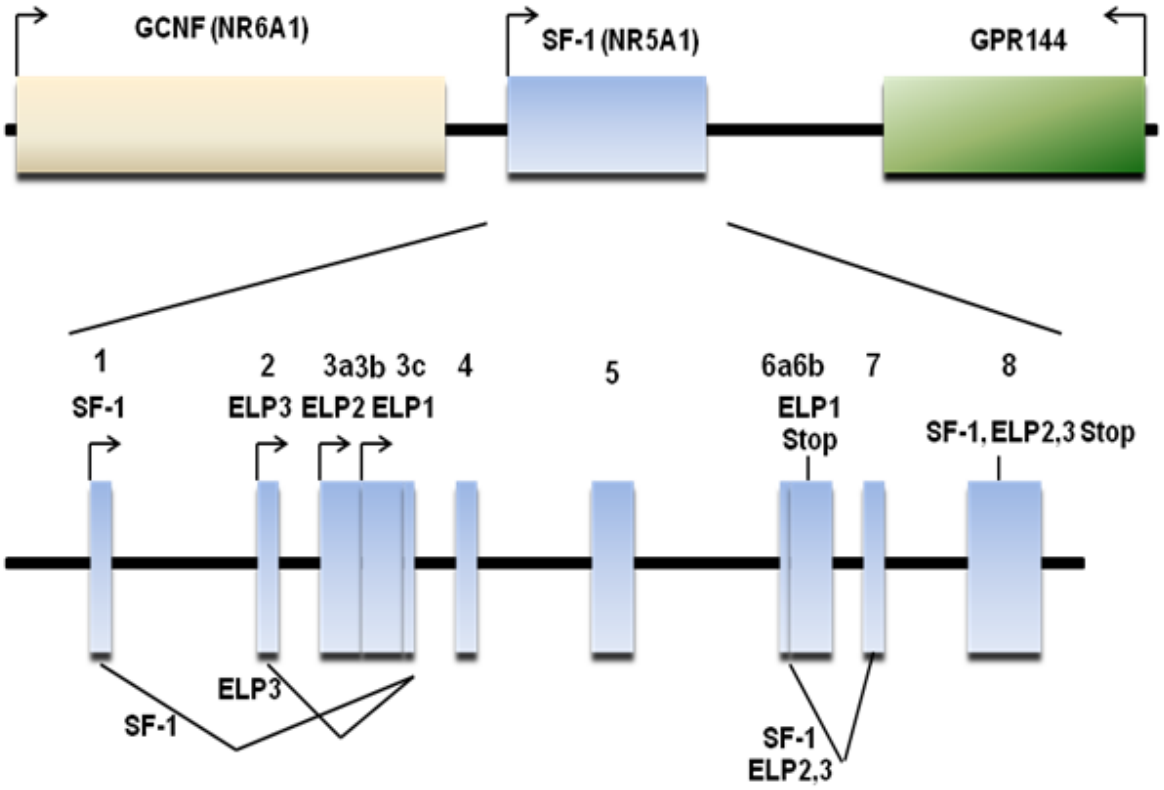
Ablation of SF-1 from the central nervous system (CNS) was directed by using a nestin-Cre transgene. These mice displayed normal development and function of the adrenal glands, pituitary, and gonads but the VMH nuclei were disrupted. In multiple behavioral tests, these CNS-specific knockout mice, both males and females, displayed decreased locomotor activity, exhibited heightened anxiety and later onset of obesity (Majdic, Young et al. 2002, Zhao, Kim et al. 2008). CNS-specific SF-1 knockout female mice showed impaired follicular maturation and decrease in lordosis, ovulation, and fertility, thus highlighting the crucial role of SF-1 in female reproduction (Kim, Li et al. 2010).

SF-1 gene and transcription

The *Nr5a1*/*NR5A1* is located on mouse chromosome 2 and human chromosome 9. The gene is flanked on its 5' side by *NR6A1*, which codes for germ cell nuclear factor, a member of the nuclear receptor family and 3' side by *GPR144* that encodes for probable G Protein-Coupled Receptor 144 (Figure 2). In addition to SF-1, the genomic sequence of *Nr5A1* also encodes for embryonal long-terminal repeat-binding protein, ELP 1 and its isoforms ELP2 and ELP3 through alternate promoter usage and splicing (Figure 1). Among these four transcripts, SF-1 has emerged as a key regulator of endocrine homeostasis. Function analysis of ELP transcripts have revealed ELP1 acts as a repressor while ELP2 and ELP3 function as transcriptional activators (Ninomiya, Okada et al. 1995).

While there is no evidence of a TATA box in the first 110bp in the 5'-flanking region of NR5A1, a number of other regulatory elements such as SRY (sex determining region Y)-box (SOX)

binding site, an E box, a CCAAT box and Sp1/Sp3 site have been identified (Schimmer and White 2010). In addition, two other Sp1/Sp3 sites are located between the +10 and +30 region upstream of the transcriptional start site (Figure 3) (Scherrer, Rice et al. 2002). Numerous other transcription factors that regulate SF-1 transcription such as GATA-4, WT1, Lhx9, SOX15, SOX30, TEAD-4 and CBX2, have been identified (Schimmer and White 2010) . To examine transcriptional requirements for *Nr5a1* in vivo, short promoter fragments were introduced into transgenic mice. Earlier work identified a 90bp proximal promoter fragment reporter construct that was sufficient to direct SF-1 expression in Y1 adrenocortical cells (Woodson, Crawford et al. 1997).


GCNF (NR6A1)
SF-1 (NR5A1)
GPR144
1
2
3a3b
3c
4
5
6a6b
7
8
SF-1
ELP3
ELP2
ELP1
ELP1
Stop
SF-1, ELP2,3 Stop
ELP3
SF-1
SF-1
ELP2,3

B

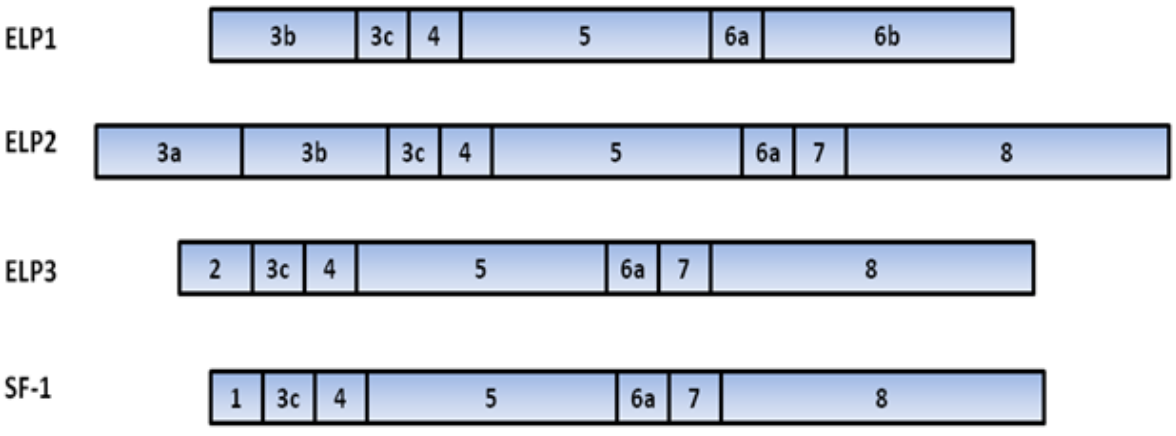

ELP1
3b
3c
4
5
6a
6b
ELP2
3a
3b
3c
4
5
6a
7
8
ELP3
2
3c
4
5
6a
7
8
SF-1
1
3c
4
5
6a
7
8

Figure 2: Schematic organization of the mouse *Nr5a1* gene. Located on chromosome 2, *Nr5a1* is flanked by *Nr6a1*, encoding germ cell nuclear factor and *Gpr144* probable G-protein coupled receptor 144. SF-1 is encoded by seven exons as depicted in B, and codes for 462 amino acids and bears 94% homology to the human SF-1. In addition to SF-1, Nr5a1 also gives rise to three other mRNA transcripts named named embryonal long-terminal repeat binding proteins (ELP1-3). These ELP proteins are encoded by exons as depicted in the figure. ELP3 bears a very close resemblance to SF-1 except for the starting exon. In mouse, ELP1 is expressed in the gonads and in Y1 mouse adrenocortical cell line. ELP2 is not expressed in any mouse tissue, but similar to ELP1 is expressed in Y1 cell line. ELP3 bears an expression profile very similar to SF-1 and is expressed in the pituitary, adrenal, ovary, testis, and spleen. ELP3 expression has also been found in the cerebrum, heart, kidney, spleen and Y-1. The potential biological role of these isoforms are unknown although in vitro studies have shown that ELP1 acts as a repressor while ELP2 and ELP3 function as transcriptional activators. Adapted from (Parker, 1997)

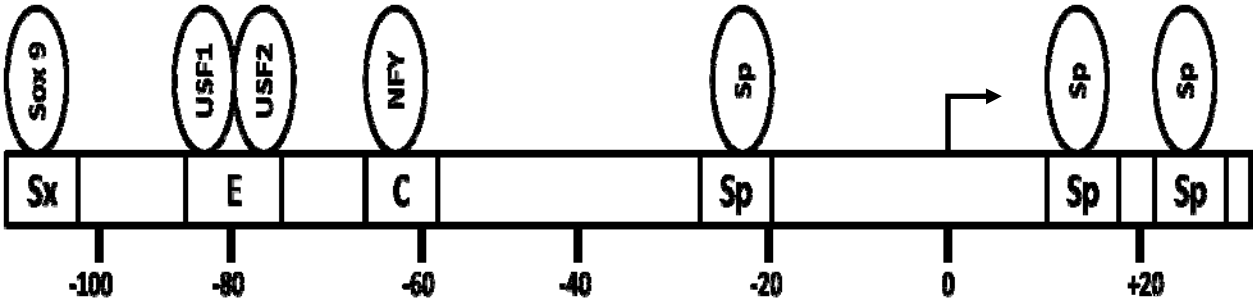


Figure 3: Schematic representation of SF-1 and its proximal promoter. The transcription start site (TSS) is depicted at base pair 0 and indicated by the bend arrow. The box regions represent conserved elements and their location from the TSS is depicted by the base pairs. Transcription factors that bind to these conserved regions are in ovals. Different SF-1 expressing tissues utilize multiple combinations of transcription factors and include many more proteins than depicted here. Sx is the binding site for Sox which is required for SF-1 expression in the Sertoli cells of the testes; The E-box is bound by the upstream stimulatory factors 1 and 2; C represents a CCAAT box that binds to nuclear factor Y (NFY) and Sp represents binding site for transcription factors Sp1 and Sp3. Adapted from (Schimmer and White 2010).

To analyze regions required for proper SF-1 expression, bacterial artificial chromosome (BAC) transgenic mouse were generated with the enhanced green fluorescent protein reporter gene (eGFP) under the control of 50kb of the *Nr5a1* locus encoding for SF-1. Analysis of transgenic SF-1/eGFP mouse embryos indicated that this transgene was able to replicate the developmental profile of SF-1 in the gonads, adrenal cortex and VMH. In adult transgenic mouse, this fragment was able to replicate SF-1 expression in the adrenal cortex, testes, ovaries, and the VMH. However, expression was not seen in the corpora lutea or the anterior pituitary gland indicating that regulatory regions beyond those expressed in the construct were required (Stallings, Hanley et al. 2002). A somewhat larger fragment that encompassed the -590 to +85 region cloned in front of the *Lac Z* gene was able to recapitulate expression of SF-1 only in the gonads (Wilhelm and Englert 2002). Transgenic mouse that expressed mouse genomic DNA extending from exon2 of *Nr5a1* into the upstream *Nr6a1* was created. This transgene showed SF-1 expression in the adrenal cortex, testicular Leydig cells, ovarian theca cells, the ventromedial hypothalamus and spleen, but not in the pituitary gland or corpus luteum (Stallings, Hanley et al. 2002). To recapitulate SF-1 expression, transgenic mouse expressing a much larger 500-kb yeast artificial chromosome (YAC) that contains the entire rat *Nr5a1* locus plus 5′ and 3′ sequence extending into the *Nr6a1* and the *Psmb7* genes was developed. These mice were able to duplicate endogenous expression of SF-1 both spatially and quantitatively indicating that the transgene contained all necessary sequences for proper SF-1 expression (Karpova, Presley et al. 2005).

In addition to the identified regulatory sequences and transcription factors, *Nr5a1* expression is regulated by epigenetic modifications. In particularly DNA methylation appears to act as another layer of regulation. In endometriotic tissues, the CPG islands in the *Nr5a1* promoter were found to be hypomethylated when compared to normal endometrial tissues leading to aberrant

expression of a number of steroidogenic genes including *Nr5a1* (Xue, Lin et al. 2007). When normal endometrial cells were treated with 5-aza-deoxycytidine, a demethylating reagent, the *Nr5a1* promoter was demethylated leading to transcription of *Nr5a1*, whereas methylation led to loss of activity (Hoivik, Aumo et al. 2008). The correlation between methylation of the *Nr5a1* promoter and activity is also seen in steroid-secreting cell lines and normal steroid-secreting tissues. The *Nr5a1* proximal promoter is unmethylated in developing testis and ovary and hypermethylated in tissues that do not express for SF-1 (Mohn and Schubeler 2009).

Chapter 4: Analysis of *Fast*, *Ftz-F1-associated transcript*, suggests functional cooperation with the nuclear receptor Steroidogenic Factor 1

Abstract

An evolutionarily conserved expressed sequence tag lacking an identifiable open reading frame was identified and named *Ftz-F1 associated transcript* (*Fast)* for its proximity and functional associations with *Ftz-f1 (Nr5a1),* the gene encoding steroidogenic factor-1 (SF-1). SF-1 is a nuclear receptor and a key determinant and regulator of the adrenal and reproductive axes. Mammals with SF-1 mutations display a range of phenotypes, including absence of adrenal glands and gonads, diminished pituitary gonadotropins, disruption of the ventromedial hypothalamus, ovarian failure, and XY-sex reversal. Directional RT-PCR demonstrated *Fast* is transcribed in opposite direction to that of *Ftz-F1* and 5' and 3' RACE showed *Fast* is composed of three exons, polyadenylated, and derived from multiple transcriptional start sites that border and extend into exon 1g of *Ftz-F1*. In addition, sequence analysis revealed two possible spliced variants that lacked part of exon 3. RT-PCR, using RNA from multiple mouse tissues and cell lines, revealed *Fast* and *Ftz-f1* share the same expression profile. Furthermore, *Fast* transcript was located predominantly in cytoplasm, when analyzed in MA-10 Leydig cell lines. In addition, in P-19 embryonal-carcinoma cells, SF-1 and *Fast* transcript levels were similarly attenuated in response to retinoic acid, while in granulosa cells, both were induced by *in vivo* treatment with pregnant mare's serum gonadotropin. To evaluate function, *Fast* knockdown and overexpression paradigms were used. Unexpectedly, neither strategy caused a change in SF-1 protein or mRNA levels.RNA structural analysis and preliminary northern blot suggests *Fast* gives rise to a small RNA and could possibly regulate other genes. Thus, *Fast* is a long non-coding RNA that is regulated like *Ftz-F1* with respect to its tissue profile and hormone response, suggesting it acts within the same biological pathway and potentially, in conjunction with SF-1, regulate development and function of adrenal glands and gonads.

Introduction

Steroidogenic factor 1 (SF-1, Ad4BP, NR5A1) is a nuclear hormone receptor and encoded by one of four known transcripts derived from *Ftz-F1* gene, officially known as *Nr5a1*. Of the four *Ftz-F1* gene products, SF-1 has received the most interest due to its critical roles in endocrine homeostasis and organ development (Luo, Ikeda et al. 1995, Ikeda 1996, Parker and Schimmer 1997). SF-1, first identified as a protein that bound a common regulatory motif (PyCAAGGTCA) within several genes associated with steroid biosynthesis, has emerged as a major regulator of the endocrine system (Lala, Rice et al. 1992, Morohashi, Honda et al. 1992, Schimmer and White 2010, Gardiner, Shima et al. 2012). Expression of SF-1 is limited to a discrete set of cells that are functionally linked by their role in the endocrine system. More specifically, SF-1 is produced in supporting cells of the testis (Sertoli) and ovary (granulosa cells), steroidogenic cells of the adrenal gland (adrenocortical), testis (Leydig), and ovary (theca), gonadotrope cells of the pituitary, and cells within the ventromedial hypothalamus (Honda, Morohashi et al. 1993, Ikeda, Lala et al. 1993, Barnhart and Mellon 1994, Ingraham, Lala et al. 1994, Morohashi, Iida et al. 1994, Morohashi 1999, Ngan, Cheng et al. 1999). During development, SF-1 is first observed on embryonic day 9 (e9) in a single population of cells in the urogenital ridge that subsequently resolve into two discrete populations that give rise to adrenocortical cells and gonadal cells (Ikeda, Shen et al. 1994, Morohashi, Hatano et al. 1995, Hanley, Ball et al. 1999, Hanley, Rainey et al. 2001). By e12.5, SF-1 transcripts are limited to the gonad, the adrenal primordium, and the diencephalon, as development continues, SF-1 expression in the gonads becomes sexually-dimorphic with higher levels observed in the developing testis (Ikeda, Shen et al. 1994, Morohashi, Hatano et al. 1995, Hanley, Ball et al. 1999). While SF-1's expression profile implicated it in transcriptional roles unrelated to

steroidogenesis, it was not until mice lacking *Ftz-F1* were generated that it was recognized as a mediator of organ development. Four different SF-1 knockout models were created and the results from each demonstrated the requirement for SF-1 in development of the adrenal glands and gonads (Ingraham, Lala et al. 1994, Luo, Ikeda et al. 1994, Sadovsky, Crawford et al. 1995). Accordingly, SF-1 deficiency resulted in male-to-female sex reversal and early postnatal death, due to the loss of testicular and adrenal steroids, respectively. Further evaluation of null embryos indicated that organ development initiated but regressed soon after mesenchymal thickening of the genital ridge (Luo, Ikeda et al. 1994).

Considering the critical role of SF-1 in development and function of the endocrine system, there has been much enthusiasm to elucidate the regulatory mechanisms that specify SF-1 transcription in the appropriate temporal and spatial manner. While DNA sequences comprising the SF-1 promoter region have been thoroughly evaluated, transgenic analysis of the promoter region in mice, revealed it was insufficient to establish the tissue-, cell- and developmental specificity of SF-1 expression(Nomura, Bartsch et al. 1995, Woodson, Crawford et al. 1997, Daggett, Rice et al. 2000, Wilhelm and Englert 2002). Other studies, using increasingly larger amounts of DNA as transgenes, expanded the location of regulatory sequences required for transcriptional competency to nearly 150kb spanning the Ftz-f1 locus (Stallings, Hanley et al. 2002, Karpova, Presley et al. 2005). Results from these studies revealed a mechanism of SF-1 regulation that emphasizes the role of distal regulatory sequences in cell-specific expression.

In an effort to resolve the location of *cis*-acting regulatory elements that control SF-1 expression, comparative genomics was used to identify highly conserved sequences between the mouse and human loci that serve as candidate regulatory elements. While several conserved, non-coding sequences were identified in this analysis and are reported elsewhere, a conserved EST located

between *Ftz-f1* and *Gcnf* was uncovered (Karpova, Presley et al. 2005). The EST was 1377 bp long, encoded by three exons and was transcribed in the opposite orientation from *Ftz-f1*. Expression profile of *Fast* was nearly identical to SF-1 and lacked open reading frames, indicating it was a long noncoding RNA (lncRNA). This study further characterizes the non-coding RNA and provides preliminary evidence that it is a precursor to a small RNA of unknown function.

Results

Comparative genomics revealed two conserved ESTs upstream of *Ftz-f1*. To identify conserved, non-coding sequences in proximity to *Ftz-F1*, genomic sequences of human and mouse *Ftz-F1* were compared using pair-wise blast analysis. The identified conserved sequences were next compared to sequences in the expressed sequence tag databases, revealing that two of the conserved sequences corresponded to novel transcripts cataloged in a RIKEN testis cDNA library (Fig. 1, Accession #AK017050 and AK007201). BLAST analysis of AK007201 confirmed its genomic location approximately 10kb 5' to *Ftz-F1* exon 1a and revealed its identity with another EST (Accession #AF390897S4), which indicated it was expressed from *Nr6a1*, the 5' proximal gene to *Ftz-F1* that encodes Germ-cell nuclear factor, Gcnf (Fig. 1). Alignment of AK017050 to mouse *Ftz-F1* revealed two exons. The larger exon (790bp) contained the identified conserved sequence, which localized 5.3kb 5' to *Ftz-F1*. The smaller exon (125bp) localized just 5' to *Ftz-F1* exon 1g (Fig. 1). The genomic location and transcript processing revealed EST AK017050 as a novel transcript derived from the *Ftz-F1* locus and it was named *Ftz-F1-associated-transcript* or *Fast*.

***Fast* transcriptional orientation opposes that of *Ftz-f1*.** The novelty of *Fast* and its association with SF-1 prompted further investigation to determine its transcript characteristics and expression profile. Transcriptional orientation of *Fast* was determined by directional RT-PCR, using cDNA generated with oligo-dT or primers of different orientations corresponding to sequences within the two identified *Fast* exons (Fig.2A). Following PCR of cDNA generated from mouse testis RNA, *Fast* products were visualized by Southern blot analysis, using an internally located probe (Fig. 2B). Two distinct amplified products were observed, when template cDNA was synthesized using oligo-dT or primer 2, but not primer 1, indicating *Fast* is transcribed in opposite orientation to *Ftz-F1*, proceeding from the shorter exon, closest to *Ftz-F1* exon 1g, towards the larger exon, located midway between *Gcnf* and *Ftz-F1* (Fig. 2 A&B). *Fast* expression and transcriptional orientation were also examined from human testis RNA, using primers directed to human sequence syntenic to the identified mouse exons. No amplified product for human *FAST* was observed using primers located in the two putative exons (primers 1 & 2, data not shown). However, amplification of human *FAST* was observed using primers (primers 1 & 3) located within the single conserved exon and cDNA synthesized with either oligo-dT or primer 1 (Fig. 2 C&D). Taken together, these data indicate that *Fast* is an evolutionarily conserved, poly-adenylated transcript that arises from syntenic regions of mouse chromosome 2 and human chromosome 9 in the opposite orientation to that of *Ftz-F1* (Brian Hermann- Regulation of FSH-receptor and SF-1: transcriptional control in reproduction). Furthermore, the studies revealed that mature mouse *Fast* (m*Fast*) is produced through splicing of at least two exons, while the complete exon structure of human *FAST* is currently unresolved.

***Fast* is a noncoding RNA consisting of multiple transcript variants.** Further sequence analysis performed on PCR-amplified *Fast* products confirmed m*Fast* as a poly-adenylated,

spliced transcript, and revealed a third exon and several potential splice variants within its transcript pool (Fig. 3). 5'RACE was used to map the transcriptional start site for m*Fast* and revealed several distinct 5' ends, three of which overlapped *Ftz-f1* exon 1g (Fig. 3). Similarly, 3'RACE was used to identify the 3' end of m*Fast* transcripts and revealed the presence a non-canonical poly-A signal sequence (CATAAA) that likely directs transcript poly-adenylation. In addition to identification of a third m*Fast* exon, sequence analysis revealed two amplified products lacking portions of exon 3. Variant 1 represents loss of 182bp from exon 3, while variant 2 lacks 71bp. While it is currently unknown if these represent true splice variants or amplification artifacts caused by strong secondary structure, it is interesting to note that the deleted sequences reside within the regions of greatest conservation. Regardless, the studies demonstrate the existence of multiple m*Fast* transcript variants that result from the use of multiple transcriptional starts sites and at least one splice variant based on inclusion of the small exon 2 (Fig. 3). Similar experiments showed the existence of human *FAST* and revealed that it contains only the large exon homologous to mouse exon 3 and has a non-canonical poly-A signal similar to mouse *Fast* (Brian Hermann- Regulation of FSH-receptor and SF-1: transcriptional control in reproduction).

***Fast* is predominantly cytoplasmic.** To delineate the function of the lncRNA, its sub-cellular location was determined (Fig. 4). RNA from mouse MA-10 cells was isolated from cytoplasmic and nuclear fractions, which were assayed by RT-PCR for *Fast* and *Gapdh* as control. RT-PCR was done subsequently using the gene-specific primers for the noncoding RNA (Table 1). PCR products of *Fast* were obtained predominantly in the cytoplasmic fraction, with a no-RT control to eliminate any false amplification that might arise due to DNA contamination.

***Fast* can produce a 7kd protein.** To determine if *Fast* is capable of producing smaller peptides, the RNA was used in an *in vitro* translation system and labeled proteins visualized by ^{35}S methionine (Fig.5). A ~7 kd protein was detected in pcDNA3 *Fast* construct, indicating that a small protein maybe translated under *in vitro* conditions. The control plasmid that harbors the luciferase and DMRT1 was transcribed in full length *in vitro*. Inspection of cDNA sequences for both human and mouse *Fast* failed to identify any open reading frames (ORFs) coding for proteins larger than 260 or 194 nucleotides, respectively. Furthermore, none of these small ORFs were conserved between mouse and human *Fast*, nor did they share homology with any known protein. This characteristic is consistent with transcripts designated as noncoding RNAs (ncRNAs), in which the average length of the longest ORF is approximately 200 nucleotides (Numata, Kanai et al. 2003). Thus, the structural studies of *Fast* reveals it to be a conserved, polyadenylated, spliced transcript that does not code for a functional protein, suggesting it is a member of a growing class of ncRNAs that contribute to or are implicated in numerous biological processes.

***Fast* is co-expressed with SF-1.** *Fast* expression was examined by RT-PCR of RNA isolated from adult mouse tissues and cell lines. Interestingly, the tissue expression pattern for *Fast* was closely identical to that of SF-1, with both transcripts detected in adrenal, testis, pituitary, and ovary (Fig. 6). Neither *Fast* nor SF-1 was consistently detected in mouse spleen, in contrast to previous reports of SF-1 in this tissue (Morohashi 1999, Kimura, Yoshii et al. 2000, Karpova, Presley et al. 2005). The conflict may reflect differences in the methodology used to detect *Ftz-F1* transcripts. To determine specific *Fast*-expressing cell types, mRNA *in situ* hybridization was performed. However, high background hybridization precluded identification of cellular expression patterns, using this method. To help circumvent this obstacle an alternative approach

was employed that examined *Fast* in SF-1-positive cell lines using RT-PCR. The results showed *Fast* and SF-1 expression in many of the same tissues and cell line, including MA10 Leydig cells (Fig. 6) (Brian Hermann- Regulation of FSH-receptor and SF-1: transcriptional control in reproduction). In summary, the expression data show identical expression patterns for *Fast* and SF-1, at both the tissue and cellular level. The similar expression patterns of SF-1 and *Fast* suggested they share similar regulatory mechanisms.

To further explore this avenue, the response of SF-1 and *Fast* to retinoic acid (RA) was evaluated in P-19 embryonal carcinoma cells. Previous studies have shown time-dependent declines in SF-1 expression in P-19 cells exposed to RA (Barnea and Bergman 2000, Gu, Goodwin et al. 2005). RNA samples from P-19 treated cells obtained from Dr. Austin Cooney were evaluated for SF-1 and *Fast* expression by semi-quantitative PCR. Results from the analysis revealed that both SF-1 and *Fast* transcripts decreased in response to RA, with notable decreases in transcript levels after 36 hours of treatment and nearly undetectable levels by 72 hours (Fig. 7). No change was observed in the L7 mRNA control (Brian Hermann- Regulation of FSH-receptor and SF-1: transcriptional control in reproduction).

In addition, hormonal response of *Fast* in mouse granulosa cells was examined. Expression analysis of *Fast* in RNA samples isolated from hormone-treated granulosa cells (graciously provided by Dr. Lane K. Christenson) was evaluated. This revealed an increase in *Fast* transcripts 48 hours after PMSG stimulation (Fig. 8). In summary the expression studies show *Fast* expression is hormonally regulated, which together with its restricted tissue profile, strongly support its role as a functional ncRNA. Its striking similarities to SF-1 suggest *Fast* and SF-1 share an important functional relationship and act within the same biological pathways.

***Fast* does not regulate SF-1 mRNA stability.** Because shared expression and regulatory profiles of SF-1 and *Fast*, suggested a functional relationship, the effect of *Fast* knockdown on SF-1 mRNA was examined. short hairpin RNAs (shRNAs) were introduced in MA10 Leydig cells to decrease *Fast* levels (Figure 9A).Semi-quantitative RT-PCR showed that despite reduction of *Fast* mRNA by >80%, there was no notable decrease in SF-1 mRNA. (Fig.9 B). Further RT-PCR analysis of genes under SF-1 regulation showed no change in Cyp17 and StAR, mRNA levels indicating that they are not regulated by *Fast* (Figure 10). Furthermore, over expression of the *Fast* transcript in MA 10 cells did not alter mRNA levels of SF-1 (Figure 11).

***Fast* does not regulate SF-1 protein levels**. SF-1 regulates steriodogenesis which is regulated by gonadotropins via cAMP. Thus, *Fast* may act in conjunction with the cAMP pathway to regulate SF-1. MA-10 cells were transfected with mu6 shRNA control and *Fast* shRNA 6, and 48 hrs post transfection cells were treated with cAMP and harvested at various time points of 0hr, 1hr, 2hr, and 6 hrs (Figure 10). No change in SF-1 protein was detected by Western blot with either mU6 or *Fast* shRNA or *Fast* over expression (Fig. 12, 13)

Discussion

In the current study, we characterized a conserved transcript within the *Ftz-F1* locus that lacks an identifiable ORF expressed and is regulated in a manner that resembles SF-1. A ~7kd protein was detected in the pcDNA3 *Fast* construct in our in vitro transcription translation assay. On closer analysis of the *Fast* transcript we were able to identify an open reading frame that may contribute for the expressed protein. However, owing to the fact that it is an *in vitro* transcription translation assay, it is possible that what we are seeing is a spurious reaction (Dinger, Pang et al. 2008). A number of bifunctional RNAs have been identified that have been found to function at

both the transcript level and at the protein level. Examples of such bifunctional RNA include *Steroid Receptor Activator*, *VegT* and *Oskar* mRNA [reviewed in(Dinger, Pang et al. 2008)]. While the data are intriguing on several levels, *Fast*'s close association with SF-1 and its status as a noncoding RNA (ncRNA) are features, which reveal its greatest potential. In particular, its expression profile and response to retinoic acid and FSH suggest that, like SF-1, it too acts to control normal development and endocrine homeostasis. Furthermore, the molecular characteristics of *Fast* place it in the family of mRNA-like ncRNAs, which not only contribute significantly to the transcribed genome but whose members are increasingly identified with functions in a variety of biological processes.

With completion of the human genome came the surprising finding that only 25,000 or so protein-coding genes were required to direct such a complex biological program (Venter, Adams et al. 2001, 2004). It was then revealed that the number of expressed transcripts was much greater than the number of protein coding genes predicted from genome annotation (Okazaki, Furuno et al. 2002, Carninci, Kasukawa et al. 2005, Mattick 2005). Resolution of this discrepancy came with the discovery that a significant portion of the genome gives rise to transcripts that do not code for proteins. Thus, tiling experiments across ten human chromosomes revealed approximately 10% of the genome is expressed as polyadenylated transcripts, a number ten times greater than that for protein-coding genes and extensive analysis of mouse transcriptional units revealed a similar difference in the number of transcripts to estimated protein-coding genes (Carninci, Kasukawa et al. 2005, Cheng, Kapranov et al. 2005). While there is evidence that many of these putative transcripts represent library contamination from genomic DNA or unprocessed pre-mRNA, or alternatively nonspecific transcriptional noise, a wealth of studies substantiate important biological roles for polyadenylated ncRNAs

(Ravasi, Suzuki et al. 2006). The few well-characterized ncRNAs identify functions in X-chromosome inactivation, chromatin structure, DNA imprinting, DNA methylation, transcription, environmental response, development, and cell differentiation, and has led to recognition of the potential magnitude of their biological importance [reviewed in (Kung, Colognori et al. 2013)]. Well known examples of such ncRNAs include *Xist, TsiX, H19, Kcnq1ot1*, *Air,* and *MALAT1*. *Xist*, found in both mice and humans, localizes to the inactive X chromosome in females to control dosage compensation (Lyon 1961, Lee 2011). It is also regulated by the ncRNA *TsiX* (Lee, Davidow et al. 1999). *H19*, *Kcnq1ot1*, and *Air* are mammalian RNAs associated with genomic imprinting (Brannan, Dees et al. 1990, Bartolomei, Zemel et al. 1991, Lee, DeBaun et al. 1999, Lyle, Watanabe et al. 2000). *MALAT1* a nuclear lncRNA, is associated with various cancers and regulates a number of cytoskeletal and extracellular matrix genes (Ji, Diederichs et al. 2003, Tano, Mizuno et al. 2010, Lin, Roychowdhury-Saha et al. 2011).

While the concept of noncoding RNAs (ncRNAs) is not novel, their extensive contribution to the expressed portion of the genome came as a surprise and the resulting biological implications has brought new appreciation to these RNAs and intensified research in order to understand their functional potential. In addition to the traditional RNA polymerase III-derived transcripts, tRNA and rRNA, there are numerous classes of noncoding RNAs including; microRNA (miRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), small non-mRNA (snmRNA, a.k.a small ncRNA), small nucleolar RNA (snoRNA), small temporal RNA (stRNA), and large, mRNA-like ncRNA [reviewed in (Guil and Esteller 2012)] . While the mechanisms and characteristics of these RNAs are still being revealed, these non-coding RNA can be classified according to their nucleotide size and function. For small noncoding RNAs, the most

extensively studied are the 21–25 nucleotide non-mRNAs, which include miRNAs, siRNAs, and the 100-200 nucleotide snoRNAs. In animals, many miRNAs and all snoRNAs are located within introns [reviewed in (Mattick and Makunin 2005)]. Both siRNAs and miRNAs act to suppress gene expression by interacting in a sequence-specific manner with target mRNAs to regulate their stability and translation. They are also both excised from larger ncRNAs by RNase III digestion, requiring the parent RNA to have double-stranded secondary structure [reviewed in (Ghildiyal and Zamore 2009)]. Small nucleolar RNAs function within the nucleolus to guide ribosome biogenesis (Eddy 2001). These RNAs fall into two general classes, C/D box and H/ACA snoRNAs, which direct ribose methylation and pseudo-uridylation of rRNA, respectively. Among small ncRNAs, miRNAs have received the greatest attention due to their participation in diverse regulatory pathways, including control of developmental timing, haematopoietic cell differentiation, apoptosis, cell proliferation and organ development (Pasquinelli 2012).

While the above studies demonstrate important functional roles for these specific ncRNAs, they do resolve the question of function with respect to the numerous uncharacterized ncRNAs identified in expression libraries and if they, in general, are biologically important or an artifact of genome analysis. However, studies have begun to take more global approaches to determine if these ncRNAs represent functional RNAs. In a study by Cawley et al. binding sites for the transcription factors Sp1, cMyc, and p53 were mapped along human chromosomes 21 and 22 to help identify promoters regulated by these proteins (Cawley, Bekiranov et al. 2004). The results revealed the surprising finding that only 22% of the binding sites associated with protein-coding genes, while 36% of the sites were correlated with ncRNAs. Further examination showed that the ncRNAs on these chromosomes acted similar to the coding genes in response to retinoic

acid-induced differentiation. Likewise, studies in mice showed that many cloned cDNAs lacking open reading frames (ORFs) were derived from genuine transcripts of unknown function and that these ncRNAs, in general, have larger exons and fewer introns than protein-coding transcripts. The study also revealed that a significant proportion of ncRNAs are expressed in a tissue-specific manner and were regulated by extracellular signals (Ravasi, Suzuki et al. 2006). More recently, a study by Willingham et al. used shRNAs to knockdown expression of selected ncRNAs from the RIKEN Fantom2 mouse cDNA collection that showed significant sequence conservation to human genomic sequence. Eight new functional ncRNAs were identified, six essential for cell viability, one a repressor of Hedgehog signaling, and one a repressor of the transcription factor NFast called NRON (Willingham, Orth et al. 2005). Together, the findings indicated that ncRNAs are an important, regulated component of the mammalian genome and represent a largely uncovered level of gene regulation in complex organisms.

Importantly, transcription of one *Fast* exon and its orientation was conserved between species, suggesting that *Fast* was not a merely non-functional artifact of abundant, nearby SF-1 transcription. The most intriguing feature of *Fast* was its shared expression profile with SF-1, which strongly implicates a functional relationship between their respective protein and RNA products. Since *Fast* and SF-1 are expressed and regulated similarly, a concerted regulatory scheme for these two genes would be most efficient, and thus, the two genes may share some of the same transcriptional regulatory mechanisms. While its expression is clear in the human testis, it is not known if the full extent of *Fast* expression completely mimics human SF-1 expression.

Although a functional relationship between *Fast* and SF-1 were suggested by their shared expression patterns, a cellular or molecular role for *Fast* was not immediately apparent. At least

one possibility, however, was based on reports of other non-coding RNAs, which suggest that *Fast* mRNA could be processed to form a miRNA or siRNA. These two types of small ncRNAs (generally 20-25nt) constitute a growing class of small regulatory RNAs that control expression of protein-coding genes via an antisense-RNA mechanism by blocking mRNA translation or directly targeting mRNA degradation by hybridization with the homologous sequence in other genes (reviewed in (Huttenhofer, Schattner et al. 2005)). Although experiments in this study indicate that *Fast* does not regulate SF-1 expression via its homology to Ftz-f1 exon 1g, it is possible that another portion of *Fast* produces a small RNA that is homologous to one or more other genes.

In support of this theory, RNA secondary structure modeling indicated several highly stable stem-loop and hairpin structures which in the *Fast* cDNA, which would form the precursors necessary for generating a small RNA by RNase III digestion. Additional evidence was obtained experimentally by detection of a small RNA containing *Fast* sequence in *Fast* expressing cells within the highly conserved segment of the gene (data not shown). Thus, these results raise the possibility that secondary structure in *Fast* is processed to form a small RNA which serves some function in SF-1 positive cells. Although this possibility is intriguing, a significant number of non-coding RNAs are emerging that have no known associated function, raising the possibility that a large number of non-functional ncRNAs are produced in eukaryotic cells. Although we cannot exclude *Fast* as a member of this latter non-functional class, its high level of evolutionary conservation, intriguing expression pattern, and opposite transcriptional orientation to Ftz-f1 suggest that *Fast* is not merely coincidental. Ongoing experiments are seeking to clarify the existence of a small RNA produced from *Fast* and to determine its function in expressing cells.

The extensive biological importance of small ncRNAs to a wide array of critical developmental and physiological processes has emerged over the last several years. If subsequent studies demonstrate that *Fast* produces a small ncRNA, it may therefore play a fundamental regulatory role in expressing tissues. Some examples of miRNAs which are expressed at specific developmental time-points or during particular processes include those shown to participate in the mechanisms controlling stem cell and hematopoietic lineage differentiation, insulin secretion, and some which are imprinted, just to name a few (Seitz, Youngson et al. 2003, Seitz, Royo et al. 2004, Suh, Lee et al. 2004). Based on these exciting findings, there is tremendous potential that *Fast* participates in a novel physiological or developmental process that is key to endocrine development and function.

Materials and Methods

Assembly, annotation and comparative sequence analysis of the human and mouse Ftz-f1 loci. A 4187947bp subsequence of human chromosome 9 genomic contig (Accession #NT_008470) containing the entire human FTZ-F1 gene and roughly 2Mb each of 5' and 3' flanking sequence was identified in the NCBI Entrez Nucleotide search engine (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide). The mouse Ftz-f1 gene was identified within a single chromosome 2 genomic contig (Accession #NT_039206) in the same way. Repetitive sequences were identified and masked using the Repeatmasker web server (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker). Exons were assigned using the Spidey mRNA-to-genomic alignment program within NCBI (http://www.ncbi.nlm.nih.gov/Sitemap/index.html#Spidey). Two expressed sequence tags (ESTs) (Accession #AK017050 & #AK007201) were identified by sequence comparison of the rat and human FTZ-F1 loci using pairwise Blast alignment through the NCBI web site (http://www.ncbi.nlm.nih.gov/; Tatusova and Madden, 1999) under the program's default parameters (Cost to open a gap [5], Cost to extend a gap [2], Penalty for a mismatch [-2], Reward for a match [1], Expectation value (E) [10], Word size, 11 for blastn) and a minimum conservation level of 75%.

The analysis used a FTZ-F1-containing subsequence of a human chromosome 9 contig (Accession #NT_008470) and a mouse chromosome 2 contig (Accession #NT_039206) containing Ftz-F1. Repetitive sequences were identified and masked using the web-based program Repeatmasker, exons assigned, and the sequences aligned using the BLAST sequence alignment program through the National Center for Biotechnology Information (NCBI). All

identified conserved sequences were examined against the nonredundant and expressed sequence tag databases at NCBI.

Cell culture conditions. Mouse P19 embryonal carcinoma (EC) cells were maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum, 2 mM glutamine, 100 U/ml penicillin, and 100 mg/ml streptomycin. For retinoic acid (RA) treatment of P19 cells, 1 uM RA was added to growth media for 12, 24, 36, 48 or 72 hours prior cell harvest. MA-10 cells, a mouse Leydig tumor cell line were cultured as described (Ascoli 1981) . Mouse granulosa cells were isolated from ovaries of 19-day-old CF-1 mice at various time points after treatment with PMSG at various time points were a gift from Dr. Lane K. Christenson.

RNA isolation and RT-PCR. C57BL/6J mice were sacrificed at 6 weeks of age and total RNA was isolated from liver, kidney, spleen, stomach, heart, brain, lung, adrenal, testis, pituitary, and ovary, from immortalized cell lines MA-10, and from primary mouse Sertoli cells, mouse granulosa cells, and P-19 EC cells (with and without RA treatment) using Trizol reagent (Invitrogen Life Technologies, Carlsbad, CA) according to manufacturer recommendations. Human testis total RNA was purchased from BD Biosciences (San Jose, CA). Samples of poly-adenylated mRNA were enriched from MA10 Leydig cell total RNA by oligo-dT cellulose chromatography as described (Sambrook and Russell 2006). Complementary DNA (cDNA) was synthesized from each RNA sample using 2 μg total RNA as described(Daggett, Rice et al. 2000).

For directional RT-PCR to determine transcriptional orientation, cDNA was synthesized from 2μg adult mouse Testis total RNA using species-specific primer 1 or primer 2 (Table 1), and separately with oligo-dT17 in the presence and absence of reverse transcriptase as described

(Heckert et al., 2000). PCR was then performed with the cDNA templates and species-specific oligodeoxynucleotide primers 1 and 2. PCR-amplified products were visualized by agarose gel electrophoresis and analyzed by Southern blot hybridization using species-specific internal primers (primer 3, Table 1) radiolabeled using T4-polynucleotide kinase (New England Biolabs, Ipswich, MA) according to manufacturer recommendations.

For expression profiling, PCR-amplification of *Fast* (primers 1 and 2) cDNAs were performed as described above, while SF-1 cDNA containing exon 1g was detected by PCR using a primer against exon 1g (primer 1; Table 1) and another against exon 5 (primer 2; Table 1). Amplification of L7 cDNA controlled for cDNA synthesis as described (Lei and Heckert, 2002). Amplified products were visualized by agarose gel electrophoresis and, in some cases, *Fast* and SF-1 products were analyzed by Southern blot hybridization with internal oligodeoxynucleotide primers (primer 3; Table 1).

***Fast* cloning.** Mouse EST transcripts were also PCR-amplified with oligodeoxynucleotide primers (primers 4 and 5, Table 1), and cloned using the pGEM T-easy vector system (Promega). The PCR product was cleaned of residual reactions by GeneJet PCR purification kit (Thermo Scientific) and ligation reaction was set up as described by the manufacturer. 2 μl of the ligation reaction was used for transformation and grown on agar plate containing ampicillin/X-gal/IPTG. The white colonies containing the PCR product was chosen and plasmid DNAs were prepared from overnight bacterial cultures using DNA plasmid columns according to the supplier's protocol (GeneJet Plasmid Mini-prep kit, Thermo Scientific). All resulting clones were sequenced. The PCR-amplified product from human testis cDNA was sequenced with human primers 1 and 2.

Mouse *Fast* transcripts were PCR-amplified with oligodeoxynucleotide primer sets engineered with Kpn1 (GCG CGG TAC CCT CGG CCT TCA CCC TCA CCT CCT GGC CCT CCA GTT CCA GCT CGA TC) /Xba1 (CGG CTC TAG ACC AGT TCT GTG CAC CCA CTT TAT GTC TGG) restriction endonuclease sites and double restriction digest performed using Kpn1/ Xba1. The cloning vectors, pcDNA3 (Invitrogen, Carlsbad, CA) was restriction digested with Kpn1/ Xba1; dephosphorylated with calf intestinal phosphatase (New England Biolabs, Ipswich, MA), and ligated with the Kpn1/ Xba1 digested PCR products, respectively, using T4 ligase (New England Biolabs, Ipswich, MA).

Table1: Oligodeoxynucleotide primers

Target	Primer name and sequence	Use
AK017050 & mouse Fast	Primer 1; 5'-CACTCTACGGCATCCCAAGG -3' Primer 2; 5'-CCAGGCAGGCACCAGACC-3' Primer 3; 5'-CCATGTGGGCACAGGGAGGTT-3'	Directional RT- PCR, RT-PCR, southern blot
human *FAST*	Primer 1; 5'-ACTTTATGTGTGGCAAGGTCC-3' Primer 2; 5'-GGAAGCTGGGGCTGGAGGTCT-3' Primer 3; 5'-gaaggaggctctcaggctggg-3'	Directional RT- PCR, southern blot
mouse *Fast*	Primer 4; 5'-CTCGGCCTTCACCCTCAC-3' Primer 5; 5'- CCAGTTCTGTGCACCCACTT -3'	Cloning
mouse SF-1 (exon1g)	Primer 1; 5'-GTCCAGTTTTTCCTTGCTCACC-3' Primer 2; 5'-GCGGTTAGAGAAGGCAGGATAG-3' Primer 3; 5'-GGGGTCTAGAGACCTGGACGAGCTGTGTCC-3'	RT-PCR
mouse *Fast* §	5'RACE1; 5'-GATGTCCCCGAGATTTGGTC-3' 5'RACE2; 5'-AGACCTTGGGATGCCGTAG-3' 5'RACE3; 5'-gcgcacgcgtgtggaattgaccatacatccc-3'	5'RACE
human *FAST* #	5'RACE1h; 5'-GCCCAGCACCATTACAGGAGG-3' 5'RACE2h; 5'-CTGTACTATTTTAAGCCAGGG-3' 5'RACE3h; 5'-gcgcaagcttGTGGGCACAGGGAGGGTAG-3'	5'RACE
mouse *Fast*	3'RACE1; 5'-ACTTTGGTCTGGTGCCTGC-3' 3'RACE2; 5'-gcgcaagcttGCCTGGCTTTCAACCAAAATG-3'	3'RACE
Mouse *Fast*	Primer 6; 5'- TCGCCACAGTCTGACTCTTC -3'	Southern blot
mouse SF-1 exon 1g *	SF-1 e1g; 5'-gcgcggatccGTCCAGTTTTTCCTTGCTCACC-3' SF-1 e4; 5'- gcgcggatccCGTGTAATGCTTGTTGTTCTGG-3'	RNase protection probe
mouse SF-1 exon 1a *	SF-1 e1a; 5'-gcgcggatccGAAGTTTCTGAGAGCCCGC-3' SF-1 e4; 5'- gcgcggatccCGTGTAATGCTTGTTGTTCTGG-3'	RNase protection probe

Mouse Gapdh	mGapdh exon-exon primer F : 5'-AACTTTGGCATTGTGGAAGG-3' mGapdh exon-exon primer R: 5' TGTGAGGGAGATGCTCAGTG-3' mGapdh splice primer F: 5'- GTGCAGGACCTCACTCATTG-3' mGapdh splice primer R: 5'- CACATTGGGGGTAGGAACAC -3'	Subnuclear fractionation
Mouse L7	F: GGGGGAAGCTTCGAAAGGCAAGGAGGAAGCT R:GGGGGGTCGACTCCTCCATGCAGATGATGCC	RT-PCR
Mouse StAR	F: CCG GAG CAG AGT GGT GTC A R: GCC AGT GGA TGA AGC ACC AT	RT-PCR
Mouse Cyp17	F:GCC TGA CAG ACA TTC TG R: TCG TGA TGC AGT GCC CAG	RT-PCR

Determination of transcriptional start site, polyadeylation site and transcript size. 5' Rapid amplification of cDNA ends (5'RACE) was performed as described elsewhere (Frohman, 1993;Frohman, 1990) with modifications using nested oligodeoxynucleotide primers generated against mouse *Fast*. cDNA was synthesized from mouse testis total RNA using a primer within the first exon of *Fast* (5'RACE1) and enhanced AMV reverse transcriptase (Promega) using a ramped synthesis incubation (55 C to 58 C over 55 minutes; Table 1). The cDNA was used for two rounds of PCR amplification using gene-specific nested primers 5'RACE2 and 5'RACE3, and products were cloned and sequenced (Table 1). Transcriptional start sites were assigned based on DNA sequence analysis of the RACE clones. 5'RACE to detect the transcriptional start site of human *FAST* was performed from human testis total RNA as above with primers 5'RACE1h, 5'RACE2h and 5'RACE3h. 5'RACE was also performed to detect 5' 7-methylguanosine capped mRNAs using the same primers described above with the GeneRacer kit (Invitrogen) according to manufacturer recommendations. 3'RACE was performed to determine the polyadenylation site as described (Frohman, 1993;Frohman, 1990) with modifications. In short, cDNA was synthesized from mouse testis RNA as described and used in two rounds of nested amplification using 3'RACE1 and 3'RACE2, and products were cloned and sequenced.

RNAi-mediated *Fast* knockdown: Targets for shRNAi were selected from within the *Fast* cDNA sequence as described (Elbashir et al., 2002) and short hairpin RNA (shRNA) double-stranded oligodeoxynucleotide inserts for each target were designed and cloned into the mU6pro vector as described (Table 2; Yu et al., 2002). A Bgl I-Pvu II fragment from each clone containing the U6 promoter-shRNA cassette or the control U6 promoter-GFP cassette was blunted with Klenow fragment (Roche). Likewise, pcDNA3 (Invitrogen) was digested with Bgl

I to remove the CMV promoter, blunted with Klenow, dephosphorylated with calf intestinal phosphatase as above, and ligated with the U6 expression cassettes with Quick Stick ligase (Bioline). Each U6 cassette- pcDNA3 plasmid was transiently transfected into MA10 Leydig cells as follows. MA-10 cells were plated at a density of 400,000 cells per well in 6-well plates and transfected using 2µg vector DNA, 2µg vector DNA, 12µl lipofectamine (Invitrogen) and 10µl PLUS reagent (Invitrogen). Lipid/DNA complexes were removed after 24h, and the cells were incubated with complete growth media (Waymouth's supplemented with 20 mM Hepes, 15% horse serum, and 50 µg/ml gentamicin). Total RNA was harvested from the transfected MA10 cells and cDNA was synthesized as above and diluted 1:20 in nuclease-free H2O. Expression of *Fast* was analyzed by RT-PCR.

Overexpression studies: *Fast* was cloned in a pcDNA3 vector under the CMV promoter. MA-10 cells were plated at a density of 400,000 cells per well in 6-well plates and transfected using 2µg vector DNA, 12µl lipofectamine (Invitrogen) and 10µl PLUS reagent (Invitrogen). The next day cells were fed with complete growth media. Cells were harvested 48 hr post-transfection, and protein was extracted in 50ul of freshly made RIPA solution.

Western Blot: Western blot analysis was performed as described (Heckert, Sawadogo et al. 2000). Briefly, whole cell extracts from transfected MA-10 cells were extracted in RIPA buffer and resolved by SDS-PAGE, transferred to polyvinylidene fluoride (Millipore Corp., Billerica, MA), blocked, and probed overnight at 4 C with SF-1 antibody (1:3,000 dilution) a kind gift from Dr. Kenichirou Morohashi, Japan,. After incubation with secondary antibody, protein complexes were visualized by chemiluminescence using the enhanced chemiluminescence (ECL) system (Amersham Life Sciences, Arlington Heights, IL). Horseradish peroxidase-

conjugated secondary antibodies were donkey antirabbit IgG (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA). Blots were stripped [100 mM 2-mercaptoethanol, 2% sodium dodecyl sulfate, and 62.5 mM Tris-HCl (pH 6.7)] for 30 min at 55 C according to manufacturer's recommendations (Millipore), and reprobed for Mouse anti-α-Tubulin (Calbiochem, CP06) which was used as loading control.

In vitro transcription translation: The in vitro coupled transcription/translation was done using the T7 rabbit reticulolysate system (Promega). Briefly, the gene was cloned in a pcDNA3 vector downstream of T7 polymerase and was added to rabbit reticulolysate along with ^{35}S-methionine, and incubated for 90 min at 30°C. The in vitro translated products were separated in 15% SDS-PAGE. The gel was dried and exposed to a PhosphorImager screen overnight.

Subnuclear fractionation: Subnuclear fractionation of RNA was done using Norgen Biotek's Cytoplasmic and Nuclear RNA isolation kit. In brief, MA-10, mouse Leydig cells were grown in a monolayer in 100cm plates. Confluent plates were washed with PBS, lysed, and lysates was spun down. The supernatant containing the cytoplasmic RNA was transferred to a new tube, and the nuclear RNA pellet was washed and eluted from the spin columns to a new eppendorf tube.

Southern Blot: RNA was isolated from mouse granulosa cells at various time points post-PMSG, and cDNA was generated as described earlier. RT-PCR was performed by using primers 4 and 5 that span the entire *Fast* locus and products were resolved in a 1% Ethidium Bromide gel for two hours. The bands were then transferred to a Protran nitrocellulose membrane for 3 hours. The membrane was washed twice, pre-hybridized, and probed overnight with Primer 6 located on *Fast* exon3, which was 5' end labeled with γ32P-ATP, by T4 Polynucleotide Kinase (New

England Biolabs). The following day the membrane was washed twice, and bands were detected by exposing it to film overnight.

Figure 1: Two highly conserved ESTs are located between *Nr5A1* (*Ftz-F1*) and *Gcnf* loci. Two conserved ESTs located between the *Ftz-f1* and *Gcnf* loci.Representation of mouse chromosome 2 in the proximity of *Ftz-f1* (*Nr5a1*; blue rectangles) and *Gcnf* (Nr6a1; grey rectangles). Conserved, non-coding ESTs identified by pairwise alignment of mouse and human genomic sequences are marked (Accession #AK007201= EST 12840611; red rectangles; and Accession# AK007201= EST 12840611; green rectangle,). A *Gcnf* variant transcript with an extended exon 12 is depicted by the hatched line extending the Nr6a1 locus.

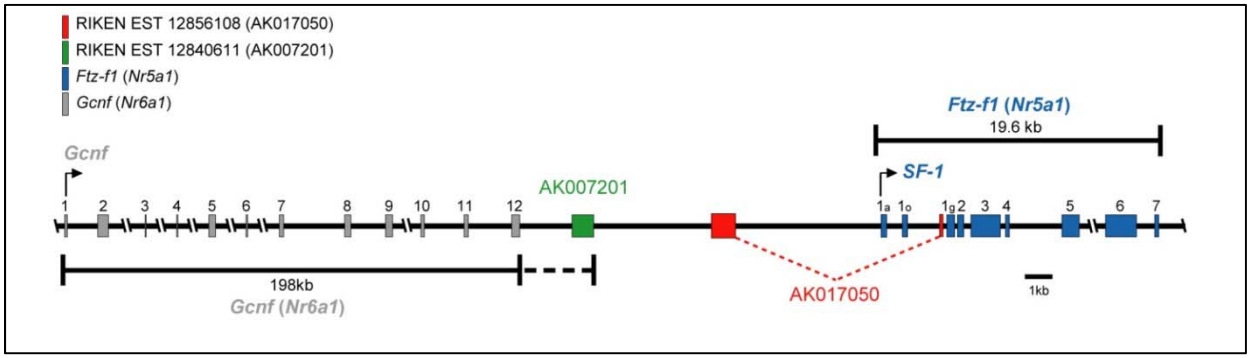


Figure 2: *Fast* is transcribed opposite to *Ftz-F1* (A) Diagram of mouse *Fast* and *Ftz-f1* 5' sequence, indicating exon positions of Ftz-f1 (blue) and FAT (red). (B) PCR amplification of Fast using primers 1 and 2 (locations shown in A) and cDNA synthesized with the indicated primer (top) in the presence (+) or absence (-) of reverse transcriptase. Upper panel shows the resulting ethidium bromide stained agarose gel and the lower panel an auto-radiogram of its Southern analysis hybridized with an internal primer (shown in A). C) Human *FAST* and *FTZ-F1* 5' flanking sequence, indicating positions of *FTZ-F1* (blue) and *FAST* exons (red). (D) PCR-amplification of *FAST* using primers 1 and 3 (locations shown in C) and cDNA synthesized with the indicated primer (top) in the presence (+) or absence (-) of reverse transcriptase . Upper panel is the ethidium bromide stained agarose gel and the lower panel an autoradiogram of its Southern analysis hybridized with an internal primer (shown in C). Bent arrows indicate transcriptional orientation and start sites. The mouse *SF-1* gene has multiple promoters accompanied by first exons, Ia, Ig, and Io, which have been identified in the figure below. (Brian Hermann-Regulation of FSH-receptor and SF-1: transcriptional control in reproduction)

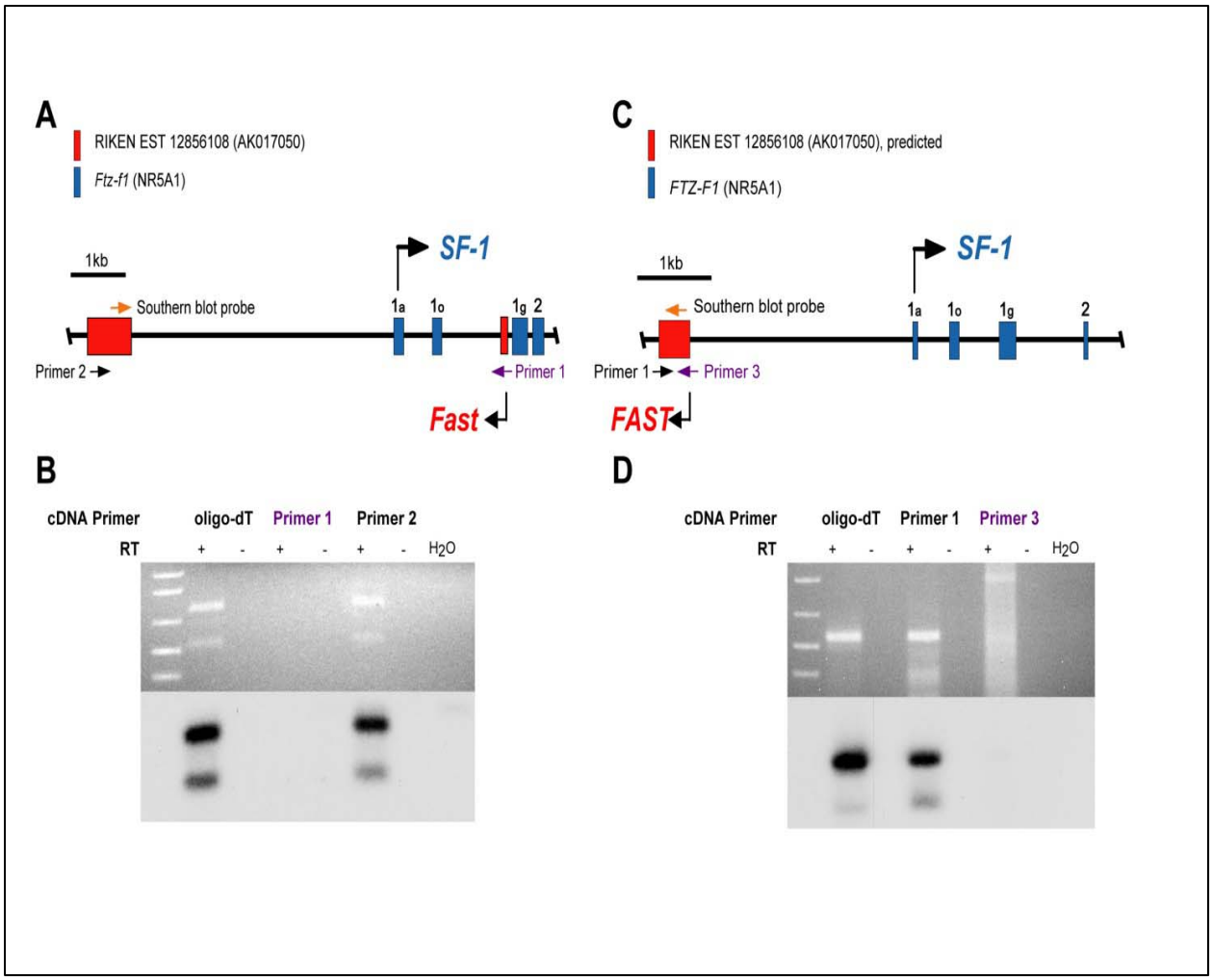


Figure 3: Diagrammatic representation of the genomic location of *Fast* and its transcripts. Top shows the relative position of *Ftz-f1* exons 1-4 (blue) and *Fast* exons 1-3 (orange). Transcriptional orientation and start sites are shown as bent arrows. Human *FAST* contains only the exon that corresponds to mouse exon 3. Bottom shows the two known *Fast* splice variants; one, containing all three known exons, indicates introns as solid lines and the other, lacking exon 2, indicates the intron as a hatched line. (Brian Hermann- Regulation of FSH-receptor and SF-1: transcriptional control in reproduction)

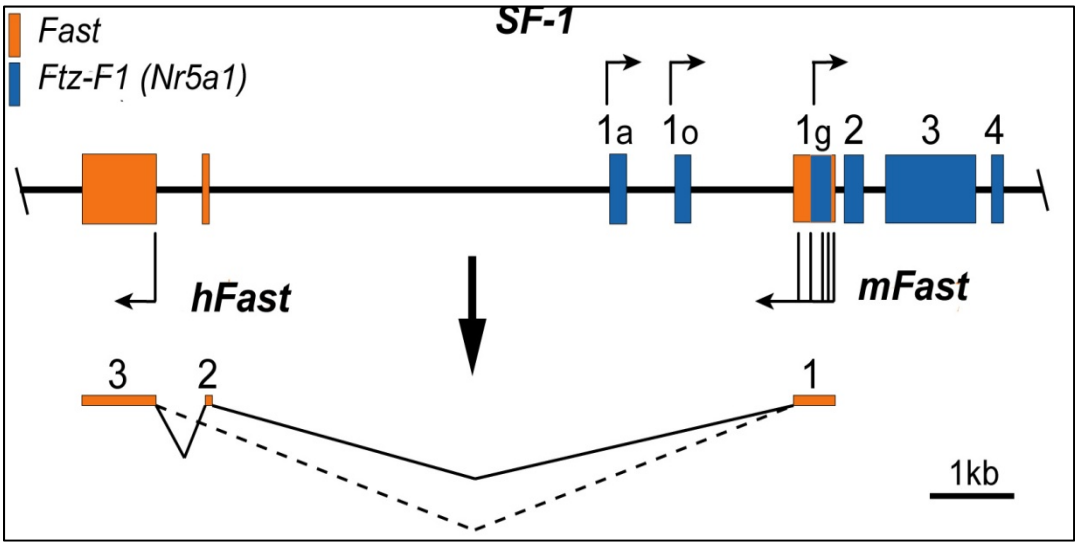
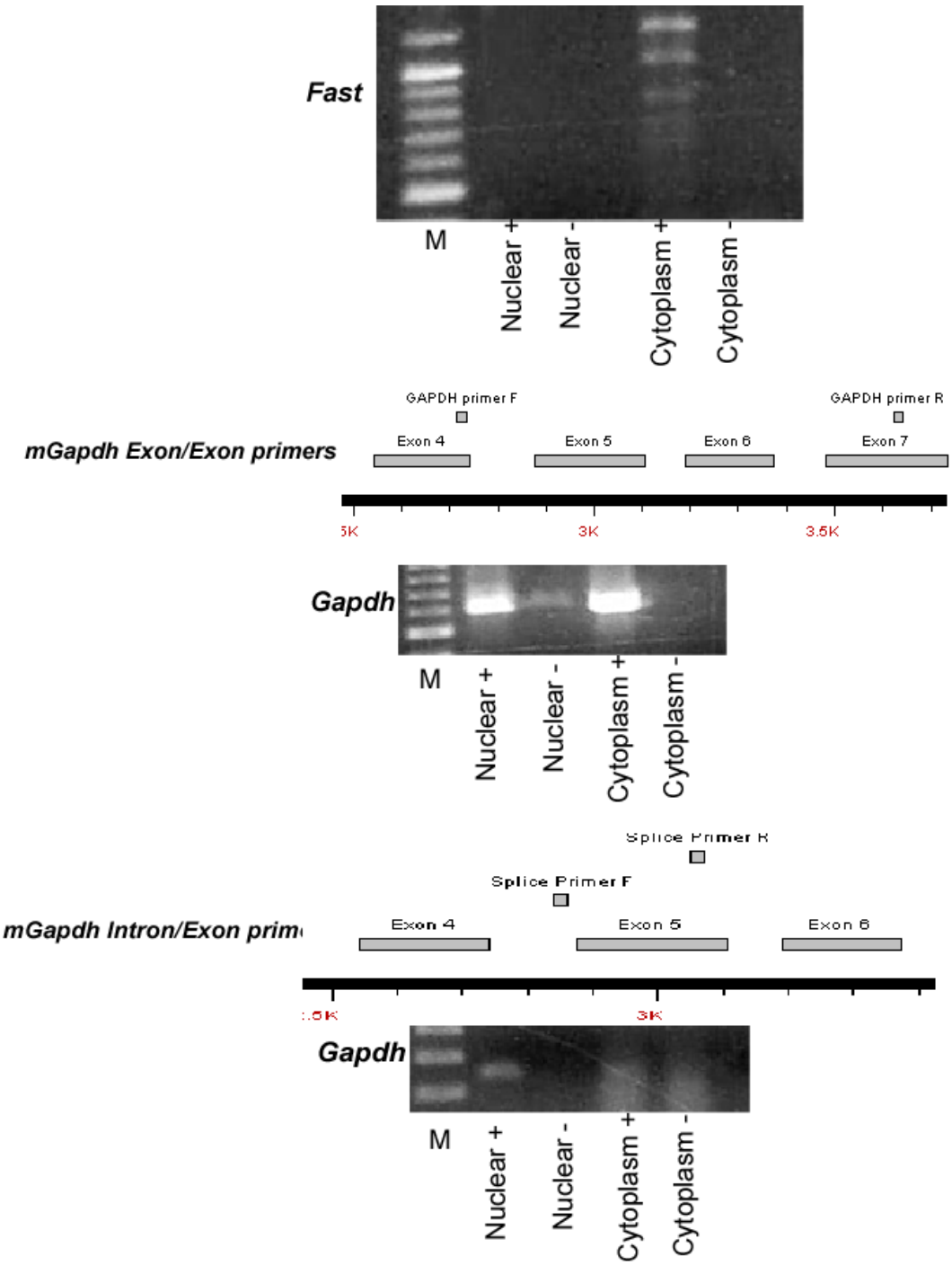


Figure 4: *Fast* localizes predominantly in the cytoplasm. Nuclear and cytoplasmic RNA fractions were isolated from MA-10, mouse Leydig cells, and examined for the presence of *Fast* by RT-PCR (Top). Mouse *Gapdh* primers were designed located in different exons (middle) and within an intron and exon (bottom), to check for purity of cytoplasmic and nuclear fractions, respectively. Experimental repeats shown below with L7 as control.



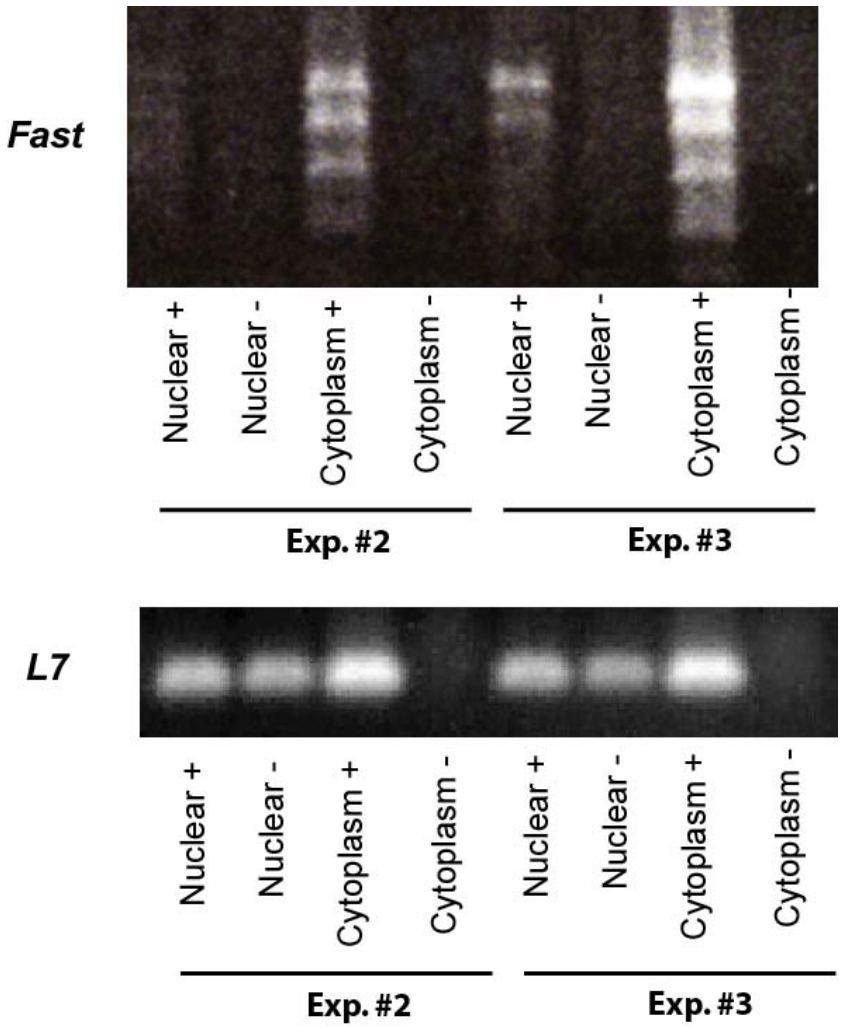
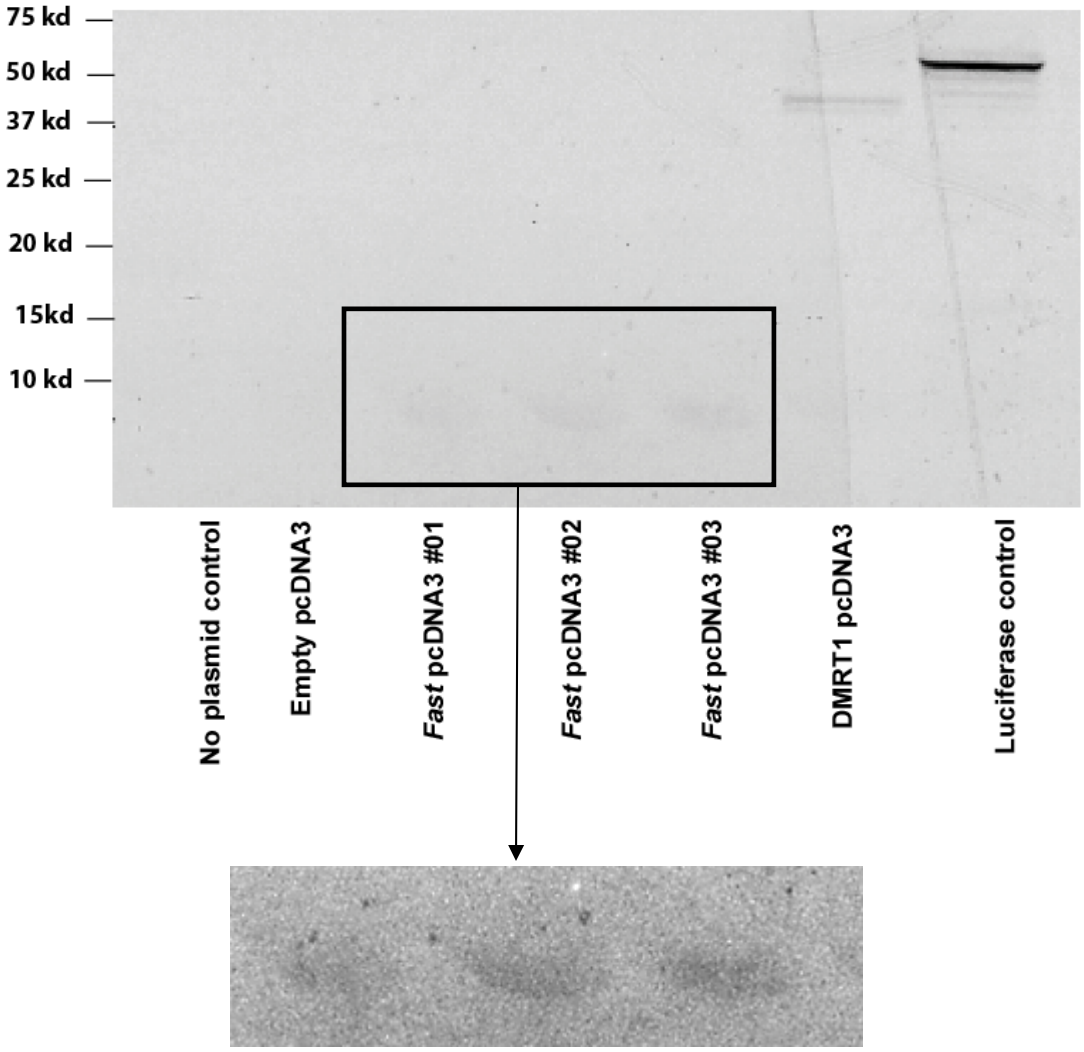

Fast
Nuclear +
Nuclear -
Cytoplasm +
Cytoplasm -
Nuclear +
Nuclear -
Cytoplasm +
Cytoplasm -
Exp. #2
Exp. #3
L7
Nuclear +
Nuclear -
Cytoplasm +
Cytoplasm -
Nuclear +
Nuclear -
Cytoplasm +
Cytoplasm -
Exp. #2
Exp. #3

Figure 5: *Fast* may code for a 7kd protein. Top: In vitro transcription-coupled translation assay of no plasmid control; empty vector negative control; triplicate of *Fast* pcDNA3 (1.7-kb gene cloned into pcDNA3 vector downstream from T7 promoter); DMRT1 pcDNA3 encoding 39-kDa DMRT1 protein; and the positive control luciferase encoding 61-kDa protein. Below: The longest ORF highlighted in red, which may possibly encode for the 7kd protein.



L G L H P H L L A L Q F Q L D P P K P A S S K R I G D P E H P R N P V R C R V R G S F G
P P P A Q G G G D L P F A S P V **Stop** G C R G L N A G R L A A **Stop** K W L R V **Stop** V
G R **Stop** A R K N W T F E L A L P H L P L Q A A G T H S S P V F **Stop** P E A S S P R V W
L K **Met** D V V Q W T A R V P G G E P H Y Q E S P S G C S P G N G **Met** Y G Q F H S T
A S Q G L H S C K N P L V F I P F P S S K P E P R L **Stop** P A L E L P G S L N **Stop** T K S R
G H L C S Y D **Stop** **Stop** P T S **Met** G R V R L W R E D Q R A R E S R V C P G N N G V T
G T F R P P S T H L G S H P **Stop** **Stop** I H E A R S G G L F H V V E E A R A **Stop** G G C R
G E G R K L G L E V L A **Stop** G P P S C L **Stop** G S D C L R K L L G E R R E A **Met** V T S G
K K A Q P P C A H **Met** G F S S P G L K **Stop** C N S A I D A G P S L N L E G L Q P W F V L
P V G A V V **Met** C E G L P S Y L V Q I R L C P V A T G T K T N K Q T N K Q P H R S **Stop**
T **Stop** S F G C K F Q L L H **Stop** S A S R L W S G A C L A F N Q N E D S G **Met** A R S G A F
Stop P N E D R A L P D I K W V H R T

Figure 6: Expression of *Fast* and SF-1 in adult mouse tissue and mouse Leydig cell line, MA10. (A) Top panel, adult mouse tissues and MA-10, cell line were examined for the presence of *Fast*, SF-1 and ribosomal protein, L7, in the presence (+) and (-) of reverse transcriptase. Southern blot utilizing an internal primer for *Fast* is shown in the lower panel. (Brian Hermann- Regulation of FSH-receptor and SF-1: transcriptional control in reproduction)

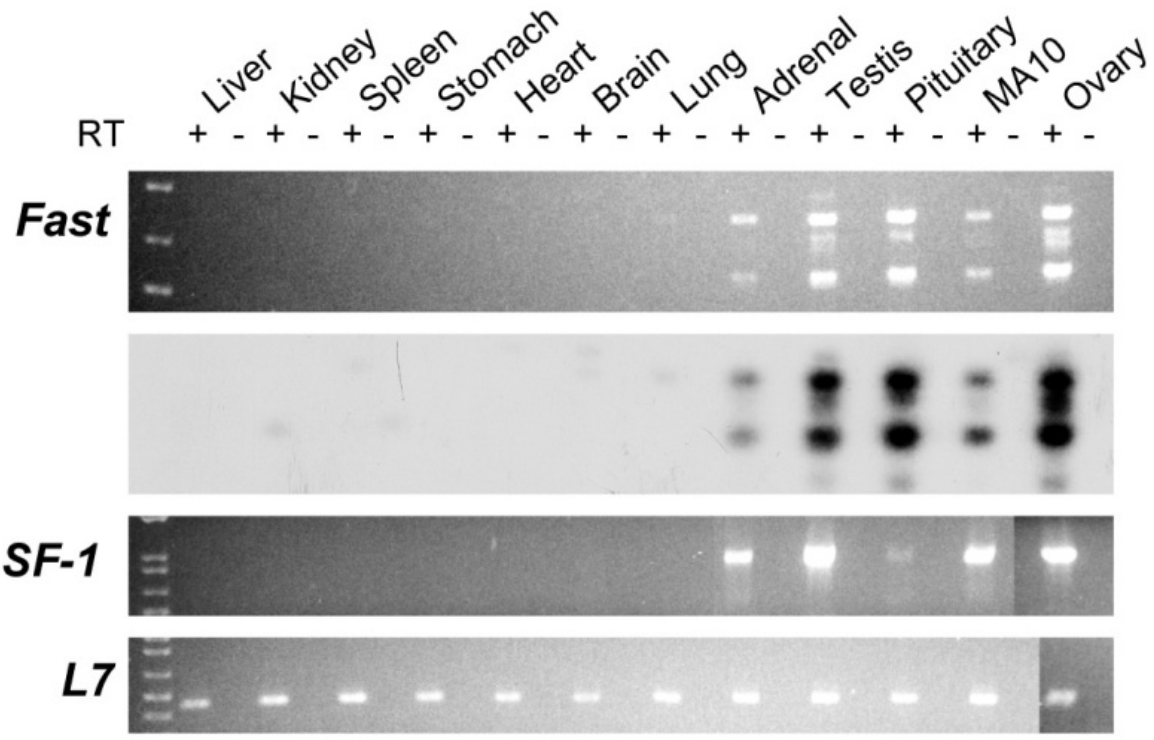


Figure 7: *Fast* and SF-1 respond to retinoic acid in P19 embryonial carcinoma cells. P19 embryonal carcinoma cells were treated with 1μM retinoic acid. RNA was isolated at various time points (top). RT-PCR-amplified products were resolved by agarose gel electrophoresis and stained with ethidium bromide. Assays were performed in the presence (+) or absence (-) of reverse transcriptase (RT). (Brian Hermann- Regulation of FSH-receptor and SF-1: transcriptional control in reproduction)

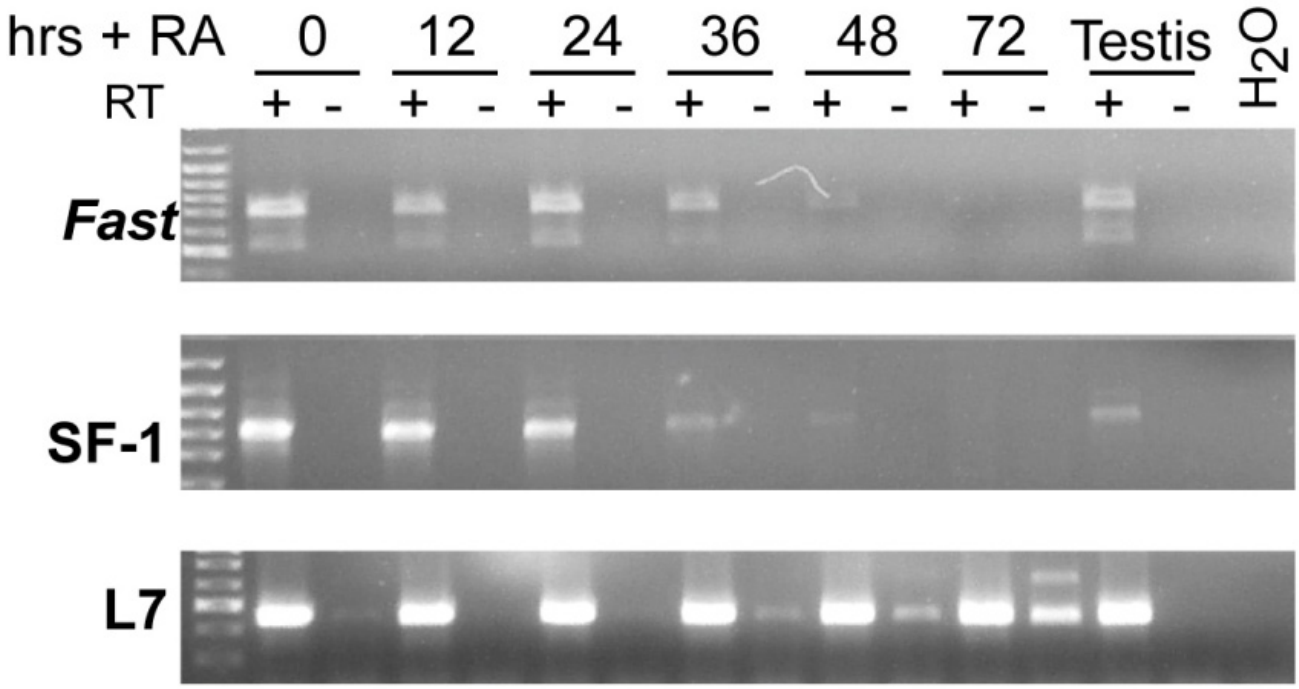


Figure 8: Hormonal response of *Fast* in mouse granulosa cells. Granulosa cells were isolated from ovaries of 19-day-old CF-1 mice at various time points after treatment with PMSG at various time pointed as indicated. RNA was isolated and assayed for fast by RT-PCR followed by Southern blot analysis.

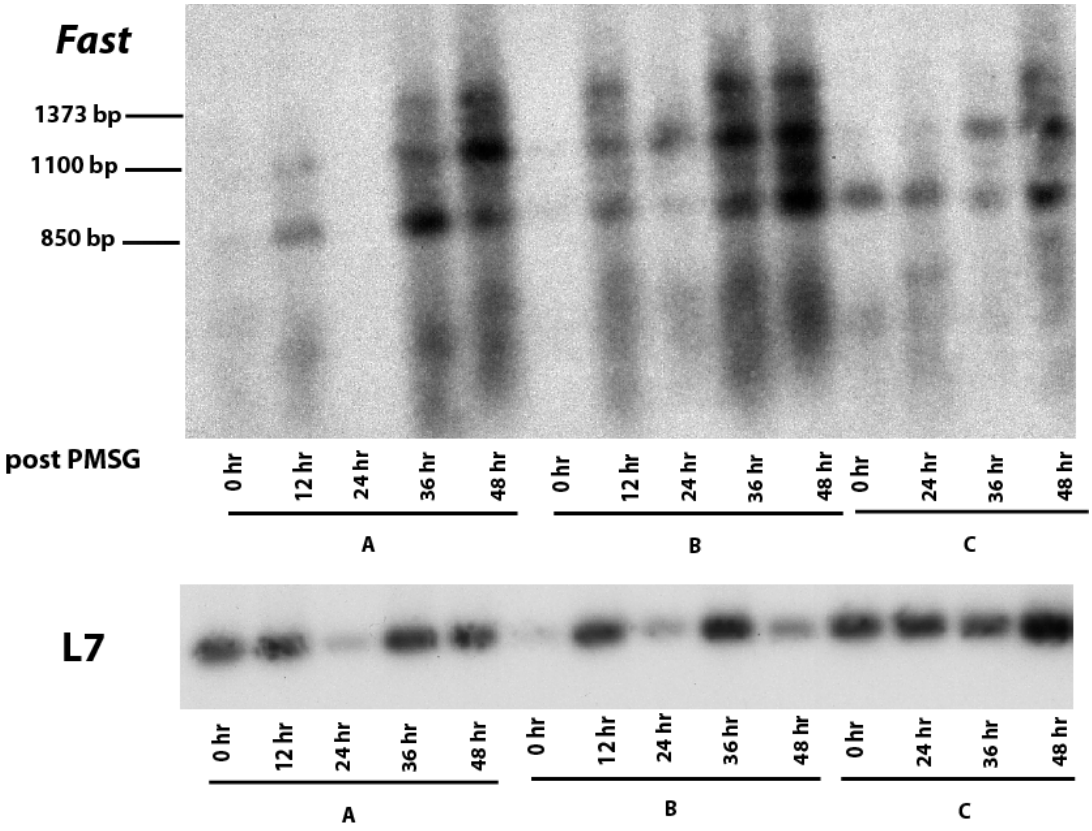


Figure 9: Knocking down *Fast* does not affect transcription of SF-1: **(A)** Position of Fast shRNA 6 target used for RNAi-mediated *Fast* knockdown (blue boxes) is shown on this diagram of the *Fat* cDNA sequence. *Fat* and SF-1 exons (shaded boxes), EST sequence (red boxes), the highly conserved segment of exon3 (85% identity, orange box), and the poly-adenylation signal (red balloon) are shown. Expression vectors containing each shRNA target were transfected intoMA10 Leydig cells and cells were harvested 48 hours post-transfection. **(B)** RNA was isolated, reverse transcribed and assayed for the presence of *Fast* and mouse SF-1. Mouse Gapdh was used as internal control.

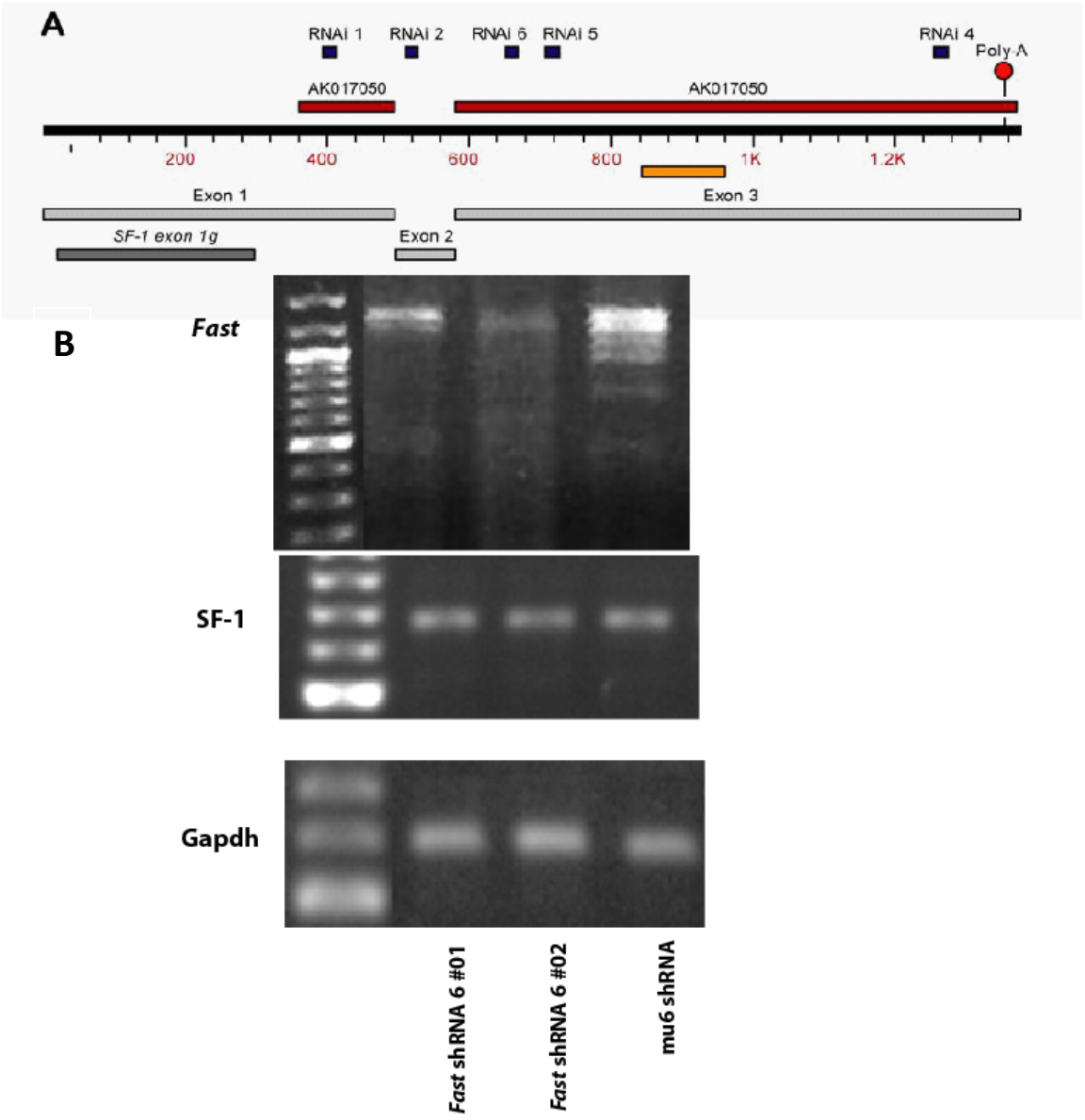


Figure 10: Knocking down *Fast* does not affect downstream targets of SF-1: RNA was isolated, reverse transcribed and assayed for the presence of Cyp17, Cyp 11a, and StAR.Mouse Gapdh was used as internal control.

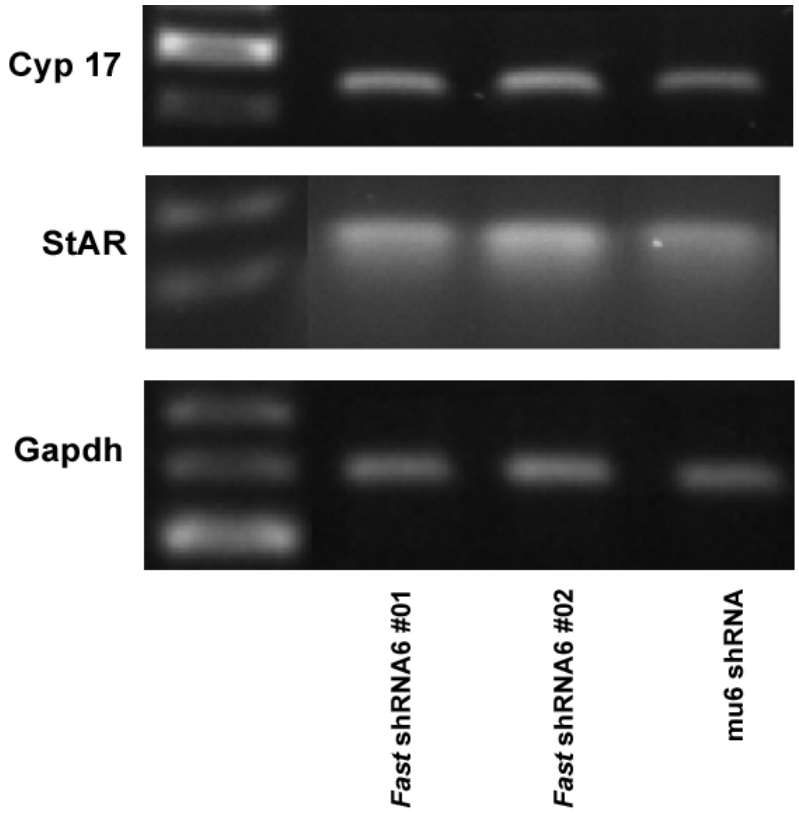


Figure 11: Over expressing *Fast* does not affect SF-1 mRNA levels: RNA was isolated from MA-10 Leydig cells transfected with empty pcDNA3 control and *Fast* pcDNA3. RNA was isolated, reverse transcribed and assayed for the *Fast* and SF-1 Gapdh served as internal control.

Figure 12: Knocking down *Fast* does not affect translation of SF-1: Protein was isolated from MA-10 Leydig cells transfected with mu6shRNA control and *Fast* shRNA#6 and probed for SF-1 levels at various time points, post cAMP as indicated. Tubulin was used as internal control

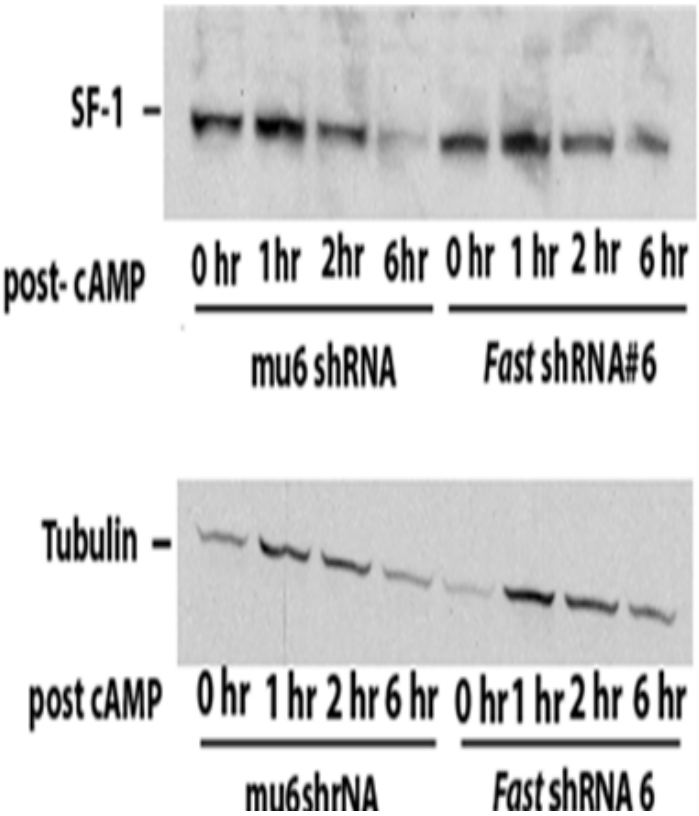


Figure 13: Over expressing *Fast* does not affect SF-1 protein levels: Protein was isolated from MA-10 Leydig cells transfected with empty pcDNA3 control and *Fast* pcDNA3 and probed for SF-1 levels 48 hours post-transfection. Tubulin was used as internal control.

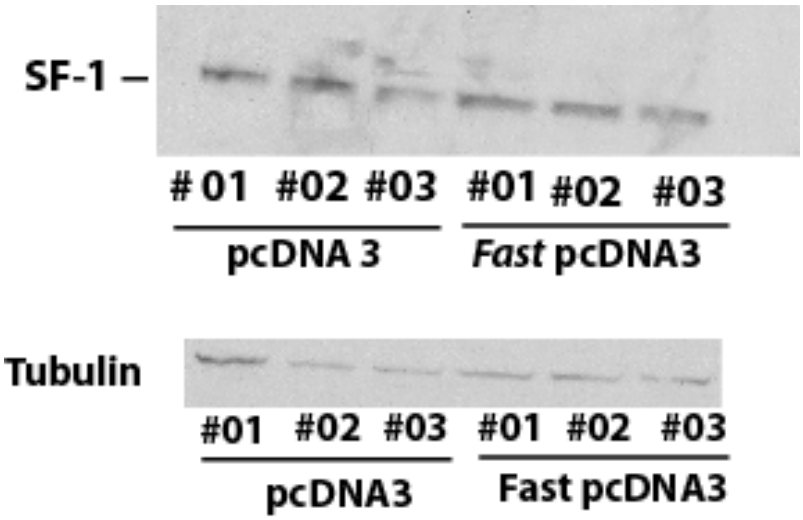


Figure 13: Over expressing *Fast* does not affect SF-1 protein levels: Protein was isolated from MA-10 Leydig cells transfected with empty pcDNA3 control and *Fast* pcDNA3 and probed for SF-1 levels 48 hours post-transfection. Actin was used as internal control.

Chapter 5: An upstream distal Evolutionary Conserved Region regulates the expression of nuclear Receptor Steroidogenic Factor 1 in the pituitary

ABSTRACT

Steroidogenic factor 1 (SF-1/Ad4BP/NR5A1) is a nuclear receptor with a pivotal role in the development of adrenal glands and gonads. Despite many studies that describe transcriptional features of *Nr5a1*, the mechanisms associated with cell-specificity are incompletely resolved for many expressing cell types. As revealed by studies in transgenic mice, these mechanisms require distal regulatory elements residing within a 153kb transgene containing the *Nr5a1* locus. This 153kb transgene defined the initiation point for the current studies. Thus, comparative sequence analysis of this region was used identify sequences of high evolutionary conservation as a means to identify candidate regulatory elements of *Nr5a1*. This identified multiple evolutionarily conserved non-coding regions (ECRs) and four of the most highly conserved were selected for additional analysis to test their transcriptional potential. Each ECR was cloned upstream of the -734/+60 SF-1 promoter directing expression of a luciferase reporter and its transcriptional activity assessed in various cell types by transient transfection analysis. One of the ECRs, ECR3, upregulated SF-1 transcription in alpha T3 (gonadotrope cell line) and Y-1 (adrenal) cells, but downregulated SF-1 in MA-10(Leydig cell line) , MSC-1 (Sertoli cell line) and primary rat Peritubular Myoid and Sertoli cells. Our data indicate that ECR3 contains control elements necessary for SF-1 transcriptional induction in pituitary and adrenal cells and repression in MA-10, MSC-1 cell lines and primary rat Peritubular Myoid and Sertoli cells. Mutagenesis and DNA/protein interaction studies of ECR3 identified sequences important for activity in alpha T3 cells. The ECR3 sequence contains an E box that binds class A basic-helix-loop-helix (bHLH) protein E2A, suggesting cell-specific class B bHLH protein as its likely dimeric partner. Co-transfection studies identified Inhibitors of DNA binding, Id2 and Id3 downregulate SF-1 transcription possibly by interacting with E2A, and preventing its binding to the E-box. Thus,

these studies show that an E box present in the ECR is required for SF-1 expression in the pituitary and involves members of the bHLH protein family.

INTRODUCTION

Steroidogenic factor 1 (SF-1), also known as Ad4BP (adrenal 4 binding protein) and officially designated NR5A1, belongs to the nuclear receptor superfamily. The gene, *NR5A1*, encodes four different proteins, ELP1, ELP2, ELP3 and SF-1. In particular, the transcription factor SF-1 has emerged as a key regulator of reproductive tract development and endocrine homeostasis (Sadovsky, Crawford et al. 1995, Shinoda, Lei et al. 1995, Parker and Schimmer 1997). SF-1 binds to a common regulatory motif (AGGTCA) within promoters and initiates and regulates the expression of various genes. SF-1 was first identified by its ability to regulate the promoter activity of several cytochrome P450 steroid hydroxylase genes (Rice, Mouw et al. 1991, Morohashi, Honda et al. 1992)

The necessity of the SF-1 protein emerged from the development of the SF-1 knockout (KO) mice. The SF-1 KO mice lacked adrenal glands and gonads and died shortly after birth due to adrenal insufficiency (Luo, Ikeda et al. 1994, Hammer, Parker et al. 2005). SF-1 KO mice also exhibited male to female sex reversal, impaired gene expression within pituitary gonadotropes and structural abnormalities of the ventromedial hypothalamic nucleus (VMH) (Ingraham, Lala et al. 1994). Data obtained from both global and pituitary-specific SF-1 KO mice support a role for SF-1 in gonadotrope function and transcriptional regulation of the gonadotropins, luteinizing hormone and follicle-stimulating hormone, as both are absent in mice deficient of SF-1(Zhao, Bakke et al. 2001). Gonad specific knockout mice demonstrated that SF-1 action in the gonads is essential for proper gonadal function in both male and female mice (Jeyasuria, Ikeda et al. 2004).

In humans, mutations in SF-1 resulting in a loss of activity are also associated with male to female sex reversal and adrenal failure (Achermann, Ito et al. 1999, Biason-Lauber and Schoenle 2000, Achermann, Ozisik et al. 2002). Thus, both *in vitro* and *in vivo* evidence establish SF-1 as a critical regulator of development and endocrine homeostasis in humans and mice.

Given the critical role played by SF-1 in development and endocrine homeostasis, many studies have focused on uncovering the regulatory mechanisms that specify *SF-1* transcription in the appropriate spatio-temporal manner. However, despite the efforts, little has been revealed about the specific transcriptional mechanisms that drive tissue-specific SF-1 expression. Transient transfection studies, using the SF-1 promoter in front of a reporter, have revealed several cis-regulatory elements that control basal transcription. In particular, the promoter has an E box, a CCAAT box, and three Sp1 binding sites (Barnhart and Mellon 1994, Nomura, Bartsch et al. 1995, Daggett, Rice et al. 2000). Further studies on the proteins binding to these elements uncovered a variety of binding complexes as well as cell-specific interactions (Daggett, Rice et al. 2000, Scherrer, Rice et al. 2002). While these findings have shed light on basal function of the SF-1 promoter, they have not revealed any mechanisms to explain how SF-1 expression is restricted to different target tissues or how its transcriptional regulation contributes to its role in development and endocrine homeostasis.

The development of transgenic mice has lead to the identification of specific DNA elements, which direct SF-1 transcription *in vivo*. Accumulated evidence indicates that proper expression of SF-1 requires a large genomic fragment that spans a 153kb region of the *Nr5a1* locus. (Wilhelm and Englert 2002) identified a 674bp region on the SF-1 promoter that partially directs expression to the gonads. However, the study also revealed that the 674bp promoter region

lacked key regulatory sequences needed for SF-1 expression in most of its target tissues (Wilhelm and Englert 2002). A transgenic study by Stallings *et al.,* targeted a 50kb region of mouse *Nr5a1*, including 45kb of 5' flanking sequence, exon 1 and part of exon 2, to direct expression of green fluorescent protein (Stallings, Hanley et al. 2002). This 50kb transgene was able to direct expression to many but not all SF-1-expressing cells. Our laboratory demonstrated that a 153kb yeast artificial chromosome (YAC) transgene containing rat *Nr5a1* completely mimicked expression of endogenous *Nr5a1*and rescued all known defects of SF-1 null mice (Karpova, Presley et al. 2005). Together, these studies indicate that tissue-specific expression of SF-1 is directed by regulatory elements located at significant distances from the transcription start site. Since the 153kb transgene used in our previous studies contains all the sequences necessary for proper SF-1 expression, we used this transgene to identify and characterize essential distal regulatory elements. Here, we report the identification and characterization of an Evolutionary Conserved Region (ECR), ECR 3. ECR3 contains regulatory elements that enhance transcriptional activity of the SF-1 promoter specifically in gonadotrope and adrenocortical-derived cells.

RESULTS

Comparison of *NR5A1* between human and chicken genomic sequences identifies Six Major Evolutionarily Conserved Regions.

To help identify potential regulatory elements essential for SF-1 expression, a web-based sequence analysis tool, ECR browser was used (Ovcharenko, Nobrega et al. 2004). The ECR browser identifies sequence conservations between various species that have remained unchanged for millions of years, suggesting that the predicted sites are functionally important.

Comparison of rat, mouse, frog and chicken *Nr5a1* identified more than twenty-five ECRs. (Figure 1). To improve functional prediction, a more distant genome (chicken) was included in the comparative analysis. This sequence conservation analysis between chicken, mouse, and human led to a more stringent identification of 6 ECRs which were chosen for further analysis, as they represent the most conserved, and thus most likely, regions to contain essential regulatory elements.

ECR3 differentially regulates SF-1 transcriptional activity in αT3 and Y-1 cells versus MA-10, Myoid, MSC-1 and primary Sertoli cells.

To identify functional importance of the ECR's, four ECRs (ECRs 1-4) were cloned into a vector that contains the SF-1 promoter, from -734 to +60, directing expression of the firefly luciferase reporter. Transcriptional activity was tested by transiently transfecting the vector in mouse Leydig tumor cell line MA-10, mouse sertoli cell line MSC-1, Y-1 murine adrenocortical cells, and mouse gonadotroph-derived alpha T3 cells, as well as primary cultures of rat Sertoli and Peritubular Myoid cells. With TK-renilla as an internal control, firefly and renilla luciferase activities were compared to determine transcriptional activity. The luciferase/renilla values of the vectors containing ECRs were normalized to the values for the SF-1 promoter alone. Of the four ECRs, ECR3 had the most dramatic effect (Figure 2). ECR3 increased SF-1 transcriptional activity in αT3 and Y-1 cells, but decreased transcriptional activity in MA-10, MSC-1 and primary rat Peritubular Myoid and Sertoli cells. This data indicates that ECR 3 likely contains elements important for SF-1 regulation and acts in a cell-specific manner, activating SF-1 transcription in pituitary and adrenal cells, while repressing its activity in Leydig, Sertoli, and myoid cells.

ECR3 contains potential regulatory elements and key regions necessary for full enhancer activity.

Next, we decided to identify regions of ECR3 that are important for the enhancer activity seen in αT3 and Y-1 cells. To this effect, deletion mutants were generated by cloning varying lengths of ECR3 based on conservation profile and size, namely ECR3 LC(less conserved), ECR C (highly conserved), ECR3-300, ECR3-350, ECR3-504, and ECR3-580 into SF-1(-734+60)Luc vector. The transcriptional activities of these mutants were then assayed in αT3 pituitary cells, as ECR3 showed highest enhancer activity in these cells (Figure 2). The less conserved region within ECR3 showed very minimal activity when compared to that of full-length ECR3 (Figure 3). ECR3 C showed increased activity compared to ECR3 LC, but not to the level of full-length ECR3. None of the additional constructs tested (ECR3-300, ECR3-350, ECR3-504, and ECR3-580) reached the activity of full-length ECR3. These results suggest that regulatory elements reside within the ECR3 locus and their interactions are necessary for complete activity.

To further categorize ECR3, DNase I footprinting was performed with and without αT3 cell nuclear extracts using a probe as indicated in Figure 4. Addition of nuclear proteins revealed a prominent hypersensitive site and two protected footprints. *In silico* sequence analysis using the web-based program rvista (http://rvista.dcode.org/) identified several evolutionarily conserved sequences that are potential transcription factor binding sites, E box, p53 tumor suppressor protein and E twenty-six (ETS)-like transcription factor 1, also known as Elk1 were chosen based on the scores.

An E box is required for full transcriptional activity of ECR3.

To evaluate the effects of these transcriptional sites in αT3 cells, potential transcription factor-binding sites identified above, E box, p53 and Elk1, were mutated and were transiently transfected in αT3 cells. These cells were then assayed for relative luciferase activity as before, and activities were graphed (Figure 5). Of the three transcription factor binding sites tested, mutation of the E box element significantly decreased the transcriptional activity, suggesting functional importance.

The E box site identified was **CATCTG** (E box elements have the consensus sequence **CANNTG)**, a sequence that binds basic helix-loop-helix (bHLH) proteins. bHLH proteins are transcription factors, divided into two broad functional groups by their patterns of expression. Class A bHLH members, often also referred to as E-proteins, which are ubiquitously expressed and bind DNA as homo-and hetero-dimers. In contrast, Class B proteins are expressed in cells of particular lineage (e.g., MyoD and Myogenin) and mainly bind as heterodimers . These proteins have a helix–loop-helix domain that mediates homo- and/or hetero- dimerization with other HLH proteins. To help identify the bHLH protein that binds to E box, we performed an electrophoretic mobility shift assay (EMSA) using αT3 nuclear extracts and a radiolabeled probe containing the ECR3 E box. Unlabeled competitors with sequence homology were added to the reaction at a concentration 100X that of probe. Similarly, antibodies to bHLH proteins, Upstream stimulatory factors 1 and 2, (USF1 and USF2) or transcription factor E2A, a ubiquitous bHLH protein, were added to specific reactions. Notably, formation of binding complexes was inhibited by inclusion of unlabeled homologous competitor DNA but not a non-specific (N sequence, indicating the proteins bind specifically to the ECR3 E-box probe sequence (Figure 6). However, a competitor containing a mutation in the E box did not compete for the fastest and slowest migrating

complexes and was a poorer competitor for some of the complexes in the middle. Thus, the top and bottom complexes are appear to be dependent on the E box sequence for binding and are candidates for eliciting the transcriptional activity associated with E box. As indicated by their cross-reactivity to included antibodies, these candidate proteins are USF1, USF2, and E2A, plus an unidentified binding partner.

Chromatin immunoprecipitation (ChIP) analysis was employed to analyze transcription factor binding toSF-1 ECR3 site, *in vivo*. Formaldehyde cross-linked chromatin from αT3 cells was used for immunoprecipitation with antibodies to E2A, USF1, USF2, RNA polymerase II, p300, histone active mark H3Ac, and repressive histone marker H3K27me3. With chromatin prepared from αT3 cells, E2A, USF1 and USF2 were found to bind to the ECR3 region and were congruent with EMSA results (Fig 7). In contrast, p300 an enhancer marker was found not enriched at Sf-1 ECR3 site. Active histone mark H3Ac was found to be significantly enriched with lower binding of H3K27me3 indicating that the region was in an "open" chromatin state.

Id2 and Id3 inhibit transcriptional activity through the ECR3 E box

Having identified that bHLH proteins bind to the E-box, we then decided to establish whether Id proteins inhibit transcriptional activity by binding to the E-box. Ids are a group of HLH proteins that lack a DNA binding domain and act as negative regulators of bHLH proteins (Ruzinova and Benezra 2003, Perk, Iavarone et al. 2005) . To investigate Id proteins that inhibit transcriptional activity, Id1, Id2, and Id3 were cloned and co-transfected in αT3 cells. Increasing concentrations of empty vector or expression vectors for Id1, Id2, and Id3, were co-transfected into αT3 cells with luciferase reporters driven by wild type (Trompouki, Bowman et al.) or E box mutant

(mEbox) SF-1(-734+60)-ECR3. Relative reporter activity was assayed for each line (Figure 7). Co-transfection of SF-1(-734+60)-ECR3 with plasmids expressing Id2 and Id3 showed a dose dependant diminished SF-1 promoter activity and this change in transcription levels was completely lost when Id plasmids were transfected with mEbox SF-1(-734+60)-ECR3 construct. In contrast, expression of Id1 expression plasmid had little or no impact on the transcription level of SF-1. This data demonstrates that Id2 and Id3 proteins inhibit the activity of the SF-1 transcription and it does so by inhibiting the binding of the bHLH proteins to the E-box in the ECR3 distal region.

DISCUSSION

Nuclear hormone receptor, SF-1, is a key regulator of endocrine homeostasis and development. Many studies have investigated the expression profile of SF-1 and the mechanisms associated with its strict tissue specificity. However, our knowledge of the transcriptional regulatory mechanisms responsible for proper spatiotemporal expression of SF-1 has been limited. A previous study from our lab established that a 153kb region of the *NR5A1* locus contains all the necessary control elements for SF-1 activity (Karpova, Presley et al. 2005). We identified Evolutionarily Conserved Regions that serve as potential regulatory sequences important for directing the spatiotemporal expression of SF-1.

DNA sequence conservation is a characteristic that has long been associated with functional regions of the genome. More recently, the importance of conserved non-coding sequences in transcriptional regulation has been validated by numerous studies (Frazer, Sheehan et al. 2001, Visel, Blow et al. 2009, Blow, McCulley et al. 2010). Sequence comparisons between different species have often shown that functional non-coding sequences are highly conserved, whereas sequences that are not functional diverge (Hardison 2000). The location and function of regulatory sequences that orchestrate gene expression remain obscure, which makes it difficult to study their role in developmental processes (Pennacchio, Ahituv et al. 2006, Prabhakar, Poulin et al. 2006, Holland, Albalat et al. 2008, Visel, Prabhakar et al. 2008). A widely used approach is comparative genomics of both closely related and highly divergent organisms to identify specific control elements (Gottgens, Barton et al. 2000, Thomas, Touchman et al. 2003, Chapman, Donaldson et al. 2004). We employed the same approach in comparing different genomes ranging from human to chicken to identify six ECRs as potential regulators of SF-1.

While non-coding sequences under evolutionary pressure could predict the location of enhancer elements in the genome, they do not reveal when and where these enhancers are active *in vivo*(Visel, Blow et al. 2009). To test the functional importance of these sequences, one can use different experimental strategies, including *in vitro* or *in vivo* assays, like transient transfection or transgenic studies. In the present study, we tested the transcriptional activity of SF-1 as regulated by four ECRs, using transient transfection reporter assays in various cell types. Of these ECRs, ECR3 showed strong cell-specific up regulation in αT3 pituitary and Y-1 adrenal cell types; however, transcriptional repression was observed in testis cell types.

ECR3, like other critical regulatory elements, contains elements identified through sequence conservation. For example, DNaseI hypersensitivity studies of cis-regulatory elements showed that about 70% of regulatory sequences are found at least 5kb from the transcriptional start site of the gene(Sabo, Humbert et al. 2004). Interestingly, ECR3 is located 8kb upstream of the transcriptional start site. The observed characteristics of ECR3, including evolutionary conservation, cell-specific activity and distance relative to SF-1 transcription start site, triggered us to further characterize ECR3.

Previous studies employing transgenics from the laboratory of Ken Morohashi identified many tissue-specific regulatory regions within the SF-1 gene (Shima, Zubair et al. 2005, Zubair, Oka et al. 2009, Shima, Miyabayashi et al. 2012). Using a 5.8kb construct containing Exon 2 and the upstream promoter region, enhancers for fetal adrenal gonad and VMH were found in introns 4 and 6, respectively, in two different studies (Shima, Zubair et al. 2008). Interestingly, these regions correspond to two of our ECRs, ECR2 and ECR1. Another transgenic study from the same group identified a 5kb fragment that contained Rathkes pouch -specific enhancer activity

(Shima, Zubair et al. 2008) . Further studies, uncovered a 486bp enhancer element in intron 6 of the gene, that is necessary for gonadotrope-specific enhancer. This enhancer element was reported to contain several conserved sequences. In contrast, however, the ECR we identified is a non-coding region (ECR3) that contains regulatory sequences specific for SF-1 expression in pituitary.

In vitro foot printing identified binding sites for several proteins that, when mutated, revealed their importance for transcriptional activity of ECR3. In addition, *in silico* analysis of footprinted sequences identified potential transcription factor binding sites for p53 and members of the ETS- and bHLH families. Further analysis of these sites revealed that an E-box element is essential for the activity of ECR3 in pituitary gonadotropes. E-box elements have been known to play an important role in the expression of many genes in different organs. Our current study, combined with the study of Shima et al, demonstrates that, pituitary-specific expression of SF-1 involves a highly complex transcriptional regulatory network and involves multiple transcriptional control sequences (Shima, Zubair et al. 2008).

We hoped to identify the protein that binds to the E-box (critical for pituitary-specific SF-1 expression), and the most likely candidate is the ubiquitous E2A, a basic helix-loop-helix protein (bHLH). bHLH proteins are transcription factors that play important roles in various developmental processes, including sex determination (Zheng, Wang et al. 2009). They contain a basic DNA binding domain and two helices that are involved in interaction with other proteins during the formation of homo- or hetero-dimers (Massari and Murre 2000). E2A hetero-dimerizes with other tissue-specific bHLH proteins before binding to DNA to regulate transcription. E2A has been extensively studied and is widely believed to play a central role in

transcriptional regulatory networks (Kee 2009). Although, E2A is expressed in pituitary, no studies account for the pituitary functions of E2A. Our current study indicates that E2A plays a transcriptional regulatory role in pituitary gonadotropes (Roberts, Steenbergen et al. 1993) .

Inhibitors of DNA binding/differentiation are another type of bHLH proteins, which inhibit helix-loop-helix activators by acting as repressors (Ruzinova and Benezra 2003, Perk, Iavarone et al. 2005) . To test whether Ids repress SF-1 activity through ECR3, we performed a co-transfection with ECR3, ECR3 mutant and increasing amounts of Id proteins. Out of three Ids tested Id1, Id2 and Id3 only Id2 and Id3 showed inhibitory effects on SF-1 promoter activity. As expected, co-transfection with the mutant ECR3 did not alter the activity of SF-1. Although, all Id proteins are closely related in structure and were thought to show similar binding affinity to other bHLH proteins, a recent study using gene targeting revealed the difference between these Id proteins. Id1 null mice were normal. However, Id2 null mice lacked lymph nodes and had severe defects in development of natural killer cell lineage, while Id3 null mice had defects in B-cell immune response (Lyden, Young et al. 1999, Pan, Sato et al. 1999, Yokota, Mori et al. 2001). Id2 and Id3, therefore have greater functional importance than Id1. The difference in the inhibitory effects of different Id proteins in the present study could be attributed to interactions of different Ids with E2A and its unidentified binding partner.

In summary, our study identified a highly evolutionarily conserved region namely ECR3. Further characterization of the region revealed that an E-box within ECR3 is essential for pituitary expression of SF-1. Additional studies will identify more or all of the regulatory sequences that are important for pituitary expression of SF-1 and increase our understanding on SF-1 gene regulation in pituitary.

MATERIALS AND METHODS

SF-1(-734+60)Luc vector preparation: The SF-1 promoter region from approximately -734 to +60 was amplified by polymerase chain reaction (PCR) from rat primary Sertoli cell DNA using primers shown in Table 1 and Bio-X-act DNA polymerase (Bioline USA Inc.MA, USA) and standard procedures. An *XhoI* site was incorporated into the 5' primer and a *HindIII* site was incorporated into the 3' primer. The resulting PCR product was digested with *XhoI* and *HindIII* and cloned into pGL3 basic (Promega) .

Identification and cloning of ECRs:

ECRs were identified using a web-based sequence analysis tool, ECR browser (http://ecrbrowser.dcode.org/),(Ovcharenko, Nobrega et al. 2004). The ECRs were amplified by PCR from rat genomic DNA isolated from primary Sertoli cells using primers shown in Table 2 and Bio-X-act DNA polymerase (Bioline USA Inc.MA, USA) and standard procedures. The primers used contained sites for the restriction endonuclease *KpnI*. The amplified ECRs, following digestion with the Kpnl enzyme, were cloned into SF-1(-734+60) Luc which contains the rat SF-1 promoter, from -734 to +60, directing expression of the firefly luciferase reporter. The ECR DNA fragments were cloned upstream of the SF-1 promoter insert. All clones were confirmed by sequence analysis.

Identification of potential transcription factor binding sites:

The web-based tool "rVISTA," which combines TFBS predictions, sequence comparisons and cluster analysis to identify non-coding DNA regions that are evolutionarily conserved, was used to analyze ECR3 sequence for potential transcription factor binding sites (Ovcharenko, Nobrega et al. 2004).

Site directed mutagenesis:

For deletion mutagenesis, specific lengths (Highly conserved, Less conserved, 300, 363, 500, and 580 bp) of the sequences of ECR3 were PCR amplified using primers containing *KpnI* sites and cloned into SF-1(-734+60)luc, as described above. For PCR based site-directed mutagenesis, a *HindIII* site was introduced into the potential transcription factor binding sites (E box element, p53 and Elk1) and the mutant ECRs were amplified using primers containing *KpnI* sites and cloned into SF-1(-734+60)Luc as for the full length ECR3. All clones were confirmed by sequence analysis.

DNase I footprint and EMSA analysis:

The preparation of nuclear extracts and the generation of DNA probes were as described elsewhere (Lei and Heckert 2004, Hermann and Heckert 2005). DNase I foot printing assay using nuclear extracts from αT3 cells (containing 20 micrograms of protein) was performed as previously described (Lei and Heckert 2004). EMSAs were performed as previously described (Lei and Heckert 2004, Hermann and Heckert 2005).

ChIP

ChIP was performed as described (Hiroi, Christenson et al. 2004, Hermann and Heckert 2005). Cross-linked chromatin was prepared from αT3 gonadotrope cells. Antibodies were purchased from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA), and 5 μg each was used for immunoprecipitation: rabbit anti-USF1 IgG (sc-229), rabbit anti-USF2 IgG (sc-862), rabbit E2A.E12 (V-18) X (sc-349X) , RNA polymerase II (sc-899), p300 (SC-6149), H3Ac (Millipore 06-599), H3K27me3 (Millipore 07-449), and normal rabbit IgG (sc-2027). PCR was performed

using 5 µl immunoprecipitated DNA or 1 µl input material, and primers were directed to sequences within the mouse SF-1 ECR3 loci. For mouse SF-1 ECR3, the primer set spanned the ECR3 region, ECR F: AGCAGCAGAGGCTGTTTCC and ECR R: AGCTGAGGCTCCCTCCAC. Products were resolved by agarose gel electrophoresis.

Cell Culture and Transfections: Primary cultures of rat Sertoli and peritubular myoid cells were prepared, and transfections performed, as previously described (Karl and Griswold 1990, Heckert, Daggett et al. 1998, Chen and Heckert 2001) Preparation and transient transfection of mouse MA-10 Leydig cells was performed as described, except that cells were plated at a density of 52,000 cells/well (Ascoli 1981). Preparation and transient transfection of the αT3 gonadotrope cells were performed as described, with modifications (Windle, Weiner et al. 1990, Wolfe and Call 1999). Preparation and transient transfection of the MSC-1 cells were performed as described earlier (Heckert, Daggett et al. 1998). Co-transfection of Ids was performed as described, with modifications (Scherrer, Rice et al. 2002).Transfection experiments included TK-renilla as an internal control and the firefly and renilla luciferase activities were measured. The luciferase/renilla values of the vectors containing ECRs were made relative to the values for the SF-1 promoter alone.

Table 1. Oligos used to amplify SF-1 Promoter (-734+60) region

Primer Sequence(5'-3')
SF-1 F GGGCTCGAGATCCGTCTAGGCCAGTTCAG
SF-1 R GGGAAGCTTCTATCGGGCTGTCAGGAACT

Table 2. Oligos used to amplify the ECRs

Primer Sequence(5'-3')
ECR 1 F GCGCGGTACCACTTCCAGTCCGCCTGCTCGTG
ECR 1 R GCGCGGTACCGGACTGGGACCCTTGCCGAG
ECR 2 F GCGCGGTACCTGCTCGGAGAGATGGTTTATTA
ECR 2 R GCGCGGTACCCCTGGCTTGGGGTCCCTGGC
ECR 3 F GCGCGGTACCGGTAATGCTGGCAGGTTGGGAT
ECR 3 R GCGCGGTACCTGGAGGCAGAAAATGAACTAA
ECR 4 F GCGCGGTACCCTCTGCCCAGGACAAACCC
ECR 4 R GCGCGGTACCCAACT TTGGTTTCTTCATTTACA

Figure 1: Alignment of NR5A1 region using ECR browser, depicting the Evolutionarily conserved regions (ECRs)**.**Alignment of NR5A1 region using ECR browser, depicting the Evolutionarily conserved regions (ECRs). In this figure, red peaks depict regions of non-coding evolutionarily conserved sequence, *NR5A1* and its neighboring gene *NR6A1* are annotated at the top (arrows indicate transcription direction), and conserved exons and untranslated sequences are marked as blue and yellow peaks, respectively.

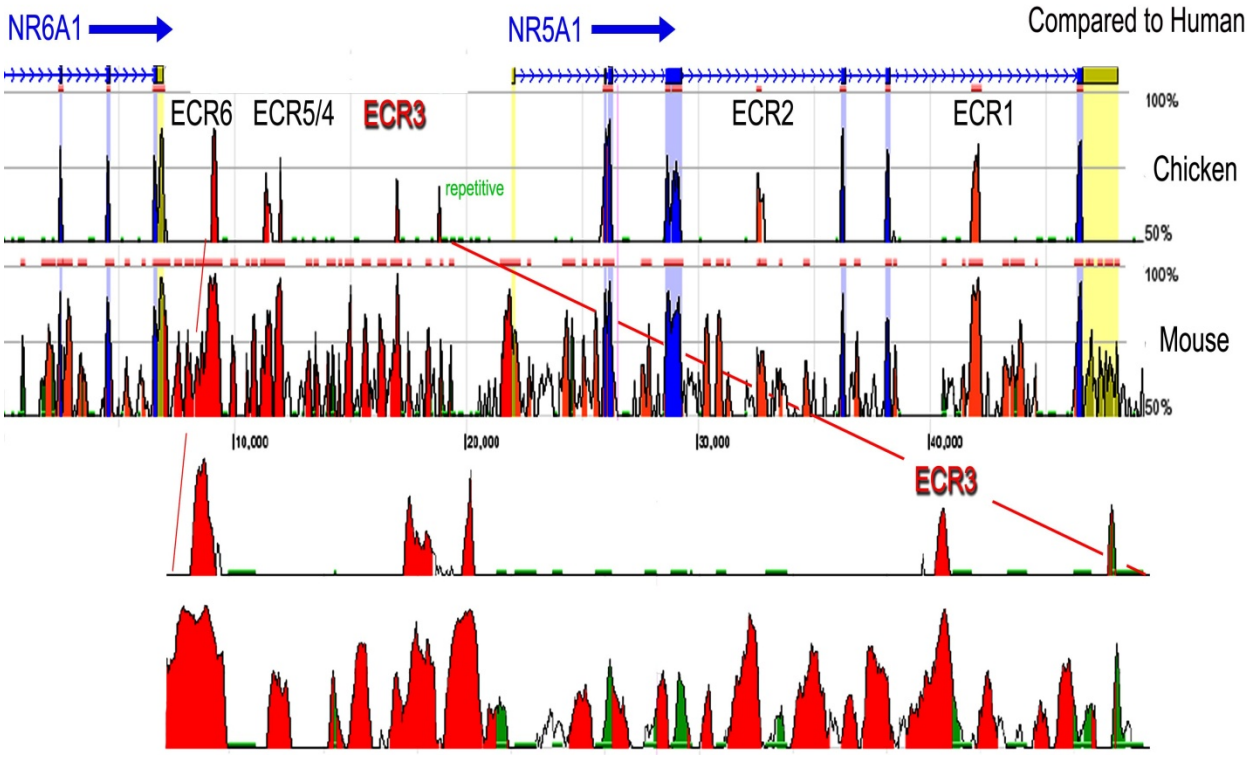


Figure 2: Schematic diagram of the vector SF-1(-734+60) Luc and the transcriptional activity of each ECR**.** Schematic diagram of the vector SF-1(-734+60) Luc and the transcriptional activity of each ECR shown as determined in multiple cell types (noted to the right of the graph) by transient transfection analysis. The bar graph shows promoter activity as the ratio of firefly/renilla luciferase activity of each construct relative to that of the SF-1 promoter without the ECR and error bars represent the SD of mean.

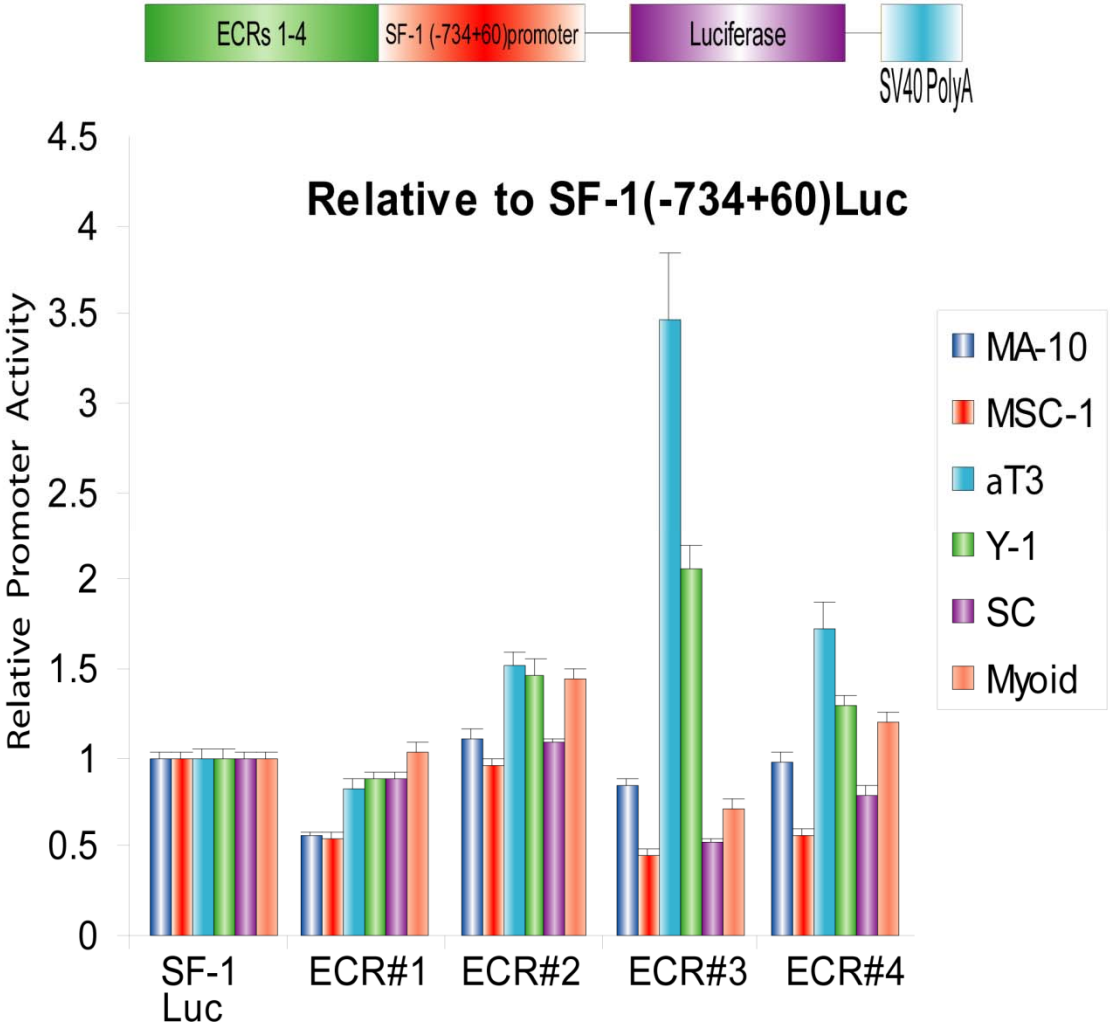
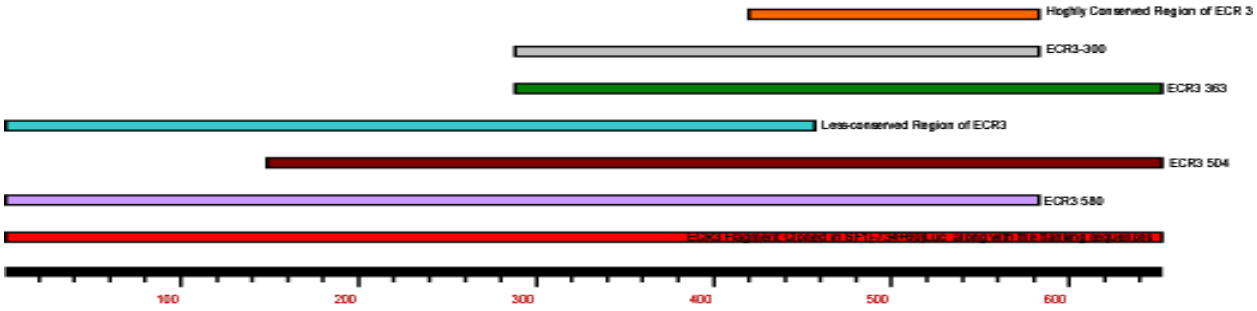


Figure 3: Deletion mutants and analysis of transcriptional activity. Deletion mutants were generated by cloning varying lengths of highly conserved sequences (shown on the top of the diagram) in SF-1(-734/+60)Luc and their transcriptional activities tested individually in alpha T3 cell type and graphed (bottom).



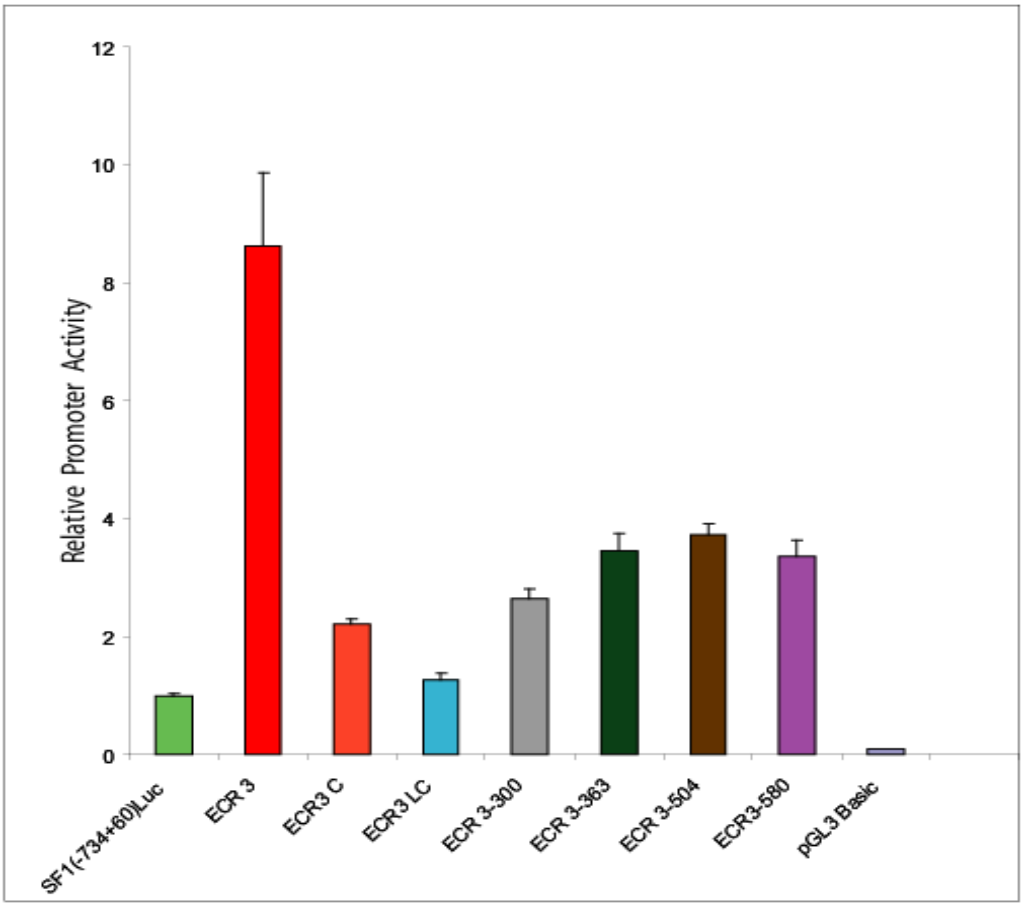


Figure 4: Dnase Footprinting gel of ECR3 region. Diagram showing the identified binding sites, hypersensitive site and footprints on the sequence of ECR3 (left and below). A picture of the footprinting gel is shown along with a hypersensitive site, footprints and the associated sequence ladder. +NE and -NE represent with and without nuclear extract respectively.

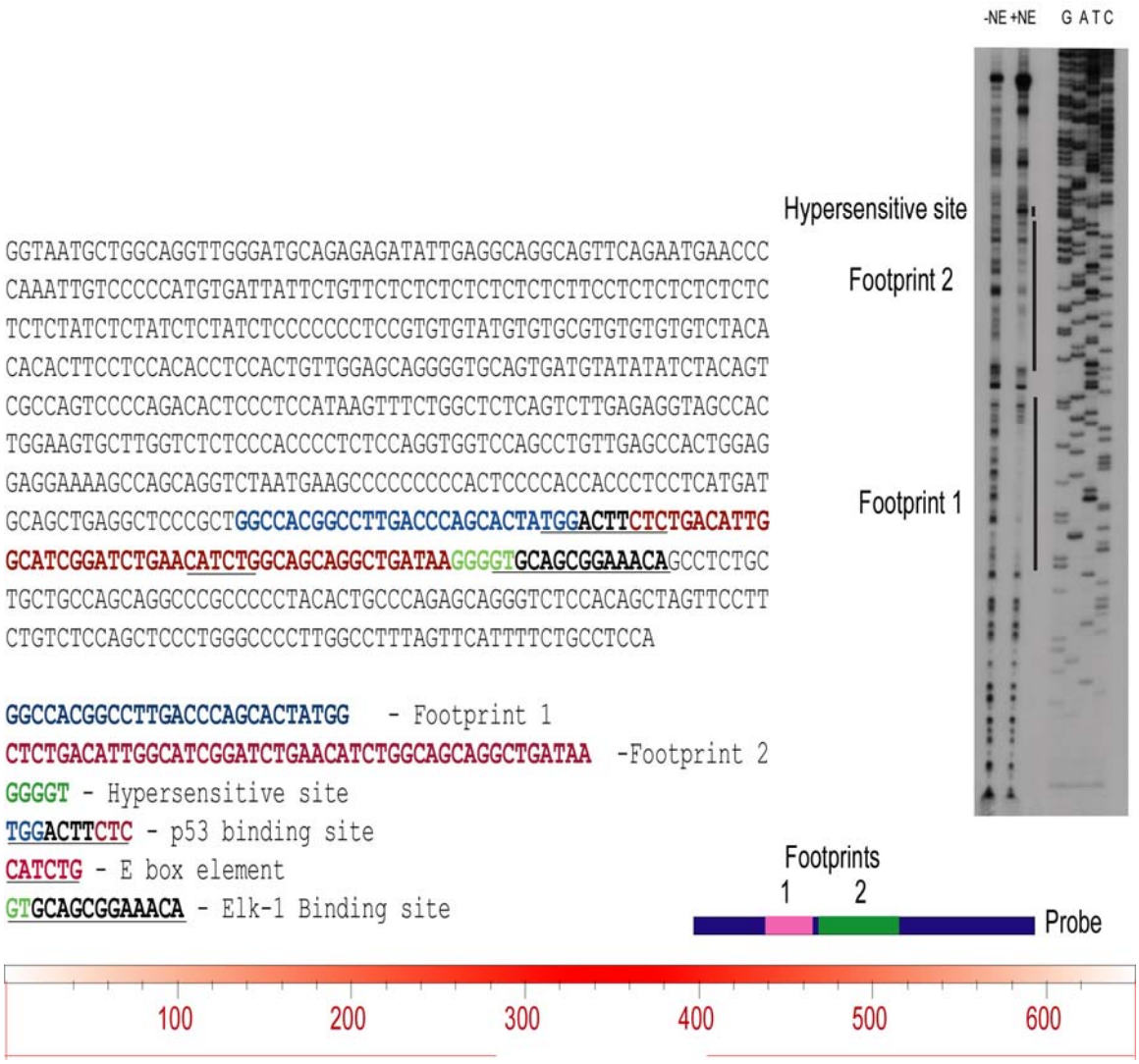


Figure 5: Transient transfection analysis to check activity of potential transcription factor binding sites. Schematic diagram of the relative positions of the generated mutations and potential transcription factor binding sites within the ECR3 fragment (top). The activities of E-box, p53 and Elk 1 mutants were tested by transient transfection analysis and compared to that of the base vector (bottom).

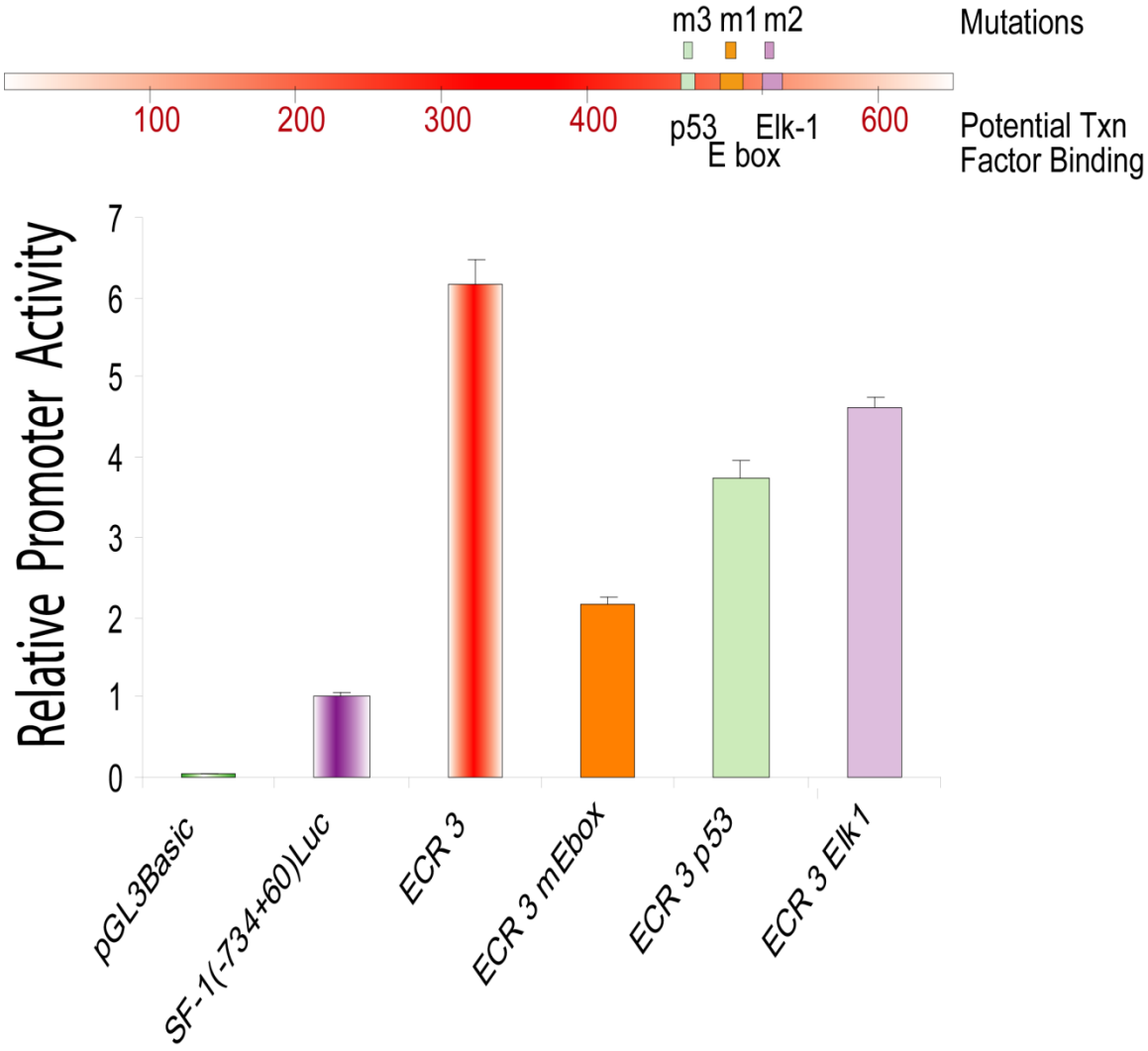


Figure 6: Electrophoretic mobility shift assay (EMSA) of ECR3 region. Electrophoretic mobility shift assay (EMSA) of ECR3 was performed using αT3 nuclear extracts and a radiolabeled probe containing the ECR3 E box (bottom). Where indicated competitors were added to the reaction at a concentration 100X that of probe. Similarly, indicated lanes contain antibodies to basic-helix-loop-helix (bHLH) proteins (E2A, USF1 and USF2), transcription factors that bind to E box elements. The results are summarized at the top with a schematic diagram indicating E box binding of E2A and its binding partner, USF1 and USF2.

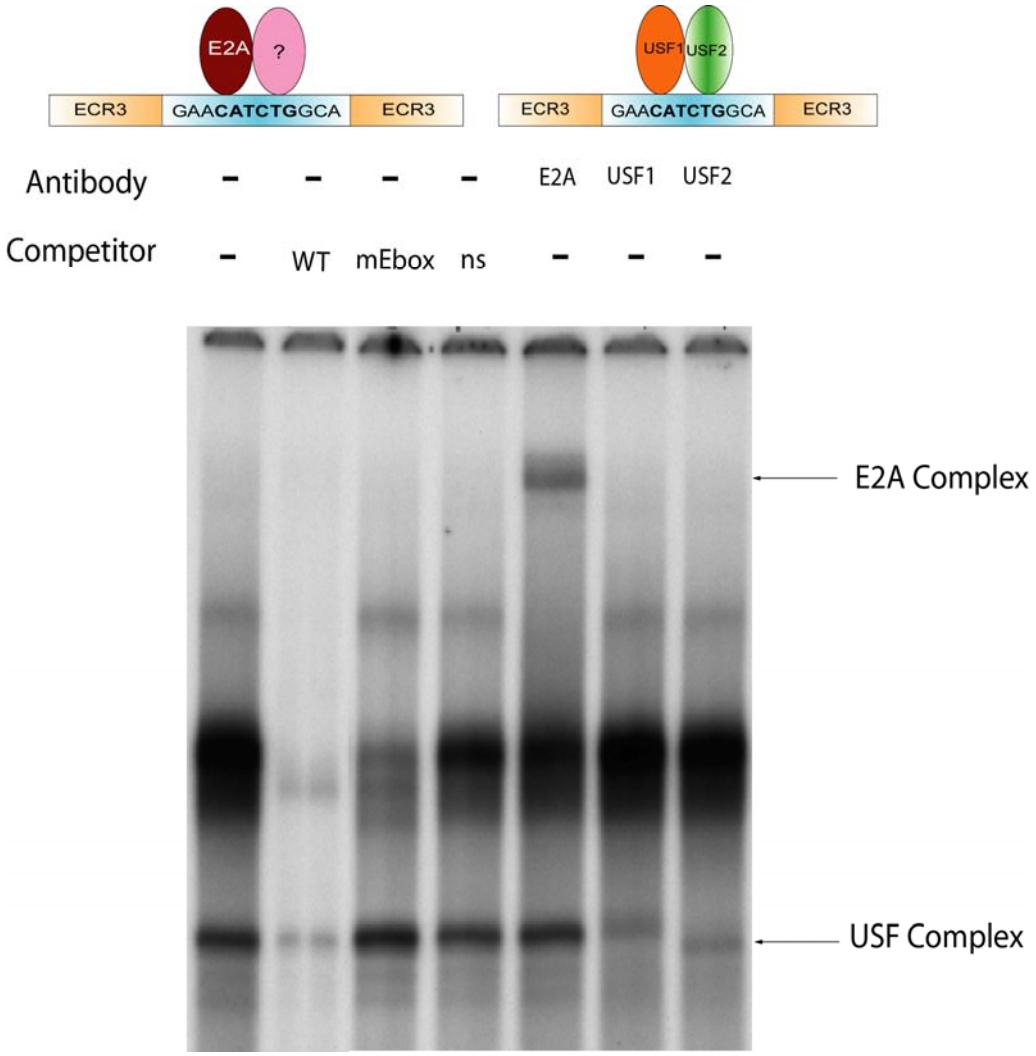


Figure 7: Evaluation of transcription factor binding by ChIP in αT3 cells. Interactions between E2A, USF1, USF2, RNA polymerase II, p300, H3Ac, H3K27me3 and ECR3 , *in vivo*, were evaluated using ChIP with cross linked chromatin from αT3 cells. Immunoprecipitations for each antibody were compared with normal rabbit IgG.

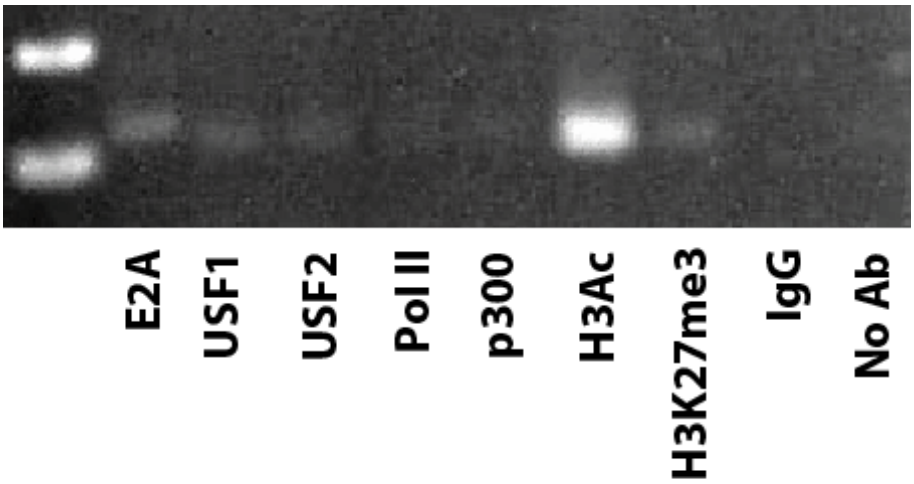
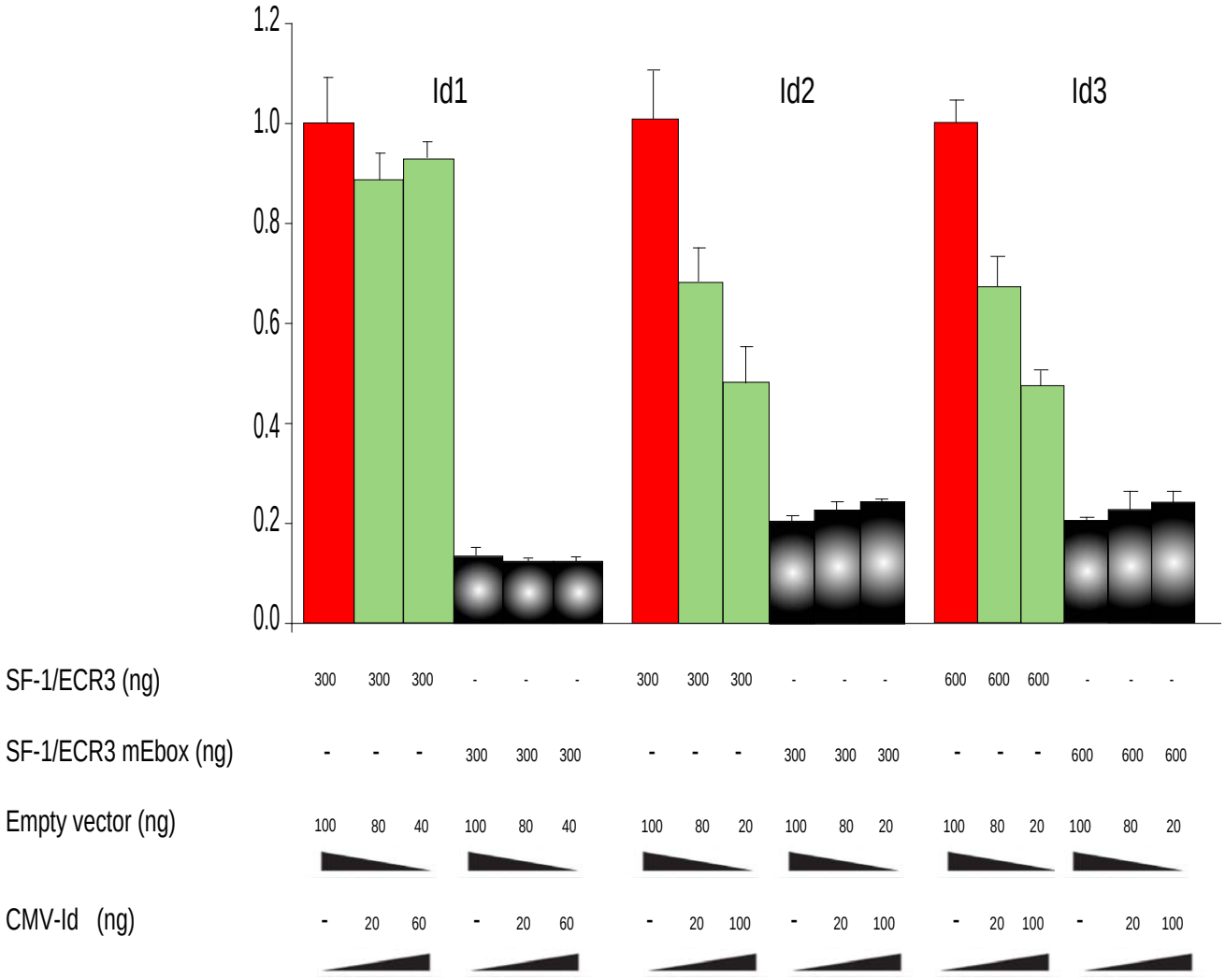


Figure 8: Co-transfection analysis to determine role of Id proteins. Negative regulators of bHLH proteins Id1, Id2, and Id3 were co transfected into αT3 cells with luciferase reporters driven by wild type (Trompouki, Bowman et al.) or E box mutant (mEbox) SF-1(-734+60)-ECR3. Relative activities were assessed and are shown below.



Chapter 6: Implications and Future Directions

Implications

This thesis delves in the transcriptional regulation of *Fshr* and *SF-1* that form an important component for the proper functioning of the HPA/HPG axis. It is now clear that regions beyond the promoter and its proximal elements are required for proper gene transcription. To identify these distal regulatory elements, sequence conservation was employed to identify genomic sequences that have remained unchanged over millions of years of evolution, and hence indicate that these regions are under evolutionary pressure to remain constant and harbor transcriptional factor binding regions. Chapter 2 of this thesis identifies highly conserved regions in the *Fshr* locus and potential binding sites for CTCF, a protein known to regulate transcription and chromatin architecture, and insinuates to a role of CTCF in transcriptional regulation of the *Fshr* gene. Chapter 4 and Chapter 5 of this thesis deals with conserved elements identified at SF-1 locus. A long non-coding RNA labeled *Fast* was found transcribed in opposite orientation of SF-1 and displayed similar tissue expression and regulatory profile as of SF-1. Knockdown and over-expression of the *Fast* transcript did not modulate mRNA or protein level of SF-1, indicating that the very act of transcription of the non-coding RNA might contribute to the opening of the chromatin for transcription of SF-1. Comparative sequence analysis identified an evolutionary conserved non-coding region, ECR3, approximately 4kb from the transcriptional start site of SF-1. ECR3 were cloned upstream of a -734/+60 SF-1 promoter directing expression of a luciferase reporter and its transcriptional activity assessed in various cell types by transient transfection analysis. Transient transfection analysis identified an ECR in the SF-1 locus, which regulates expression of SF-1 in the pituitary. Further characterization of the ECR region identified an E-box and binding of USF-1 and USF-2, along with E2A transcription factor,

confirmed by EMSA and ChIP. In addition, co-transfection studies identified Inhibitors of DNA binding, Id2 and Id3 that interfere with bHLH proteins to downregulate SF-1 transcription.

These studies highlight the importance of conserved elements and sequence conservation as a tool to identify regulatory elements distal to the promoter and lay the groundwork for further characterization of these identified regions and the role they might play in modulating gene transcription.

Future directions

Defining CTCF regions and its role in regulating *Fshr* transcription

Initiation of transcription starts from the promoter and involves recruitment of general transcriptional factors (Lee and Young 2000, Butler and Kadonaga 2002). Since the characterization of *Fshr*, twenty years ago, we have come a long way in identifying transcriptional mechanisms that regulate gene expression (Heckert, Daley et al. 1992, Gromoll, Pekel et al. 1996, George, Dille et al. 2011). The extensive characterization of the *Fshr* gene, while underlining the importance of the promoter and its proximal regions, fall short in revealing cell-specific expression of the receptor. Multiple transgenic reporter studies expressing the promoter and large swathe of the surrounding region were unable to replicate the cell-specific expression, supporting the theory that distal regulatory elements are required for proper spatio-temporal expression (Heckert, Sawadogo et al. 2000, Nordhoff, Gromoll et al. 2003, Hermann, Hornbaker et al. 2007). However, identification of distal regulatory elements possess great challenge as they can act over large distances and the *Fshr* gene is located in a long genomic

stretch devoid of protein-coding sequences and no known biological functions (Ovcharenko, Loots et al. 2005, Akalin, Fredman et al. 2009).

To facilitate the search of these regulatory elements comparative genomics and web based prediction tools were employed as described in Chapter 2. Among the varied distal regulatory elements known to regulate gene transcription, we looked at the ubiquitously expressed "multivalent" protein CTCF, with known multiple functions, including transcription (Filippova, Fagerlie et al. 1996, Phillips and Corces 2009). Evaluation of these evolutionary conserved regions (ECR) in the CTCF prediction software led to the identification of six ECRs; ECR1, ECR1d, ECR1f, ECR2 and ECR15, with high probability of CTCF binding. Depletion of CTCF in rat granulosa cells increased transcription of *Fshr* two-fold while the *Lhcgr* remained unchanged; indicating that depletion of CTCF removed its repressor effect on *Fshr* transcription. While these findings implicate CTCF in regulating *Fshr* transcription, a series of experiments will have to be performed to evaluate the location and the mechanism involved.

A good starting point will be to identify if CTCF binds to the ECR regions using electrophoretic mobility shift assay (EMSA) and further confirmation by chromatin immunoprecipitation (ChIP). ChiP-chip, a technique that combines ChIP with microarray technology (chip), can be used to identify a number of important DNA-protein interactions leading to important discoveries on transcriptional regulation (Ren, Robert et al. 2000). This technology will help identify CTCF binding regions on the *Fshr* locus.

Colony assay is an established assay to determine enhancer-blocking activity of insulator elements (Gombert, Farris et al. 2003). This assay employs a bacterial neomycin resistance gene (neo) reporter construct, stably integrated into the genomic site and measures the ability of the proposed insulator element to repress or stimulate transcription of selectable marker construct.

As shown in Figure 1, the insulator element will be placed between the enhancer and the promoter region or upstream of the enhancer. The number of colonies obtained after transfection into mouse transformed granulosa cells, GRMO2, will be proportional to the number of cells expressing the neomycin resistance gene (Neo^r) at levels sufficient to confer resistance to the drug (G418). CMV promoter and the enhancer run the Neo^r transcription. In absence of the enhancer, there will be reduced number of colonies. To test the ability of the construct to measure enhancer blocking activity, Drosophila scs and scs' boundary elements will be introduced at E-scs-P-neo-scs' and also as scs-E-P-Neo-scs'. Insertion of the scs should insulate the promoter activity and reduce the number of colonies. However, when the scs will be inserted upstream of the enhancer, the colony number will remain high. Insertion of the scs element will increase the distance between the enhancer and the promoter and this might be the cause of lower number of colonies.

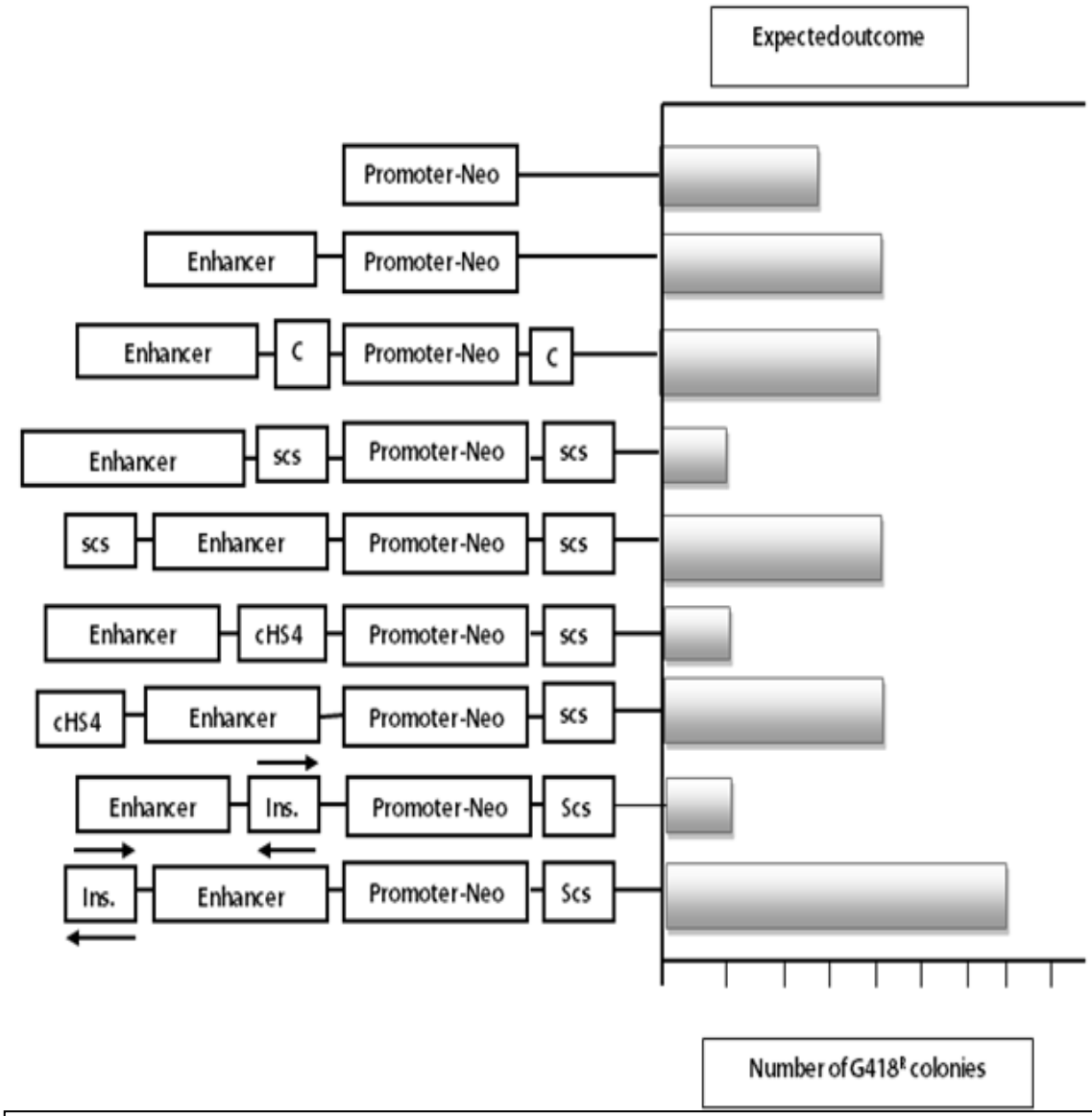


Figure 1: A schematic diagram depicting colony assay. As shown, the CMV promoter and neomycin resistance gene (neo) reporter construct denotes control and bar graph denotes expected trend for each reporter construct. Enhancer element placed upstream of the Promoter-Neo construct should yield in increased G418 resistant mouse GRMO2 cell colonies. C; control should not affect the colony number. *Drosophila* scs and scs' boundary elements will affect colonies when placed between the enhancer and the Promoter-Neo construct and will be another positive control. cHS4 a known insulator, when placed between the enhancer and promoter should have a similar affect as seen with *Drosophila* scs. This insulator effect shall be relieved when cHS4 will be placed upstream of the enhancer. To check if the regions identified function as insulators, they will be placed in both orientations and both upstream and in between the enhancer and promoter-Neo construct to check if they block enhancer activity, which will be measured by the number of resistant colonies

ascertain the enhancer blocking property of the insulator when placed between the CMV enhancer and the promoter. As a control, 1.2-kb chicken β-globin HS4 insulator sequence will be placed between the enhancer and the promoter. As HS4 is a previously reported enhancer-blocking element, it will reduce the number of colonies (Chung, Bell et al. 1997).

Chromosome conformation capture (3C) analysis is another method to evaluate the interaction between distal regulatory *cis*-acting elements and the promoter (Dekker, Rippe et al. 2002). In 3C, transient chromatin interactions are stabilized by formaldehyde cross-linking, followed by restriction digestion and intramolecular ligation and subsequent analysis of the ligation products by PCR. 3C has been used successfully to analyze spatial organization of small genomic domains such as mouse interferon gamma gene domain to larger genomic areas as in the case of mammalian alpha and beta globin gene domains (Palstra, Tolhuis et al. 2003, Eivazova and Aune 2004, Zhou, Xin et al. 2006). These techniques can identify CTCF binding regions on the *Fshr* locus, determine if there is a physical interaction between these distal regulatory elements and the *Fshr* promoter, and identify its role in modulating transcriptional regulation of *Fshr*.

Identifying functional role of *Fast*

The long noncoding RNA, *Fast*, identified at the *Nr5a1* locus did not affect transcription or translation of SF-1 indicating that it could function by acting in *trans* as in the case of HOX antisense intergenic RNA (*HOTAIR*). *HOTAIR* does not affect at the site of its expression at the HOX C locus, but interacts with PRC2 and the (LSD1)–CoReST–ReST complex and represses HOXD and several other loci (Rinn, Kertesz et al. 2007, Tsai, Manor et al. 2010). To identify targets of *Fast*, mU6 ShRNA and *Fast* shRNA plasmids will be transfected into MA-10 Leydig cell line, selected for neomycin resistance, expanded, and RNA extracted using Trizol. RT-PCR will be utilized to analyze the efficacy of the RNAi construct, and constructs that decrease

expression of *Fast* by 80-90% will be used for further experimentation. Six samples will be used, three for *Fast* knockdown and three controls and expression profiling can be performed using Affymetrix Gene Chip Mouse Genome 430 2.0 platform. Ingenuity Pathway Analysis (IPA) can then be utilized to analyze the raw data and promising gene candidates can be identified on basis of minimum two-fold difference and p-value less than 0.05. Once identified, these candidate genes can be confirmed by utilizing RT-PCR.

Fast shares similar expression and regulatory functions as SF-1 indicating a functional role of the non-coding RNA, which it might relay through interaction with cellular proteins. It has been frequently observed that a number of lncRNA recruit or interact with protein complexes to bring about their associated function. The lncRNA, *RepA* recruits the Polycomb Repressive Complex 2 (PRC2) to the future inactivated X chromosome resulting in trimethylation of Histone 3 lysine27 (H3K27me3)(Zhao, Sun et al. 2008). Similarly, lncRNA *Air* and *Kcnqlot1* interact with H3K9 histone methyltransferases G9a to epigenetically silence transcription. Identification of the proteins that interact with the non-coding RNA will help elucidate its functional significance and any potential mechanism involved in regulation (Nagano, Mitchell et al. 2008, Pandey, Mondal et al. 2008).

In order to identify proteins interacting with *Fast*, a biochemical approach can be taken as described previously for purification of proteins interacting with non-coding RNA, NRON (Willingham, Orth et al. 2005). Full length *Fast* will be cloned into pGEM5z and *in vitro* transcribed, following which a 3-hairpin RNA (MS2 loop) epitope bound to a MS2-MBP will be added to the 3' termini of the full length *Fast* RNA (Figure 2). MS2 is a bacteriophage coat protein that is bound to the maltose binding protein (MBP). MS2-MBP, *Fast* and a nonspecific (NS) RNA without the hairpin loop

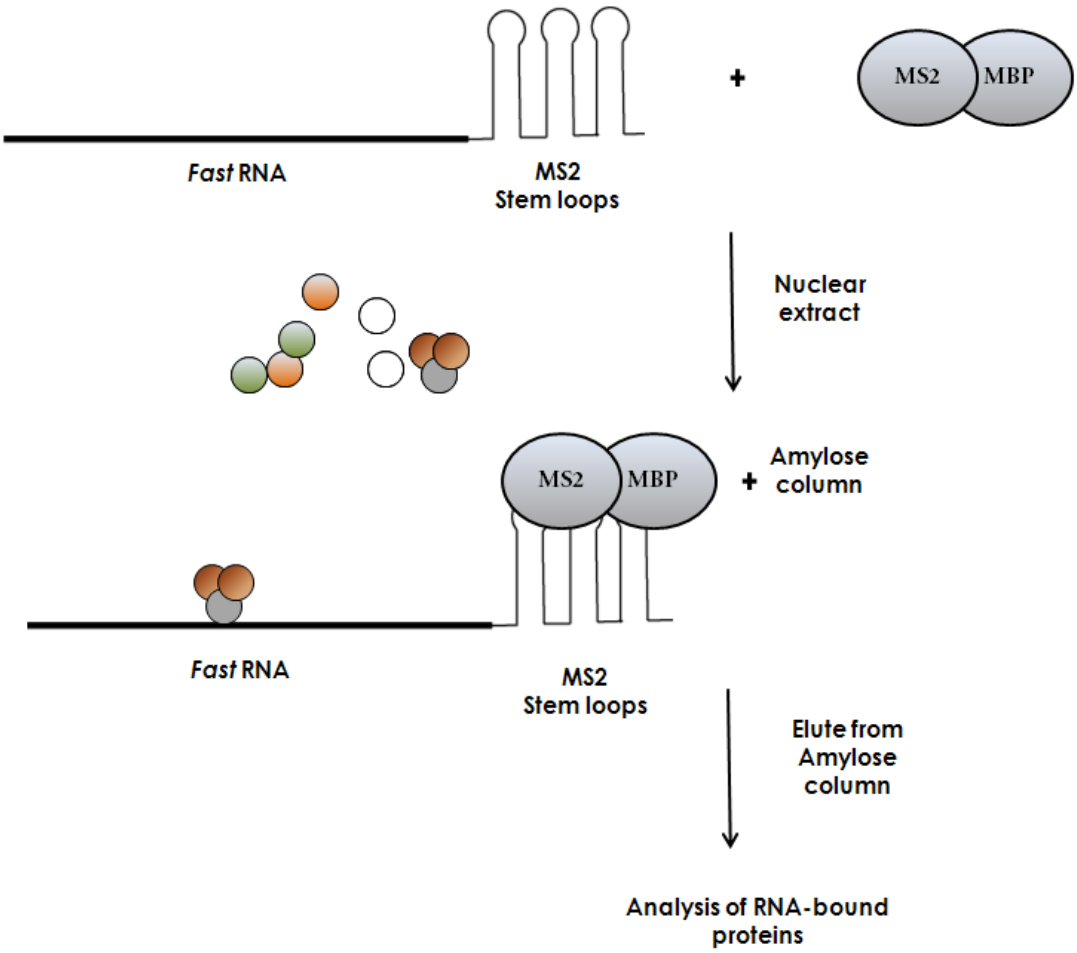


Figure 2: Purification of RNA bound proteins by MS2-MBP pull down method. The *Fast* RNA will be tagged with the three MS2 stem loops and then bound with the MS2:MBP fusion protein. Nuclear extracts are now incubated with the *Fast* RNA-MS2 stem loop and MS2:MBP fusion will be then affinity selected by binding to amylose resin and eluted.

will be prebound and incubated with protein extract from mice granulosa cells (*Fast* expressing) and NIH3T3 cells (*Fast* non-expressing). Samples will then be washed and eluted in an amylose binding resin. The resulting proteins can then be concentrated, resolved, analyzed by Mass Spectrometry and peptide sequence identified.

Over the past decade, the role of non-coding RNA in gene transcription and regulation has been appreciated. The so called "dark matter" has been resolved to play important biological role in proper regulation of the cellular machinery and our appreciation grows as new functions are described. The identification of a long non-coding RNA transcribed from the same locus as SF-1, in the opposite direction, with similar expression profile and tissue-specific transcript variants indicates a functional role. The data generated will provide insight to the functional role of the non-coding RNA and the role it may play in adrenal and gonadal development and will reveal a new chapter both in the role of the non-coding RNA and of SF-1 function.

Further characterization of ECR3 and role of identified transcription factors

Comparative genome sequencing between chicken, mouse, and human led to the identification and characterization of ECR3 located upstream of SF-1. An E-box was identified and it was found to be bound by upstream stimulatory factors 1 and 2 (USF1 and USF2), and transcription factor E2A, a ubiquitous bHLH protein. bHLH proteins contain a basic DNA binding domain and two helices that are involved in interaction with other proteins during the formation of homo- or hetero-dimers and play important roles in various developmental processes, including sex determination and is widely believed to play a central role in transcriptional regulatory networks (Massari and Murre 2000, Kee 2009, Zheng, Wang et al. 2009). Although, E2A is expressed in pituitary, no studies account for the pituitary functions of E2A. Our current study

indicates that E2A plays a transcriptional regulatory role in pituitary gonadotropes (Roberts, Steenbergen et al. 1993) . The bHLH protein, E2A hetero-dimerizes with other tissue-specific bHLH proteins before binding to DNA to regulate transcription. This protein can be identified by employing *p*roteomics of *i*solated *ch*romatin segments (PICh), that allows isolation and identification of proteins bound at a chromatin locus (Dejardin and Kingston 2009). PICh involves fixing cells with formaldehyde, followed by solubilization of chromatin, hybridization of a specific probe, capture using magnetic beads, elution of hybrids and identifying associated proteins by Mass spectrometric analysis. Once the proteins are identified, they can be confirmed by standard ChIP. PICh has been used successfully to identify binding of orphan nuclear receptor COUP-TF2 at telomeres (Dejardin and Kingston 2009).

Earlier work in mouse and rat, identified an E box motif in the proximal promoter region of SF-1 and it was found to be bound by USF1 and USF2 (Harris and Mellon 1998, Daggett, Rice et al. 2000). In the current study, another E-box was identified in the distal regions upstream of SF-1. ECR3 lies ~4kb upstream of the SF-1 promoter region, indicating that distal looping is involved and the two E-box might work in concert to regulate SF-1 transcription in the gonadotrope pituitary. 3C can be employed, as described earlier, to evaluate the interaction between ECR3 and the SF-1 promoter.

This is the first time that E2A has been implicated to play a role in transcriptional regulation of SF-1 in the pituitary gonadotropes. To identify the role of E2A in the transcriptional regulation of SF-1 in the pituitary gonadotropes, α-T3 cells stably transfected with shRNA directed against E2A, can be analyzed to identify the function of E2A on SF-1 transcription.

In summary, our study identified a highly evolutionarily conserved region namely ECR3 and an E-box within ECR3 that is essential for pituitary expression of SF-1. Additional studies will identify more or all of the regulatory sequences that are important for pituitary expression of SF-1 and increase our understanding on SF-1 gene regulation in pituitary.

REFERENCES

(1999). "A unified nomenclature system for the nuclear receptor superfamily." Cell **97**(2): 161-163.

(2004). "Finishing the euchromatic sequence of the human genome." Nature **431**(7011): 931-945.

(2004). "Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution." Nature **432**(7018): 695-716.

Abou-Issa, H. and L. E. Reichert, Jr. (1977). "Solubilization and some characteristics of the follitropin receptor from calf testis." J Biol Chem **252**(12): 4166-4174.

Achermann, J. C., M. Ito, P. C. Hindmarsh and J. L. Jameson (1999). "A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal and adrenal failure in humans." Nat Genet **22**(2): 125-126.

Achermann, J. C. and J. L. Jameson (1999). "Fertility and infertility: genetic contributions from the hypothalamic-pituitary-gonadal axis." Mol Endocrinol **13**(6): 812-818.

Achermann, J. C., G. Ozisik, M. Ito, U. A. Orun, K. Harmanci, B. Gurakan and J. L. Jameson (2002). "Gonadal determination and adrenal development are regulated by the orphan nuclear receptor steroidogenic factor-1, in a dose-dependent manner." J Clin Endocrinol Metab **87**(4): 1829-1833.

Adlam, M. and G. Siu (2003). "Hierarchical interactions control CD4 gene expression during thymocyte development." Immunity **18**(2): 173-184.

Akalin, A., D. Fredman, E. Arner, X. Dong, J. C. Bryne, H. Suzuki, C. O. Daub, Y. Hayashizaki and B. Lenhard (2009). "Transcriptional features of genomic regulatory blocks." Genome Biol **10**(4): R38.

Alliston, T. N., A. C. Maiyar, P. Buse, G. L. Firestone and J. S. Richards (1997). "Follicle stimulating hormone-regulated expression of serum/glucocorticoid-inducible kinase in rat ovarian granulosa cells: a functional role for the Sp1 family in promoter activity." Mol Endocrinol **11**(13): 1934-1949.

Amoss, M., R. Burgus, R. Blackwell, W. Vale, R. Fellows and R. Guillemin (1971). "Purification, amino acid composition and N-terminus of the hypothalamic luteinizing hormone releasing factor (LRF) of ovine origin." Biochem Biophys Res Commun **44**(1): 205-210.

Apaja, P. M., J. T. Aatsinki, H. J. Rajaniemi and U. E. Petaja-Repo (2005). "Expression of the mature luteinizing hormone receptor in rodent urogenital and adrenal tissues is developmentally regulated at a posttranslational level." Endocrinology **146**(8): 3224-3232.

Aravin, A., D. Gaidatzis, S. Pfeffer, M. Lagos-Quintana, P. Landgraf, N. Iovino, P. Morris, M. J. Brownstein, S. Kuramochi-Miyagawa, T. Nakano, M. Chien, J. J. Russo, J. Ju, R. Sheridan, C. Sander, M. Zavolan and T. Tuschl (2006). "A novel class of small RNAs bind to MILI protein in mouse testes." Nature **442**(7099): 203-207.

Arnosti, D. N., S. Barolo, M. Levine and S. Small (1996). "The eve stripe 2 enhancer employs multiple modes of transcriptional synergy." Development **122**(1): 205-214.

Arnosti, D. N. and M. M. Kulkarni (2005). "Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards?" J Cell Biochem **94**(5): 890-898.

Aronow, B. J., R. N. Silbiger, M. R. Dusing, J. L. Stock, K. L. Yager, S. S. Potter, J. J. Hutton and D. A. Wiginton (1992). "Functional analysis of the human adenosine deaminase gene thymic regulatory region and its ability to generate position-independent transgene expression." Mol Cell Biol **12**(9): 4170-4185.
Ascoli, M. (1981). "Characterization of several clonal lines of cultured Leydig tumor cells: gonadotropin receptors and steroidogenic responses." Endocrinology **108**(1): 88-95.

Azzi, S., S. Rossignol, V. Steunou, T. Sas, N. Thibaud, F. Danton, M. Le Jule, C. Heinrichs, S. Cabrol, C. Gicquel, Y. Le Bouc and I. Netchine (2009). "Multilocus methylation analysis in a large cohort of 11p15-related foetal growth disorders (Russell Silver and Beckwith Wiedemann syndromes) reveals simultaneous loss of methylation at paternal and maternal imprinted loci." Hum Mol Genet **18**(24): 4724-4733.

Bank, A. (2006). "Regulation of human fetal hemoglobin: new players, new complexities." Blood **107**(2): 435-443.

Bao, L., M. Zhou and Y. Cui (2008). "CTCFBSDB: a CTCF-binding site database for characterization of vertebrate genomic insulators." Nucleic Acids Res **36**(Database issue): D83-87.

Barnea, E. and Y. Bergman (2000). "Synergy of SF1 and RAR in activation of Oct-3/4 promoter." J Biol Chem **275**(9): 6608-6619.

Barnhart, K. M. and P. L. Mellon (1994). "The orphan nuclear receptor, steroidogenic factor-1, regulates the glycoprotein hormone alpha-subunit gene in pituitary gonadotropes." Mol Endocrinol **8**(7): 878-885.

Barnhart, K. M. and P. L. Mellon (1994). "The orphan nuclear receptor, steroidogenic factor-1, regulates the glycoprotein hormone alpha-subunit gene in pituitary gonadotropes." Molecular endocrinology **8**(7): 878-885.

Barrett, T., D. B. Troup, S. E. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, I. F. Kim, A. Soboleva, M. Tomashevsky and R. Edgar (2007). "NCBI GEO: mining tens of millions of expression profiles--database and tools update." Nucleic Acids Res **35**(Database issue): D760-765.

Barrett, T., D. B. Troup, S. E. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, I. F. Kim, A. Soboleva, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, R. N. Muertter and R. Edgar (2009). "NCBI GEO: archive for high-throughput functional genomic data." Nucleic Acids Res **37**(Database issue): D885-890.

Bartel, B. (2005). "MicroRNAs directing siRNA biogenesis." Nat Struct Mol Biol **12**(7): 569-571.

Bartel, D. P. (2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." Cell **116**(2): 281-297.
Bartolomei, M. S., S. Zemel and S. M. Tilghman (1991). "Parental imprinting of the mouse H19 gene." Nature **351**(6322): 153-155.

Baumann, M., J. Pontiller and W. Ernst (2010). "Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview." Mol Biotechnol **45**(3): 241-247.
Bejerano, G., C. B. Lowe, N. Ahituv, B. King, A. Siepel, S. R. Salama, E. M. Rubin, W. J. Kent and D. Haussler (2006). "A distal enhancer and an ultraconserved exon are derived from a novel retroposon." Nature **441**(7089): 87-90.

Bejerano, G., M. Pheasant, I. Makunin, S. Stephen, W. J. Kent, J. S. Mattick and D. Haussler (2004). "Ultraconserved elements in the human genome." Science **304**(5675): 1321-1325.

Bell, A. C. and G. Felsenfeld (2000). "Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene." Nature **405**(6785): 482-485.

Bell, A. C., A. G. West and G. Felsenfeld (1999). "The protein CTCF is required for the enhancer blocking activity of vertebrate insulators." Cell **98**(3): 387-396.

Ben-Josef, E., S. Y. Yang, T. H. Ji, J. M. Bidart, S. V. Garde, D. P. Chopra, A. T. Porter and D. G. Tang (1999). "Hormone-refractory prostate cancer cells express functional follicle-stimulating hormone receptor (FSHR)." J Urol **161**(3): 970-976.

Benson, B., S. Sorrentino and J. S. Evans (1969). "Increase in serum FSH following unilateral ovariectomy in the rat." Endocrinology **84**(2): 369-374.

Bernard, D., K. V. Prasanth, V. Tripathi, S. Colasse, T. Nakamura, Z. Xuan, M. Q. Zhang, F. Sedel, L. Jourdren, F. Coulpier, A. Triller, D. L. Spector and A. Bessis (2010). "A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression." EMBO J **29**(18): 3082-3093.

Biason-Lauber, A. and E. J. Schoenle (2000). "Apparently normal ovarian differentiation in a prepubertal girl with transcriptionally inactive steroidogenic factor 1 (NR5A1/SF-1) and adrenocortical insufficiency." Am J Hum Genet **67**(6): 1563-1568.

Billi, A. C., A. F. Alessi, V. Khivansara, T. Han, M. Freeberg, S. Mitani and J. K. Kim (2012). "The Caenorhabditis elegans HEN1 ortholog, HENN-1, methylates and stabilizes select subclasses of germline small RNAs." PLoS Genet **8**(4): e1002617.

Birney, E., J. A. Stamatoyannopoulos, A. Dutta, R. Guigo, T. R. Gingeras, E. H. Margulies, Z. Weng, M. Snyder, E. T. Dermitzakis, R. E. Thurman, M. S. Kuehn, C. M. Taylor, S. Neph, C. M. Koch, S. Asthana, A. Malhotra, I. Adzhubei, J. A. Greenbaum, R. M. Andrews, P. Flicek, P. J. Boyle, H. Cao, N. P. Carter, G. K. Clelland, S. Davis, N. Day, P. Dhami, S. C. Dillon, M. O. Dorschner, H. Fiegler, P. G. Giresi, J. Goldy, M. Hawrylycz, A. Haydock, R. Humbert, K. D. James, B. E. Johnson, E. M. Johnson, T. T. Frum, E. R. Rosenzweig, N. Karnani, K. Lee, G. C. Lefebvre, P. A. Navas, F. Neri, S. C. Parker, P. J. Sabo, R. Sandstrom, A. Shafer, D. Vetrie, M. Weaver, S. Wilcox, M. Yu, F. S. Collins, J. Dekker, J. D. Lieb, T. D. Tullius, G. E. Crawford, S. Sunyaev, W. S. Noble, I. Dunham, F. Denoeud, A. Reymond, P. Kapranov, J. Rozowsky, D. Zheng, R. Castelo, A. Frankish, J. Harrow, S. Ghosh, A. Sandelin, I. L. Hofacker, R. Baertsch, D. Keefe, S. Dike, J. Cheng, H. A. Hirsch, E. A. Sekinger, J. Lagarde, J. F. Abril, A. Shahab, C. Flamm, C. Fried, J. Hackermuller, J. Hertel, M. Lindemeyer, K. Missal, A. Tanzer, S. Washietl, J. Korbel, O. Emanuelsson, J. S. Pedersen, N. Holroyd, R. Taylor, D. Swarbreck, N. Matthews, M. C. Dickson, D. J. Thomas, M. T. Weirauch, J. Gilbert, J. Drenkow, I. Bell, X. Zhao, K. G. Srinivasan, W. K. Sung, H. S. Ooi, K. P. Chiu, S. Foissac, T. Alioto, M. Brent, L. Pachter, M. L. Tress, A. Valencia, S. W. Choo, C. Y. Choo, C. Ucla, C. Manzano, C. Wyss, E. Cheung, T. G. Clark, J. B. Brown, M. Ganesh, S. Patel, H. Tammana, J. Chrast, C. N. Henrichsen, C. Kai, J. Kawai, U. Nagalakshmi, J. Wu, Z. Lian, J. Lian, P. Newburger, X. Zhang, P. Bickel, J. S. Mattick, P. Carninci, Y. Hayashizaki, S. Weissman, T. Hubbard, R. M. Myers, J. Rogers, P. F. Stadler, T. M. Lowe, C. L. Wei, Y. Ruan, K. Struhl, M. Gerstein, S. E. Antonarakis, Y. Fu, E. D. Green, U. Karaoz, A. Siepel, J. Taylor, L. A. Liefer, K. A. Wetterstrand, P. J. Good, E. A. Feingold, M. S. Guyer, G. M. Cooper, G. Asimenos, C. N. Dewey, M. Hou, S. Nikolaev, J. I. Montoya-Burgos, A. Loytynoja, S. Whelan, F. Pardi, T. Massingham, H. Huang, N. R. Zhang, I. Holmes, J. C. Mullikin, A. Ureta-Vidal, B. Paten, M. Seringhaus, D. Church, K. Rosenbloom, W. J. Kent, E. A. Stone, S. Batzoglou, N. Goldman, R. C. Hardison, D. Haussler, W. Miller, A. Sidow, N. D. Trinklein, Z.

D. Zhang, L. Barrera, R. Stuart, D. C. King, A. Ameur, S. Enroth, M. C. Bieda, J. Kim, A. A. Bhinge, N. Jiang, J. Liu, F. Yao, V. B. Vega, C. W. Lee, P. Ng, A. Yang, Z. Moqtaderi, Z. Zhu, X. Xu, S. Squazzo, M. J. Oberley, D. Inman, M. A. Singer, T. A. Richmond, K. J. Munn, A. Rada-Iglesias, O. Wallerman, J. Komorowski, J. C. Fowler, P. Couttet, A. W. Bruce, O. M. Dovey, P. D. Ellis, C. F. Langford, D. A. Nix, G. Euskirchen, S. Hartman, A. E. Urban, P. Kraus, S. Van Calcar, N. Heintzman, T. H. Kim, K. Wang, C. Qu, G. Hon, R. Luna, C. K. Glass, M. G. Rosenfeld, S. F. Aldred, S. J. Cooper, A. Halees, J. M. Lin, H. P. Shulha, M. Xu, J. N. Haidar, Y. Yu, V. R. Iyer, R. D. Green, C. Wadelius, P. J. Farnham, B. Ren, R. A. Harte, A. S. Hinrichs, H. Trumbower, H. Clawson, J. Hillman-Jackson, A. S. Zweig, K. Smith, A. Thakkapallayil, G. Barber, R. M. Kuhn, D. Karolchik, L. Armengol, C. P. Bird, P. I. de Bakker, A. D. Kern, N. Lopez-Bigas, J. D. Martin, B. E. Stranger, A. Woodroffe, E. Davydov, A. Dimas, E. Eyras, I. B. Hallgrimsdottir, J. Huppert, M. C. Zody, G. R. Abecasis, X. Estivill, G. G. Bouffard, X. Guan, N. F. Hansen, J. R. Idol, V. V. Maduro, B. Maskeri, J. C. McDowell, M. Park, P. J. Thomas, A. C. Young, R. W. Blakesley, D. M. Muzny, E. Sodergren, D. A. Wheeler, K. C. Worley, H. Jiang, G. M. Weinstock, R. A. Gibbs, T. Graves, R. Fulton, E. R. Mardis, R. K. Wilson, M. Clamp, J. Cuff, S. Gnerre, D. B. Jaffe, J. L. Chang, K. Lindblad-Toh, E. S. Lander, M. Koriabine, M. Nefedov, K. Osoegawa, Y. Yoshinaga, B. Zhu and P. J. de Jong (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project." Nature **447**(7146): 799-816.

Bishop, C. E., D. J. Whitworth, Y. Qin, A. I. Agoulnik, I. U. Agoulnik, W. R. Harrison, R. R. Behringer and P. A. Overbeek (2000). "A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse." Nat Genet **26**(4): 490-494.

Blackwood, E. M. and J. T. Kadonaga (1998). "Going the distance: a current view of enhancer action." Science **281**(5373): 60-63.

Blow, M. J., D. J. McCulley, Z. Li, T. Zhang, J. A. Akiyama, A. Holt, I. Plajzer-Frick, M. Shoukry, C. Wright, F. Chen, V. Afzal, J. Bristow, B. Ren, B. L. Black, E. M. Rubin, A. Visel and L. A. Pennacchio (2010). "ChIP-Seq identification of weakly conserved heart enhancers." Nat Genet **42**(9): 806-810.

Bogerd, J. (2007). "Ligand-selective determinants in gonadotropin receptors." Mol Cell Endocrinol **260-262**: 144-152.

Boitani, C., M. Stefanini, A. Fragale and A. R. Morena (1995). "Activin stimulates Sertoli cell proliferation in a defined period of rat testis development." Endocrinology **136**(12): 5438-5444.

Bonfil, D., D. Chuderland, S. Kraus, D. Shahbazian, I. Friedberg, R. Seger and Z. Naor (2004). "Extracellular signal-regulated kinase, Jun N-terminal kinase, p38, and c-Src are involved in gonadotropin-releasing hormone-stimulated activity of the glycoprotein hormone follicle-stimulating hormone beta-subunit promoter." Endocrinology **145**(5): 2228-2244.

Brand, A. H., L. Breeden, J. Abraham, R. Sternglanz and K. Nasmyth (1985). "Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer." Cell **41**(1): 41-48.
Brannan, C. I., E. C. Dees, R. S. Ingram and S. M. Tilghman (1990). "The product of the H19 gene may function as an RNA." Mol Cell Biol **10**(1): 28-36.

Braun, T., P. R. Schofield and R. Sprengel (1991). "Amino-terminal leucine-rich repeats in gonadotropin receptors determine hormone selectivity." EMBO J **10**(7): 1885-1890.

Brennecke, J., A. A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam and G. J. Hannon (2007). "Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila." Cell **128**(6): 1089-1103.

Brennecke, J., C. D. Malone, A. A. Aravin, R. Sachidanandam, A. Stark and G. J. Hannon (2008). "An epigenetic role for maternally inherited piRNAs in transposon silencing." Science **322**(5906): 1387-1392.
Britten, R. J. and E. H. Davidson (1969). "Gene regulation for higher cells: a theory." Science **165**(3891): 349-357.

Brockdorff, N., A. Ashworth, G. F. Kay, V. M. McCabe, D. P. Norris, P. J. Cooper, S. Swift and S. Rastan (1992). "The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus." Cell **71**(3): 515-526.

Brown, C. J., B. D. Hendrich, J. L. Rupert, R. G. Lafreniere, Y. Xing, J. Lawrence and H. F. Willard (1992). "The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus." Cell **71**(3): 527-542.

Brune, M., C. Adams and J. Gromoll (2010). "Primate FSH-receptor promoter nucleotide sequence heterogeneity affects FSH-receptor transcription." Mol Cell Endocrinol **317**(1-2): 90-98.

Bulger, M. and M. Groudine (1999). "Looping versus linking: toward a model for long-distance gene activation." Genes Dev **13**(19): 2465-2477.

Buratowski, S. (2008). "Transcription. Gene expression--where to start?" Science **322**(5909): 1804-1805.
Burger, L. L., A. C. Dalkin, K. W. Aylor, L. J. Workman, D. J. Haisenleder and J. C. Marshall (2001). "Regulation of gonadotropin subunit transcription after ovariectomy in the rat: measurement of subunit primary transcripts reveals differential roles of GnRH and inhibin." Endocrinology **142**(8): 3435-3442.

Burger, L. L., D. J. Haisenleder, A. C. Dalkin and J. C. Marshall (2004). "Regulation of gonadotropin subunit gene transcription." J Mol Endocrinol **33**(3): 559-584.

Butler, J. E. and J. T. Kadonaga (2002). "The RNA polymerase II core promoter: a key component in the regulation of gene expression." Genes Dev **16**(20): 2583-2592.

Cabili, M. N., C. Trapnell, L. Goff, M. Koziol, B. Tazon-Vega, A. Regev and J. L. Rinn (2011). "Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses." Genes Dev **25**(18): 1915-1927.

Camp, T. A., J. O. Rahal and K. E. Mayo (1991). "Cellular-Localization and Hormonal-Regulation of Follicle-Stimulating-Hormone and Luteinizing-Hormone Receptor Messenger-Rnas in the Rat Ovary." Molecular Endocrinology **5**(10): 1405-1417.

Cannell, I. G., Y. W. Kong and M. Bushell (2008). "How do microRNAs regulate gene expression?" Biochem Soc Trans **36**(Pt 6): 1224-1231.

Cao, Y., Z. Yao, D. Sarkar, M. Lawrence, G. J. Sanchez, M. H. Parker, K. L. MacQuarrie, J. Davison, M. T. Morgan, W. L. Ruzzo, R. C. Gentleman and S. J. Tapscott (2010). "Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming." Dev Cell **18**(4): 662-674.

Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, N. Maeda, R. Oyama, T. Ravasi, B. Lenhard, C. Wells, R. Kodzius, K. Shimokawa, V. B. Bajic, S. E. Brenner, S. Batalov, A. R. Forrest, M.

Zavolan, M. J. Davis, L. G. Wilming, V. Aidinis, J. E. Allen, A. Ambesi-Impiombato, R. Apweiler, R. N. Aturaliya, T. L. Bailey, M. Bansal, L. Baxter, K. W. Beisel, T. Bersano, H. Bono, A. M. Chalk, K. P. Chiu, V. Choudhary, A. Christoffels, D. R. Clutterbuck, M. L. Crowe, E. Dalla, B. P. Dalrymple, B. de Bono, G. Della Gatta, D. di Bernardo, T. Down, P. Engstrom, M. Fagiolini, G. Faulkner, C. F. Fletcher, T. Fukushima, M. Furuno, S. Futaki, M. Gariboldi, P. Georgii-Hemming, T. R. Gingeras, T. Gojobori, R. E. Green, S. Gustincich, M. Harbers, Y. Hayashi, T. K. Hensch, N. Hirokawa, D. Hill, L. Huminiecki, M. Iacono, K. Ikeo, A. Iwama, T. Ishikawa, M. Jakt, A. Kanapin, M. Katoh, Y. Kawasawa, J. Kelso, H. Kitamura, H. Kitano, G. Kollias, S. P. Krishnan, A. Kruger, S. K. Kummerfeld, I. V. Kurochkin, L. F. Lareau, D. Lazarevic, L. Lipovich, J. Liu, S. Liuni, S. McWilliam, M. Madan Babu, M. Madera, L. Marchionni, H. Matsuda, S. Matsuzawa, H. Miki, F. Mignone, S. Miyake, K. Morris, S. Mottagui-Tabar, N. Mulder, N. Nakano, H. Nakauchi, P. Ng, R. Nilsson, S. Nishiguchi, S. Nishikawa, F. Nori, O. Ohara, Y. Okazaki, V. Orlando, K. C. Pang, W. J. Pavan, G. Pavesi, G. Pesole, N. Petrovsky, S. Piazza, J. Reed, J. F. Reid, B. Z. Ring, M. Ringwald, B. Rost, Y. Ruan, S. L. Salzberg, A. Sandelin, C. Schneider, C. Schonbach, K. Sekiguchi, C. A. Semple, S. Seno, L. Sessa, Y. Sheng, Y. Shibata, H. Shimada, K. Shimada, D. Silva, B. Sinclair, S. Sperling, E. Stupka, K. Sugiura, R. Sultana, Y. Takenaka, K. Taki, K. Tammoja, S. L. Tan, S. Tang, M. S. Taylor, J. Tegner, S. A. Teichmann, H. R. Ueda, E. van Nimwegen, R. Verardo, C. L. Wei, K. Yagi, H. Yamanishi, E. Zabarovsky, S. Zhu, A. Zimmer, W. Hide, C. Bult, S. M. Grimmond, R. D. Teasdale, E. T. Liu, V. Brusic, J. Quackenbush, C. Wahlestedt, J. S. Mattick, D. A. Hume, C. Kai, D. Sasaki, Y. Tomaru, S. Fukuda, M. Kanamori-Katayama, M. Suzuki, J. Aoki, T. Arakawa, J. Iida, K. Imamura, M. Itoh, T. Kato, H. Kawaji, N. Kawagashira, T. Kawashima, M. Kojima, S. Kondo, H. Konno, K. Nakano, N. Ninomiya, T. Nishio, M. Okada, C. Plessy, K. Shibata, T. Shiraki, S. Suzuki, M. Tagami, K. Waki, A. Watahiki, Y. Okamura-Oho, H. Suzuki, J. Kawai and Y. Hayashizaki (2005). "The transcriptional landscape of the mammalian genome." Science **309**(5740): 1559-1563.

Carthew, R. W. and E. J. Sontheimer (2009). "Origins and Mechanisms of miRNAs and siRNAs." Cell **136**(4): 642-655.

Cawley, S., S. Bekiranov, H. H. Ng, P. Kapranov, E. A. Sekinger, D. Kampa, A. Piccolboni, V. Sementchenko, J. Cheng, A. J. Williams, R. Wheeler, B. Wong, J. Drenkow, M. Yamanaka, S. Patel, S. Brubaker, H. Tammana, G. Helt, K. Struhl and T. R. Gingeras (2004). "Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs." Cell **116**(4): 499-509.

Chapman, M. A., I. J. Donaldson, J. Gilbert, D. Grafham, J. Rogers, A. R. Green and B. Gottgens (2004). "Analysis of multiple genomic sequence alignments: a web resource, online tools, and lessons learned from analysis of mammalian SCL loci." Genome Res **14**(2): 313-318.

Chauvigne, F., A. Tingaud-Sequeira, M. Agulleiro, M. Calusinska, A. Gomez, R. N. Finn and J. Cerda (2010). "Functional and Evolutionary Analysis of Flatfish Gonadotropin Receptors Reveals Cladal- and Lineage-Level Divergence of the Teleost Glycoprotein Receptor Family." Biol Reprod.

Chen, H. P., A. Lin, J. S. Bloom, A. H. Khan, C. C. Park and D. J. Smith (2008). "Screening reveals conserved and nonconserved transcriptional regulatory elements including an E3/E4 allele-dependent APOE coding region enhancer." Genomics **92**(5): 292-300.

Chen, J. K. and L. L. Heckert (2001). "Dmrt1 expression is regulated by follicle-stimulating hormone and phorbol esters in postnatal Sertoli cells." Endocrinology **142**(3): 1167-1178.

Chen, J. L. and C. W. Greider (2004). "Telomerase RNA structure and function: implications for dyskeratosis congenita." Trends Biochem Sci **29**(4): 183-192.

Chen, L. and J. Widom (2005). "Mechanism of transcriptional silencing in yeast." Cell **120**(1): 37-48.
Cheng, J., P. Kapranov, J. Drenkow, S. Dike, S. Brubaker, S. Patel, J. Long, D. Stern, H. Tammana, G. Helt, V. Sementchenko, A. Piccolboni, S. Bekiranov, D. K. Bailey, M. Ganesh, S. Ghosh, I. Bell, D. S. Gerhard and T. R. Gingeras (2005). "Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution." Science **308**(5725): 1149-1154.

Choi, J. H., K. C. Choi, N. Auersperg and P. C. Leung (2004). "Overexpression of follicle-stimulating hormone receptor activates oncogenic pathways in preneoplastic ovarian surface epithelial cells." J Clin Endocrinol Metab **89**(11): 5508-5516.

Chua, J. H., A. Armugam and K. Jeyaseelan (2009). "MicroRNAs: biogenesis, function and applications." Curr Opin Mol Ther **11**(2): 189-199.

Chung, J. H., A. C. Bell and G. Felsenfeld (1997). "Characterization of the chicken beta-globin insulator." Proc Natl Acad Sci U S A **94**(2): 575-580.
Chung, J. H., M. Whiteley and G. Felsenfeld (1993). "A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila." Cell **74**(3): 505-514.

Chureau, C., M. Prissette, A. Bourdet, V. Barbe, L. Cattolico, L. Jones, A. Eggen, P. Avner and L. Duret (2002). "Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine." Genome Res **12**(6): 894-908.

Ciccone, N. A., C. T. Lacza, M. Y. Hou, S. J. Gregory, K. Y. Kam, S. Xu and U. B. Kaiser (2008). "A composite element that binds basic helix loop helix and basic leucine zipper transcription factors is important for gonadotropin-releasing hormone regulation of the follicle-stimulating hormone beta gene." Mol Endocrinol **22**(8): 1908-1923.

Clark, M. B., P. P. Amaral, F. J. Schlesinger, M. E. Dinger, R. J. Taft, J. L. Rinn, C. P. Ponting, P. F. Stadler, K. V. Morris, A. Morillon, J. S. Rozowsky, M. B. Gerstein, C. Wahlestedt, Y. Hayashizaki, P. Carninci, T. R. Gingeras and J. S. Mattick (2011). "The reality of pervasive transcription." PLoS Biol **9**(7): e1000625; discussion e1001102.

Core, L. J. and J. T. Lis (2008). "Transcription regulation through promoter-proximal pausing of RNA polymerase II." Science **319**(5871): 1791-1792.

Coss, D., S. B. Jacobs, C. E. Bender and P. L. Mellon (2004). "A novel AP-1 site is critical for maximal induction of the follicle-stimulating hormone beta gene by gonadotropin-releasing hormone." J Biol Chem **279**(1): 152-162.

Crawford, G. E., S. Davis, P. C. Scacheri, G. Renaud, M. J. Halawi, M. R. Erdos, R. Green, P. S. Meltzer, T. G. Wolfsberg and F. S. Collins (2006). "DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays." Nat Methods **3**(7): 503-509.

Cullen, K. E., M. P. Kladde and M. A. Seyfred (1993). "Interaction between transcription regulatory regions of prolactin chromatin." Science **261**(5118): 203-206.

D'Haene, B., C. Attanasio, D. Beysen, J. Dostie, E. Lemire, P. Bouchard, M. Field, K. Jones, B. Lorenz, B. Menten, K. Buysse, F. Pattyn, M. Friedli, C. Ucla, C. Rossier, C. Wyss, F. Speleman, A. De Paepe, J. Dekker, S. E. Antonarakis and E. De Baere (2009). "Disease-causing 7.4 kb cis-regulatory deletion

disrupting conserved non-coding sequences and their interaction with the FOXL2 promotor: implications for mutation screening." PLoS Genet **5**(6): e1000522.

Daggett, M. A., D. A. Rice and L. L. Heckert (2000). "Expression of steroidogenic factor 1 in the testis requires an E box and CCAAT box in its promoter proximal region." Biol Reprod **62**(3): 670-679.

Dankbar, B., M. H. Brinkworth, S. Schlatt, G. F. Weinbauer, E. Nieschlag and J. Gromoll (1995). "Ubiquitous expression of the androgen receptor and testis-specific expression of the FSH receptor in the cynomolgus monkey (Macaca fascicularis) revealed by a ribonuclease protection assay." J Steroid Biochem Mol Biol **55**(1): 35-41.

de Kok, Y. J., G. F. Merkx, S. M. van der Maarel, I. Huber, S. Malcolm, H. H. Ropers and F. P. Cremers (1995). "A duplication/paracentric inversion associated with familial X-linked deafness (DFN3) suggests the presence of a regulatory element more than 400 kb upstream of the POU3F4 gene." Hum Mol Genet **4**(11): 2145-2150.

de la Calle-Mustienes, E., C. G. Feijoo, M. Manzanares, J. J. Tena, E. Rodriguez-Seguel, A. Letizia, M. L. Allende and J. L. Gomez-Skarmeta (2005). "A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate Iroquois cluster gene deserts." Genome Res **15**(8): 1061-1072.
de Laat, W., P. Klous, J. Kooren, D. Noordermeer, R. J. Palstra, M. Simonis, E. Splinter and F. Grosveld (2008). "Three-dimensional organization of gene expression in erythroid cells." Curr Top Dev Biol **82**: 117-139.

DeChiara, T. M., E. J. Robertson and A. Efstratiadis (1991). "Parental imprinting of the mouse insulin-like growth factor II gene." Cell **64**(4): 849-859.
Dejardin, J. and R. E. Kingston (2009). "Purification of proteins associated with specific genomic Loci." Cell **136**(1): 175-186.

Dekker, J., K. Rippe, M. Dekker and N. Kleckner (2002). "Capturing chromosome conformation." Science **295**(5558): 1306-1311.

Dermitzakis, E. T., A. Reymond and S. E. Antonarakis (2005). "Conserved non-genic sequences - an unexpected feature of mammalian genomes." Nat Rev Genet **6**(2): 151-157.

Dermitzakis, E. T., A. Reymond, R. Lyle, N. Scamuffa, C. Ucla, S. Deutsch, B. J. Stevenson, V. Flegel, P. Bucher, C. V. Jongeneel and S. E. Antonarakis (2002). "Numerous potentially functional but non-genic conserved sequences on human chromosome 21." Nature **420**(6915): 578-582.

Dias, J. A., B. D. Cohen, B. Lindau-Shepard, C. A. Nechamen, A. J. Peterson and A. Schmidt (2002). "Molecular, structural, and cellular biology of follitropin and follitropin receptor." Vitam Horm **64**: 249-322.

Dickinson, R. E., A. J. Stewart, M. Myers, R. P. Millar and W. C. Duncan (2009). "Differential expression and functional characterization of luteinizing hormone receptor splice variants in human luteal cells: implications for luteolysis." Endocrinology **150**(6): 2873-2881.

Dinger, M. E., K. C. Pang, T. R. Mercer and J. S. Mattick (2008). "Differentiating protein-coding and noncoding RNA: challenges and ambiguities." PLoS Comput Biol **4**(11): e1000176.

Dorn, C., Q. Ou, J. Svaren, P. A. Crawford and Y. Sadovsky (1999). "Activation of luteinizing hormone beta gene by gonadotropin-releasing hormone requires the synergy of early growth response-1 and steroidogenic factor-1." J Biol Chem **274**(20): 13870-13876.

Dorrington, J. H. and D. T. Armstrong (1979). "Effects of FSH on gonadal functions." Recent Prog Horm Res **35**: 301-342.

Driscoll, M. C., C. S. Dobkin and B. P. Alter (1989). "Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites." Proc Natl Acad Sci U S A **86**(19): 7470-7474.

Drummond, A. E., M. Dyson, J. E. Mercer and J. K. Findlay (1996). "Differential responses of post-natal rat ovarian cells to FSH and activin." Mol Cell Endocrinol **122**(1): 21-32.

Du, T. and P. D. Zamore (2005). "microPrimer: the biogenesis and function of microRNA." Development **132**(21): 4645-4652.

Dufau, M. L., C. H. Tsai-Morris, Z. Z. Hu and E. Buczko (1995). "Structure and regulation of the luteinizing hormone receptor gene." J Steroid Biochem Mol Biol **53**(1-6): 283-291.

Dunkel, L., J. L. Tilly, T. Shikone, K. Nishimori and A. J. Hsueh (1994). "Follicle-stimulating hormone receptor expression in the rat ovary: increases during prepubertal development and regulation by the opposing actions of transforming growth factors beta and alpha." Biol Reprod **50**(4): 940-948.

Ebisuya, M., T. Yamamoto, M. Nakajima and E. Nishida (2008). "Ripples from neighbouring transcription." Nat Cell Biol **10**(9): 1106-1113.

Eddy, S. R. (2001). "Non-coding RNA genes and the modern RNA world." Nat Rev Genet **2**(12): 919-929.

Edgar, R., M. Domrachev and A. E. Lash (2002). "Gene Expression Omnibus: NCBI gene expression and hybridization array data repository." Nucleic Acids Res **30**(1): 207-210.

Eivazova, E. R. and T. M. Aune (2004). "Dynamic alterations in the conformation of the Ifng gene region during T helper cell differentiation." Proc Natl Acad Sci U S A **101**(1): 251-256.

Engstrom, P. G., S. J. Ho Sui, O. Drivenes, T. S. Becker and B. Lenhard (2007). "Genomic regulatory blocks underlie extensive microsynteny conservation in insects." Genome Res **17**(12): 1898-1908.

Epstein, D. J. (2009). "Cis-regulatory mutations in human disease." Brief Funct Genomic Proteomic **8**(4): 310-316.

Esteller, M. (2011). "Non-coding RNAs in human disease." Nat Rev Genet **12**(12): 861-874.

Ferguson-Smith, A. C., B. M. Cattanach, S. C. Barton, C. V. Beechey and M. A. Surani (1991). "Embryological and molecular investigations of parental imprinting on mouse chromosome 7." Nature **351**(6328): 667-670.

Filippova, G. N., S. Fagerlie, E. M. Klenova, C. Myers, Y. Dehner, G. Goodwin, P. E. Neiman, S. J. Collins and V. V. Lobanenkov (1996). "An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes." Mol Cell Biol **16**(6): 2802-2813.

Findlay, J. K. and A. E. Drummond (1999). "Regulation of the FSH Receptor in the Ovary." Trends Endocrinol Metab **10**(5): 183-188.

Fletcher, P. W. and L. E. Reichert, Jr. (1984). "Cellular processing of follicle-stimulating hormone by Sertoli cells in serum-free culture." Mol Cell Endocrinol **34**(1): 39-49.

Flores, J. A., J. D. Veldhuis and D. A. Leong (1990). "Follicle-stimulating hormone evokes an increase in intracellular free calcium ion concentrations in single ovarian (granulosa) cells." Endocrinology **127**(6): 3172-3179.

Fraser, P. and W. Bickmore (2007). "Nuclear organization of the genome and the potential for gene regulation." Nature **447**(7143): 413-417.

Frazer, K. A., J. B. Sheehan, R. P. Stokowski, X. Chen, R. Hosseini, J. F. Cheng, S. P. Fodor, D. R. Cox and N. Patil (2001). "Evolutionarily conserved sequences on human chromosome 21." Genome Res **11**(10): 1651-1659.

Gardiner, J. R., Y. Shima, K. Morohashi and A. Swain (2012). "SF-1 expression during adrenal development and tumourigenesis." Mol Cell Endocrinol **351**(1): 12-18.

Gaudet, J. and S. E. Mango (2002). "Regulation of organogenesis by the Caenorhabditis elegans FoxA protein PHA-4." Science **295**(5556): 821-825.

Geng, Y., C. H. Tsai-Morris, Y. Zhang and M. L. Dufau (1999). "The human luteinizing hormone receptor gene promoter: activation by Sp1 and Sp3 and inhibitory regulation." Biochem Biophys Res Commun **263**(2): 366-371.

George, J. W., E. A. Dille and L. L. Heckert (2011). "Current concepts of follicle-stimulating hormone receptor gene regulation." Biol Reprod **84**(1): 7-17.

Gershenzon, N. I. and I. P. Ioshikhes (2005). "Synergy of human Pol II core promoter elements revealed by statistical sequence analysis." Bioinformatics **21**(8): 1295-1300.

Ghildiyal, M. and P. D. Zamore (2009). "Small silencing RNAs: an expanding universe." Nat Rev Genet **10**(2): 94-108.

Gibb, E. A., C. J. Brown and W. L. Lam (2011). "The functional role of long non-coding RNA in human carcinomas." Mol Cancer **10**: 38.

Gibbs, R. A., G. M. Weinstock, M. L. Metzker, D. M. Muzny, E. J. Sodergren, S. Scherer, G. Scott, D. Steffen, K. C. Worley, P. E. Burch, G. Okwuonu, S. Hines, L. Lewis, C. DeRamo, O. Delgado, S. Dugan-Rocha, G. Miner, M. Morgan, A. Hawes, R. Gill, Celera, R. A. Holt, M. D. Adams, P. G. Amanatides, H. Baden-Tillson, M. Barnstead, S. Chin, C. A. Evans, S. Ferriera, C. Fosler, A. Glodek, Z. Gu, D. Jennings, C. L. Kraft, T. Nguyen, C. M. Pfannkoch, C. Sitter, G. G. Sutton, J. C. Venter, T. Woodage, D. Smith, H. M. Lee, E. Gustafson, P. Cahill, A. Kana, L. Doucette-Stamm, K. Weinstock, K. Fechtel, R. B. Weiss, D. M. Dunn, E. D. Green, R. W. Blakesley, G. G. Bouffard, P. J. De Jong, K. Osoegawa, B. Zhu, M. Marra, J. Schein, I. Bosdet, C. Fjell, S. Jones, M. Krzywinski, C. Mathewson, A. Siddiqui, N. Wye, J. McPherson, S. Zhao, C. M. Fraser, J. Shetty, S. Shatsman, K. Geer, Y. Chen, S. Abramzon, W. C. Nierman, P. H. Havlak, R. Chen, K. J. Durbin, A. Egan, Y. Ren, X. Z. Song, B. Li, Y. Liu, X. Qin, S. Cawley, A. J. Cooney, L. M. D'Souza, K. Martin, J. Q. Wu, M. L. Gonzalez-Garay, A. R. Jackson, K. J. Kalafus, M. P. McLeod, A. Milosavljevic, D. Virk, A. Volkov, D. A. Wheeler, Z. Zhang, J. A. Bailey, E.

E. Eichler, E. Tuzun, E. Birney, E. Mongin, A. Ureta-Vidal, C. Woodwark, E. Zdobnov, P. Bork, M. Suyama, D. Torrents, M. Alexandersson, B. J. Trask, J. M. Young, H. Huang, H. Wang, H. Xing, S. Daniels, D. Gietzen, J. Schmidt, K. Stevens, U. Vitt, J. Wingrove, F. Camara, M. Mar Alba, J. F. Abril, R. Guigo, A. Smit, I. Dubchak, E. M. Rubin, O. Couronne, A. Poliakov, N. Hubner, D. Ganten, C. Goesele, O. Hummel, T. Kreitler, Y. A. Lee, J. Monti, H. Schulz, H. Zimdahl, H. Himmelbauer, H. Lehrach, H. J. Jacob, S. Bromberg, J. Gullings-Handley, M. I. Jensen-Seaman, A. E. Kwitek, J. Lazar, D. Pasko, P. J. Tonellato, S. Twigger, C. P. Ponting, J. M. Duarte, S. Rice, L. Goodstadt, S. A. Beatson, R. D. Emes, E. E. Winter, C. Webber, P. Brandt, G. Nyakatura, M. Adetobi, F. Chiaromonte, L. Elnitski, P. Eswara, R. C. Hardison, M. Hou, D. Kolbe, K. Makova, W. Miller, A. Nekrutenko, C. Riemer, S. Schwartz, J. Taylor, S. Yang, Y. Zhang, K. Lindpaintner, T. D. Andrews, M. Caccamo, M. Clamp, L. Clarke, V. Curwen, R. Durbin, E. Eyras, S. M. Searle, G. M. Cooper, S. Batzoglou, M. Brudno, A. Sidow, E. A. Stone, B. A. Payseur, G. Bourque, C. Lopez-Otin, X. S. Puente, K. Chakrabarti, S. Chatterji, C. Dewey, L. Pachter, N. Bray, V. B. Yap, A. Caspi, G. Tesler, P. A. Pevzner, D. Haussler, K. M. Roskin, R. Baertsch, H. Clawson, T. S. Furey, A. S. Hinrichs, D. Karolchik, W. J. Kent, K. R. Rosenbloom, H. Trumbower, M. Weirauch, D. N. Cooper, P. D. Stenson, B. Ma, M. Brent, M. Arumugam, D. Shteynberg, R. R. Copley, M. S. Taylor, H. Riethman, U. Mudunuri, J. Peterson, M. Guyer, A. Felsenfeld, S. Old, S. Mockrin and F. Collins (2004). "Genome sequence of the Brown Norway rat yields insights into mammalian evolution." Nature **428**(6982): 493-521.

Gifford, C. A., K. Racicot, D. S. Clark, K. J. Austin, T. R. Hansen, M. C. Lucy, C. J. Davies and T. L. Ott (2007). "Regulation of interferon-stimulated genes in peripheral blood leukocytes in pregnant and bred, nonpregnant dairy cows." J Dairy Sci **90**(1): 274-280.

Girard, A., R. Sachidanandam, G. J. Hannon and M. A. Carmell (2006). "A germline-specific class of small RNAs binds mammalian Piwi proteins." Nature **442**(7099): 199-202.

Giresi, P. G., J. Kim, R. M. McDaniell, V. R. Iyer and J. D. Lieb (2007). "FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin." Genome Res **17**(6): 877-885.

Glazko, G. V., E. V. Koonin, I. B. Rogozin and S. A. Shabalina (2003). "A significant fraction of conserved noncoding DNA in human and mouse consists of predicted matrix attachment regions." Trends Genet **19**(3): 119-124.

Goetz, T. L., T. L. Lloyd and M. D. Griswold (1996). "Role of E box and initiator region in the expression of the rat follicle-stimulating hormone receptor." J Biol Chem **271**(52): 33317-33324.

Gombert, W. M., S. D. Farris, E. D. Rubio, K. M. Morey-Rosler, W. H. Schubach and A. Krumm (2003). "The c-myc insulator element and matrix attachment regions define the c-myc chromosomal domain." Mol Cell Biol **23**(24): 9338-9348.

Gong, X. and E. A. McGee (2009). "Smad3 Is Required for Normal Follicular Follicle-Stimulating Hormone Responsiveness in the Mouse." Biol Reprod.

Good, M. C., J. G. Zalatan and W. A. Lim (2011). "Scaffold proteins: hubs for controlling the flow of cellular information." Science **332**(6030): 680-686.

Gorbman, A. and S. A. Sower (2003). "Evolution of the role of GnRH in animal (Metazoan) biology." Gen Comp Endocrinol **134**(3): 207-213.

Gottgens, B., L. M. Barton, J. G. Gilbert, A. J. Bench, M. J. Sanchez, S. Bahn, S. Mistry, D. Grafham, A. McMurray, M. Vaudin, E. Amaya, D. R. Bentley, A. R. Green and A. M. Sinclair (2000). "Analysis of vertebrate SCL loci identifies conserved enhancers." Nat Biotechnol **18**(2): 181-186.

Grasso, P. and L. E. Reichert, Jr. (1990). "Follicle-stimulating hormone receptor-mediated uptake of 45Ca2+ by cultured rat Sertoli cells does not require activation of cholera toxin- or pertussis toxin-sensitive guanine nucleotide binding proteins or adenylate cyclase." Endocrinology **127**(2): 949-956.

Griswold, M. D. and J. S. Kim (2001). "Site-specific methylation of the promoter alters deoxyribonucleic acid-protein interactions and prevents follicle-stimulating hormone receptor gene transcription." Biol Reprod **64**(2): 602-610.

Grivna, S. T., B. Pyhtila and H. Lin (2006). "MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis." Proc Natl Acad Sci U S A **103**(36): 13415-13420.

Gromoll, J., B. Dankbar and T. Gudermann (1994). "Characterization of the 5' flanking region of the human follicle-stimulating hormone receptor gene." Mol Cell Endocrinol **102**(1-2): 93-102.

Gromoll, J., E. Pekel and E. Nieschlag (1996). "The structure and organization of the human follicle-stimulating hormone receptor (FSHR) gene." Genomics **35**(2): 308-311.

Gromoll, J., T. Ried, H. Holtgreve-Grez, E. Nieschlag and T. Gudermann (1994). "Localization of the human FSH receptor to chromosome 2 p21 using a genomic probe comprising exon 10." J Mol Endocrinol **12**(3): 265-271.

Grosveld, F., G. B. van Assendelft, D. R. Greaves and G. Kollias (1987). "Position-independent, high-level expression of the human beta-globin gene in transgenic mice." Cell **51**(6): 975-985.

Gu, P., B. Goodwin, A. C. Chung, X. Xu, D. A. Wheeler, R. R. Price, C. Galardi, L. Peng, A. M. Latour, B. H. Koller, J. Gossen, S. A. Kliewer and A. J. Cooney (2005). "Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development." Mol Cell Biol **25**(9): 3492-3505.

Guil, S. and M. Esteller (2012). "Cis-acting noncoding RNAs: friends and foes." Nat Struct Mol Biol **19**(11): 1068-1075.

Guil, S., M. Soler, A. Portela, J. Carrere, E. Fonalleras, A. Gomez, A. Villanueva and M. Esteller (2012). "Intronic RNAs mediate EZH2 regulation of epigenetic targets." Nat Struct Mol Biol **19**(7): 664-670.

Gupta, R. A., N. Shah, K. C. Wang, J. Kim, H. M. Horlings, D. J. Wong, M. C. Tsai, T. Hung, P. Argani, J. L. Rinn, Y. Wang, P. Brzoska, B. Kong, R. Li, R. B. West, M. J. van de Vijver, S. Sukumar and H. Y. Chang (2010). "Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis." Nature **464**(7291): 1071-1076.

Guttman, M., I. Amit, M. Garber, C. French, M. F. Lin, D. Feldser, M. Huarte, O. Zuk, B. W. Carey, J. P. Cassady, M. N. Cabili, R. Jaenisch, T. S. Mikkelsen, T. Jacks, N. Hacohen, B. E. Bernstein, M. Kellis, A. Regev, J. L. Rinn and E. S. Lander (2009). "Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals." Nature **458**(7235): 223-227.

Hadjur, S., L. M. Williams, N. K. Ryan, B. S. Cobb, T. Sexton, P. Fraser, A. G. Fisher and M. Merkenschlager (2009). "Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus." Nature **460**(7253): 410-413.

Haisenleder, D. J., H. A. Ferris and M. A. Shupnik (2003). "The calcium component of gonadotropin-releasing hormone-stimulated luteinizing hormone subunit gene transcription is mediated by calcium/calmodulin-dependent protein kinase type II." Endocrinology **144**(6): 2409-2416.

Hamalainen, T., M. Poutanen and I. Huhtaniemi (1999). "Age- and sex-specific promoter function of a 2-kilobase 5'-flanking sequence of the murine luteinizing hormone receptor gene in transgenic mice." Endocrinology **140**(11): 5322-5329.

Hamalainen, T., M. Poutanen and I. Huhtaniemi (2001). "Promoter function of different lengths of the murine luteinizing hormone receptor gene 5'-flanking region in transfected gonadal cells and in transgenic mice." Endocrinology **142**(6): 2427-2434.

Hammer, G. D., K. L. Parker and B. P. Schimmer (2005). "Minireview: transcriptional regulation of adrenocortical development." Endocrinology **146**(3): 1018-1024.

Hanley, N. A., S. G. Ball, M. Clement-Jones, D. M. Hagan, T. Strachan, S. Lindsay, S. Robson, H. Ostrer, K. L. Parker and D. I. Wilson (1999). "Expression of steroidogenic factor 1 and Wilms' tumour 1 during early human gonadal development and sex determination." Mech Dev **87**(1-2): 175-180.

Hanley, N. A., S. G. Ball, M. Clement-Jones, D. M. Hagan, T. Strachan, S. Lindsay, S. Robson, H. Ostrer, K. L. Parker and D. I. Wilson (1999). "Expression of steroidogenic factor 1 and Wilms' tumour 1 during early human gonadal development and sex determination." Mechanisms of development **87**(1-2): 175-180.

Hanley, N. A., W. E. Rainey, D. I. Wilson, S. G. Ball and K. L. Parker (2001). "Expression profiles of SF-1, DAX1, and CYP17 in the human fetal adrenal gland: potential interactions in gene regulation." Mol Endocrinol **15**(1): 57-68.

Hanley, N. A., W. E. Rainey, D. I. Wilson, S. G. Ball and K. L. Parker (2001). "Expression profiles of SF-1, DAX1, and CYP17 in the human fetal adrenal gland: potential interactions in gene regulation." Molecular endocrinology **15**(1): 57-68.

Hardison, R. C. (2000). "Conserved noncoding sequences are reliable guides to regulatory elements." Trends Genet **16**(9): 369-372.

Hare, E. E., B. K. Peterson, V. N. Iyer, R. Meier and M. B. Eisen (2008). "Sepsid even-skipped enhancers are functionally conserved in Drosophila despite lack of sequence conservation." PLoS Genet **4**(6): e1000106.

Hark, A. T., C. J. Schoenherr, D. J. Katz, R. S. Ingram, J. M. Levorse and S. M. Tilghman (2000). "CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus." Nature **405**(6785): 486-489.

Harris, A. N. and P. L. Mellon (1998). "The basic helix-loop-helix, leucine zipper transcription factor, USF (upstream stimulatory factor), is a key regulator of SF-1 (steroidogenic factor-1) gene expression in pituitary gonadotrope and steroidogenic cells." Mol Endocrinol **12**(5): 714-726.

Harris, D., D. Bonfil, D. Chuderland, S. Kraus, R. Seger and Z. Naor (2002). "Activation of MAPK cascades by GnRH: ERK and Jun N-terminal kinase are involved in basal and GnRH-stimulated activity of the glycoprotein hormone LHbeta-subunit promoter." Endocrinology **143**(3): 1018-1025.

Harris, G. W. (1937). "The Induction of Ovulation in the Rabbit, by Electrical Stimulation of the Hypothalamo-hypophysial Mechanism." Proceedings of the Royal Society of London. Series B, Biological Sciences **122**(828): 374-394.

Harris, M. B., J. Mostecki and P. B. Rothman (2005). "Repression of an interleukin-4-responsive promoter requires cooperative BCL-6 function." J Biol Chem **280**(13): 13114-13121.

Hatano, O., A. Takakusu, M. Nomura and K. Morohashi (1996). "Identical origin of adrenal cortex and gonad revealed by expression profiles of Ad4BP/SF-1." Genes to cells : devoted to molecular & cellular mechanisms **1**(7): 663-671.

Hatano, O., K. Takayama, T. Imai, M. R. Waterman, A. Takakusu, T. Omura and K. Morohashi (1994). "Sex-dependent expression of a transcription factor, Ad4BP, regulating steroidogenic P-450 genes in the gonads during prenatal and postnatal rat development." Development **120**(10): 2787-2797.

Havlak, P., R. Chen, K. J. Durbin, A. Egan, Y. Ren, X. Z. Song, G. M. Weinstock and R. A. Gibbs (2004). "The Atlas genome assembly system." Genome Res **14**(4): 721-732.

He, L. and G. J. Hannon (2004). "MicroRNAs: small RNAs with a big role in gene regulation." Nat Rev Genet **5**(7): 522-531.

Heath, H., C. Ribeiro de Almeida, F. Sleutels, G. Dingjan, S. van de Nobelen, I. Jonkers, K. W. Ling, J. Gribnau, R. Renkawitz, F. Grosveld, R. W. Hendriks and N. Galjart (2008). "CTCF regulates cell cycle progression of alphabeta T cells in the thymus." EMBO J **27**(21): 2839-2850.

Heckert, L. and M. D. Griswold (1993). "Expression of the FSH receptor in the testis." Recent Prog Horm Res **48**: 61-77.

Heckert, L. L. (2001). "Activation of the rat follicle-stimulating hormone receptor promoter by steroidogenic factor 1 is blocked by protein kinase a and requires upstream stimulatory factor binding to a proximal E box element." Mol Endocrinol **15**(5): 704-715.

Heckert, L. L. (2005). Structure and Regulation of the FSH Receptor Gene. Sertoli Cell Biology. M. K. a. G. Skinner, M.D.**:** 281-299.

Heckert, L. L., M. A. Daggett and J. Chen (1998). "Multiple promoter elements contribute to activity of the follicle-stimulating hormone receptor (FSHR) gene in testicular Sertoli cells." Mol Endocrinol **12**(10): 1499-1512.

Heckert, L. L., I. J. Daley and M. D. Griswold (1992). "Structural organization of the follicle-stimulating hormone receptor gene." Mol Endocrinol **6**(1): 70-80.

Heckert, L. L. and M. D. Griswold (1991). "Expression of follicle-stimulating hormone receptor mRNA in rat testes and Sertoli cells." Mol Endocrinol **5**(5): 670-677.

Heckert, L. L. and M. D. Griswold (2002). "The expression of the follicle-stimulating hormone receptor in spermatogenesis." Recent Prog Horm Res **57**: 129-148.

Heckert, L. L., M. Sawadogo, M. A. Daggett and J. K. Chen (2000). "The USF proteins regulate transcription of the follicle-stimulating hormone receptor but are insufficient for cell-specific expression." Mol Endocrinol **14**(11): 1836-1848.

Heindel, J. J., R. Rothenberg, G. A. Robison and A. Steinberger (1975). "LH and FSH stimulation of cyclic AMP in specific cell types isolated from the testes." J Cyclic Nucleotide Res **1**(2): 69-79.

Heo, J. B. and S. Sung (2011). "Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA." Science **331**(6013): 76-79.

Hermann, B. P. and L. L. Heckert (2005). "Silencing of Fshr occurs through a conserved, hypersensitive site in the first intron." Mol Endocrinol **19**(8): 2112-2131.

Hermann, B. P., K. Hornbaker, D. A. Rice, M. Sawadogo and L. L. Heckert (2008). "In vivo regulation of follicle-stimulating hormone receptor by the transcription factors upstream stimulatory factor 1 and upstream stimulatory factor 2 is cell specific." Endocrinology **149**(10): 5297-5306.

Hermann, B. P., K. I. Hornbaker, R. R. Maran and L. L. Heckert (2007). "Distal regulatory elements are required for Fshr expression, in vivo." Mol Cell Endocrinol **260-262**: 49-58.

Herold, M., M. Bartkuhn and R. Renkawitz (2012). "CTCF: insights into insulator function during development." Development **139**(6): 1045-1057.

Hiroi, H., L. K. Christenson and J. F. Strauss, 3rd (2004). "Regulation of transcription of the steroidogenic acute regulatory protein (StAR) gene: temporal and spatial changes in transcription factor binding and histone modification." Mol Cell Endocrinol **215**(1-2): 119-126.

Hohlweg, W. (1975). The Regulatory Centers of Endocrine Glands in the Hypothalamus. Pioneers in Neuroendocrinology. J. Meites, B. Donovan and S. McCann, Springer US. **1:** 159-172.

Hohlweg, W. and K. Junkmann (1932). "The hormonal and neurogenic regulation of the function of the anterior pituitary." Klinische Wochenschrift **11**(8): 321-323.

Hoivik, E. A., L. Aumo, R. Aesoy, H. Lillefosse, A. E. Lewis, R. M. Perrett, N. R. Stallings, N. A. Hanley and M. Bakke (2008). "Deoxyribonucleic acid methylation controls cell type-specific expression of steroidogenic factor 1." Endocrinology **149**(11): 5599-5609.

Holland, L. Z., R. Albalat, K. Azumi, E. Benito-Gutierrez, M. J. Blow, M. Bronner-Fraser, F. Brunet, T. Butts, S. Candiani, L. J. Dishaw, D. E. Ferrier, J. Garcia-Fernandez, J. J. Gibson-Brown, C. Gissi, A. Godzik, F. Hallbook, D. Hirose, K. Hosomichi, T. Ikuta, H. Inoko, M. Kasahara, J. Kasamatsu, T. Kawashima, A. Kimura, M. Kobayashi, Z. Kozmik, K. Kubokawa, V. Laudet, G. W. Litman, A. C. McHardy, D. Meulemans, M. Nonaka, R. P. Olinski, Z. Pancer, L. A. Pennacchio, M. Pestarino, J. P. Rast, I. Rigoutsos, M. Robinson-Rechavi, G. Roch, H. Saiga, Y. Sasakura, M. Satake, Y. Satou, M. Schubert, N. Sherwood, T. Shiina, N. Takatori, J. Tello, P. Vopalensky, S. Wada, A. Xu, Y. Ye, K. Yoshida, F. Yoshizaki, J. K. Yu, Q. Zhang, C. M. Zmasek, P. J. de Jong, K. Osoegawa, N. H. Putnam, D. S. Rokhsar, N. Satoh and P. W. Holland (2008). "The amphioxus genome illuminates vertebrate origins and cephalochordate biology." Genome Res **18**(7): 1100-1111.

Honda, S., K. Morohashi, M. Nomura, H. Takeya, M. Kitajima and T. Omura (1993). "Ad4BP regulating steroidogenic P-450 gene is a member of steroid hormone receptor superfamily." The Journal of biological chemistry **268**(10): 7494-7502.

Honda, S., K. Morohashi, M. Nomura, H. Takeya, M. Kitajima and T. Omura (1993). "Ad4BP regulating steroidogenic P-450 gene is a member of steroid hormone receptor superfamily." J Biol Chem **268**(10): 7494-7502.
Horwich, M. D., C. Li, C. Matranga, V. Vagin, G. Farley, P. Wang and P. D. Zamore (2007). "The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC." Curr Biol **17**(14): 1265-1272.

Houwing, S., L. M. Kamminga, E. Berezikov, D. Cronembold, A. Girard, H. van den Elst, D. V. Filippov, H. Blaser, E. Raz, C. B. Moens, R. H. Plasterk, G. J. Hannon, B. W. Draper and R. F. Ketting (2007). "A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish." Cell **129**(1): 69-82.

Huarte, M., M. Guttman, D. Feldser, M. Garber, M. J. Koziol, D. Kenzelmann-Broz, A. M. Khalil, O. Zuk, I. Amit, M. Rabani, L. D. Attardi, A. Regev, E. S. Lander, T. Jacks and J. L. Rinn (2010). "A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response." Cell **142**(3): 409-419.

Huhtaniemi, I. T., V. Eskola, P. Pakarinen, T. Matikainen and R. Sprengel (1992). "The murine luteinizing hormone and follicle-stimulating hormone receptor genes: transcription initiation sites, putative promoter sequences and promoter activity." Mol Cell Endocrinol **88**(1-3): 55-66.

Huhtaniemi, I. T. and A. P. Themmen (2005). "Mutations in human gonadotropin and gonadotropin-receptor genes." Endocrine **26**(3): 207-217.

Human Genome Sequencing, C. (2004). "Finishing the euchromatic sequence of the human genome." Nature **431**(7011): 931-945.

Hummler, E., T. J. Cole, J. A. Blendy, R. Ganss, A. Aguzzi, W. Schmid, F. Beermann and G. Schutz (1994). "Targeted mutation of the CREB gene: compensation within the CREB/ATF family of transcription factors." Proc Natl Acad Sci U S A **91**(12): 5647-5651.

Huttenhofer, A., P. Schattner and N. Polacek (2005). "Non-coding RNAs: hope or hype?" Trends Genet **21**(5): 289-297.

Ikeda, Y. (1996). "SF-1: a key regulator of development and function in the mammalian reproductive system." Acta Paediatr Jpn **38**(4): 412-419.

Ikeda, Y., D. S. Lala, X. Luo, E. Kim, M. P. Moisan and K. L. Parker (1993). "Characterization of the mouse FTZ-F1 gene, which encodes a key regulator of steroid hydroxylase gene expression." Molecular endocrinology **7**(7): 852-860.

Ikeda, Y., D. S. Lala, X. Luo, E. Kim, M. P. Moisan and K. L. Parker (1993). "Characterization of the mouse FTZ-F1 gene, which encodes a key regulator of steroid hydroxylase gene expression." Mol Endocrinol **7**(7): 852-860.

Ikeda, Y., X. Luo, R. Abbud, J. H. Nilson and K. L. Parker (1995). "The nuclear receptor steroidogenic factor 1 is essential for the formation of the ventromedial hypothalamic nucleus." Mol Endocrinol **9**(4): 478-486.

Ikeda, Y., W. H. Shen, H. A. Ingraham and K. L. Parker (1994). "Developmental expression of mouse steroidogenic factor-1, an essential regulator of the steroid hydroxylases." Molecular endocrinology **8**(5): 654-662.

Ikeda, Y., W. H. Shen, H. A. Ingraham and K. L. Parker (1994). "Developmental expression of mouse steroidogenic factor-1, an essential regulator of the steroid hydroxylases." Mol Endocrinol **8**(5): 654-662.
Ikeda, Y., W. H. Shen, H. A. Ingraham and K. L. Parker (1994). "Developmental expression of mouse steroidogenic factor-1, an essential regulator of the steroid hydroxylases." Mol Endocrinol **8**(5): 654-662.
Ingraham, H. A., D. S. Lala, Y. Ikeda, X. Luo, W. H. Shen, M. W. Nachtigal, R. Abbud, J. H. Nilson and K. L. Parker (1994). "The nuclear receptor steroidogenic factor 1 acts at multiple levels of the reproductive axis." Genes Dev **8**(19): 2302-2312.

Ingraham, H. A., D. S. Lala, Y. Ikeda, X. Luo, W. H. Shen, M. W. Nachtigal, R. Abbud, J. H. Nilson and K. L. Parker (1994). "The nuclear receptor steroidogenic factor 1 acts at multiple levels of the reproductive axis." Genes & development **8**(19): 2302-2312.

Jackson, R. J. and N. Standart (2007). "How do microRNAs regulate gene expression?" Sci STKE **2007**(367): re1.
Jacq, C., J. R. Miller and G. G. Brownlee (1977). "A pseudogene structure in 5S DNA of Xenopus laevis." Cell **12**(1): 109-120.

Jahnsen, T., J. O. Gordeladze, P. A. Torjesen and V. Hansson (1980). "FSH-response adenylyl cyclase in rat testes: desensitization by homologous hormone." Arch Androl **5**(2): 169-177.

Jakobsen, J. S., M. Braun, J. Astorga, E. H. Gustafson, T. Sandmann, M. Karzynski, P. Carlsson and E. E. Furlong (2007). "Temporal ChIP-on-chip reveals Biniou as a universal regulator of the visceral muscle transcriptional network." Genes Dev **21**(19): 2448-2460.

Jamieson, R. V., R. Perveen, B. Kerr, M. Carette, J. Yardley, E. Heon, M. G. Wirth, V. van Heyningen, D. Donnai, F. Munier and G. C. Black (2002). "Domain disruption and mutation of the bZIP transcription factor, MAF, associated with cataract, ocular anterior segment dysgenesis and coloboma." Hum Mol Genet **11**(1): 33-42.

Jeyasuria, P., Y. Ikeda, S. P. Jamin, L. Zhao, D. G. De Rooij, A. P. Themmen, R. R. Behringer and K. L. Parker (2004). "Cell-specific knockout of steroidogenic factor 1 reveals its essential roles in gonadal function." Molecular endocrinology **18**(7): 1610-1619.

Jeyasuria, P., Y. Ikeda, S. P. Jamin, L. Zhao, D. G. De Rooij, A. P. Themmen, R. R. Behringer and K. L. Parker (2004). "Cell-specific knockout of steroidogenic factor 1 reveals its essential roles in gonadal function." Mol Endocrinol **18**(7): 1610-1619.

Ji, P., S. Diederichs, W. Wang, S. Boing, R. Metzger, P. M. Schneider, N. Tidow, B. Brandt, H. Buerger, E. Bulk, M. Thomas, W. E. Berdel, H. Serve and C. Muller-Tidow (2003). "MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer." Oncogene **22**(39): 8031-8041.

Ji, Q., P. I. Liu, P. K. Chen and C. Aoyama (2004). "Follicle stimulating hormone-induced growth promotion and gene expression profiles on ovarian surface epithelial cells." Int J Cancer **112**(5): 803-814.
Johnson, R. S., B. M. Spiegelman and V. Papaioannou (1992). "Pleiotropic effects of a null mutation in the c-fos proto-oncogene." Cell **71**(4): 577-586.

Johnsson, P., A. Ackley, L. Vidarsdottir, W. O. Lui, M. Corcoran, D. Grander and K. V. Morris (2013). "A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells." Nat Struct Mol Biol **20**(4): 440-446.

Johnston, C. M., A. E. Newall, N. Brockdorff and T. B. Nesterova (2002). "Enox, a novel gene that maps 10 kb upstream of Xist and partially escapes X inactivation." Genomics **80**(2): 236-244.

Jothi, R., S. Cuddapah, A. Barski, K. Cui and K. Zhao (2008). "Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data." Nucleic Acids Res **36**(16): 5221-5231.

Junion, G., M. Spivakov, C. Girardot, M. Braun, E. H. Gustafson, E. Birney and E. E. Furlong (2012). "A transcription factor collective defines cardiac cell fate and reflects lineage history." Cell **148**(3): 473-486.
Juven-Gershon, T., J. Y. Hsu, J. W. Theisen and J. T. Kadonaga (2008). "The RNA polymerase II core promoter - the gateway to transcription." Curr Opin Cell Biol **20**(3): 253-259.

Kaiser, U. B., P. M. Conn and W. W. Chin (1997). "Studies of gonadotropin-releasing hormone (GnRH) action using GnRH receptor-expressing pituitary cell lines." Endocr Rev **18**(1): 46-70.

Kakar, S. S., S. J. Winters, W. Zacharias, D. M. Miller and S. Flynn (2003). "Identification of distinct gene expression profiles associated with treatment of LbetaT2 cells with gonadotropin-releasing hormone agonist using microarray analysis." Gene **308**: 67-77.

Kaminski, T., B. Gawronska, K. Derecka, S. Okrasa and J. Przala (2000). "Gene expression and peptide localization for LH/hCG receptor in porcine small and large luteal cells: possible regulation by opioid peptides." J Physiol Pharmacol **51**(2): 359-368.

Kanduri, C., N. Thakur and R. R. Pandey (2006). "The length of the transcript encoded from the Kcnq1ot1 antisense promoter determines the degree of silencing." EMBO J **25**(10): 2096-2106.

Karl, A. F. and M. D. Griswold (1990). "Sertoli cells of the testis: preparation of cell cultures and effects of retinoids." Methods Enzymol **190**: 71-75.
Karpova, T., J. Presley, R. R. Manimaran, S. P. Scherrer, L. Tejada, K. R. Peterson and L. L. Heckert (2005). "A FTZ-F1-containing yeast artificial chromosome recapitulates expression of steroidogenic factor 1 in vivo." Mol Endocrinol **19**(10): 2549-2563.

Kee, B. L. (2009). "E and ID proteins branch out." Nat Rev Immunol **9**(3): 175-184.

Ketelslegers, J. M. and K. J. Catt (1974). "Receptor binding properties of 125I-hFSH prepared by enzymatic iodination." J Clin Endocrinol Metab **39**(6): 1159-1162.

Khalil, A. M., M. Guttman, M. Huarte, M. Garber, A. Raj, D. Rivea Morales, K. Thomas, A. Presser, B. E. Bernstein, A. van Oudenaarden, A. Regev, E. S. Lander and J. L. Rinn (2009). "Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression." Proc Natl Acad Sci U S A **106**(28): 11667-11672.

Kikuta, H., M. Laplante, P. Navratilova, A. Z. Komisarczuk, P. G. Engstrom, D. Fredman, A. Akalin, M. Caccamo, I. Sealy, K. Howe, J. Ghislain, G. Pezeron, P. Mourrain, S. Ellingsen, A. C. Oates, C. Thisse, B. Thisse, I. Foucher, B. Adolf, A. Geling, B. Lenhard and T. S. Becker (2007). "Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates." Genome Res **17**(5): 545-555.

Kiltie, A. E. (2010). "Common predisposition alleles for moderately common cancers: bladder cancer." Curr Opin Genet Dev.

Kim, J. S. and M. D. Griswold (2001). "E2F and GATA-1 are required for the Sertoli cell-specific promoter activity of the follicle-stimulating hormone receptor gene." J Androl **22**(4): 629-639.

Kim, K. W., S. Li, H. Zhao, B. Peng, S. A. Tobet, J. K. Elmquist, K. L. Parker and L. Zhao (2010). "CNS-specific ablation of steroidogenic factor 1 results in impaired female reproductive function." Mol Endocrinol **24**(6): 1240-1250.

Kim, S., W. Hwang do and D. S. Lee (2009). "A study of microRNAs in silico and in vivo: bioimaging of microRNA biogenesis and regulation." FEBS J **276**(8): 2165-2174.

Kim, T. H., Z. K. Abdullaev, A. D. Smith, K. A. Ching, D. I. Loukinov, R. D. Green, M. Q. Zhang, V. V. Lobanenkov and B. Ren (2007). "Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome." Cell **128**(6): 1231-1245.

Kimura, R., H. Yoshii, M. Nomura, N. Kotomura, T. Mukai, S. Ishihara, K. Ohba, T. Yanase, O. Gotoh, H. Nawata and K. Morohashi (2000). "Identification of novel first exons in Ad4BP/SF-1 (NR5A1) gene and their tissue- and species-specific usage." Biochem Biophys Res Commun **278**(1): 63-71.

King, D. C., J. Taylor, L. Elnitski, F. Chiaromonte, W. Miller and R. C. Hardison (2005). "Evaluation of regulatory potential and conservation scores for detecting cis-regulatory modules in aligned mammalian genome sequences." Genome Res **15**(8): 1051-1060.

King, D. C., J. Taylor, Y. Zhang, Y. Cheng, H. A. Lawson, J. Martin, F. Chiaromonte, W. Miller and R. C. Hardison (2007). "Finding cis-regulatory elements using comparative genomics: some lessons from ENCODE data." Genome Res **17**(6): 775-786.

Kioussis, D., E. Vanin, T. deLange, R. A. Flavell and F. G. Grosveld (1983). "Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia." Nature **306**(5944): 662-666.

Kishi, H., T. Minegishi, M. Tano, T. Kameda, Y. Ibuki and K. Miyamoto (1998). "The effect of activin and FSH on the differentiation of rat granulosa cells." FEBS Lett **422**(2): 274-278.

Kliesch, S., T. L. Penttila, J. Gromoll, P. T. K. Saunders, E. Nieschlag and M. Parvinen (1992). "Fsh Receptor Messenger-Rna Is Expressed Stage-Dependently during Rat Spermatogenesis." Molecular and Cellular Endocrinology **84**(3): R45-R49.

Knecht, M., T. Ranta and K. J. Catt (1983). "Granulosa cell differentiation in vitro: induction and maintenance of follicle-stimulating hormone receptors by adenosine 3',5'-monophosphate." Endocrinology **113**(3): 949-956.

Kolbe, D., J. Taylor, L. Elnitski, P. Eswara, J. Li, W. Miller, R. Hardison and F. Chiaromonte (2004). "Regulatory potential scores from genome-wide three-way alignments of human, mouse, and rat." Genome Res **14**(4): 700-707.

Kollias, G., N. Wrighton, J. Hurst and F. Grosveld (1986). "Regulated expression of human A gamma-, beta-, and hybrid gamma beta-globin genes in transgenic mice: manipulation of the developmental expression patterns." Cell **46**(1): 89-94.

Koo, Y. B., I. Ji, R. G. Slaughter and T. H. Ji (1991). "Structure of the luteinizing hormone receptor gene and multiple exons of the coding sequence." Endocrinology **128**(5): 2297-2308.

Koonin, E. V. (2003). "Comparative genomics, minimal gene-sets and the last universal common ancestor." Nat Rev Microbiol **1**(2): 127-136.

Krylova, I. N., E. P. Sablin, J. Moore, R. X. Xu, G. M. Waitt, J. A. MacKay, D. Juzumiene, J. M. Bynum, K. Madauss, V. Montana, L. Lebedeva, M. Suzawa, J. D. Williams, S. P. Williams, R. K. Guy, J. W. Thornton, R. J. Fletterick, T. M. Willson and H. A. Ingraham (2005). "Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1." Cell **120**(3): 343-355.

Kulkarni, M. M. and D. N. Arnosti (2003). "Information display by transcriptional enhancers." Development **130**(26): 6569-6575.

Kulkarni, M. M. and D. N. Arnosti (2005). "cis-regulatory logic of short-range transcriptional repression in Drosophila melanogaster." Mol Cell Biol **25**(9): 3411-3420.

Kumar, T. R., Y. Wang, N. Lu and M. M. Matzuk (1997). "Follicle stimulating hormone is required for ovarian follicle maturation but not male fertility." Nat Genet **15**(2): 201-204.

Kung, J. T., D. Colognori and J. T. Lee (2013). "Long noncoding RNAs: past, present, and future." Genetics **193**(3): 651-669.

Lagerstrom, M. C. and H. B. Schioth (2008). "Structural diversity of G protein-coupled receptors and significance for drug discovery." Nature Reviews Drug Discovery **7**(4): 339-357.

Lala, D. S., D. A. Rice and K. L. Parker (1992). "Steroidogenic factor I, a key regulator of steroidogenic enzyme expression, is the mouse homolog of fushi tarazu-factor I." Mol Endocrinol **6**(8): 1249-1258.

Lala, D. S., D. A. Rice and K. L. Parker (1992). "Steroidogenic factor I, a key regulator of steroidogenic enzyme expression, is the mouse homolog of fushi tarazu-factor I." Molecular endocrinology **6**(8): 1249-1258.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H.

Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi and Y. J. Chen (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.

Lang, G., C. Mamalaki, D. Greenberg, N. Yannoutsos and D. Kioussis (1991). "Deletion analysis of the human CD2 gene locus control region in transgenic mice." Nucleic Acids Res **19**(21): 5851-5856.

Lau, N. C., A. G. Seto, J. Kim, S. Kuramochi-Miyagawa, T. Nakano, D. P. Bartel and R. E. Kingston (2006). "Characterization of the piRNA complex from rat testes." Science **313**(5785): 363-367.

Lee, G. R., P. E. Fields, T. J. Griffin and R. A. Flavell (2003). "Regulation of the Th2 cytokine locus by a locus control region." Immunity **19**(1): 145-153.

Lee, J. T. (2011). "Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control." Nat Rev Mol Cell Biol **12**(12): 815-826.

Lee, J. T., L. S. Davidow and D. Warshawsky (1999). "Tsix, a gene antisense to Xist at the X-inactivation centre." Nat Genet **21**(4): 400-404.

Lee, M. P., M. R. DeBaun, K. Mitsuya, H. L. Galonek, S. Brandenburg, M. Oshimura and A. P. Feinberg (1999). "Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting." Proc Natl Acad Sci U S A **96**(9): 5203-5208.

Lee, M. P., K. Howcroft, A. Kotekar, H. H. Yang, K. H. Buetow and D. S. Singer (2005). "ATG deserts define a novel core promoter subclass." Genome Res **15**(9): 1189-1197.

Lee, S. L., Y. Sadovsky, A. H. Swirnoff, J. A. Polish, P. Goda, G. Gavrilina and J. Milbrandt (1996). "Luteinizing hormone deficiency and female infertility in mice lacking the transcription factor NGFI-A (Egr-1)." Science **273**(5279): 1219-1221.

Lee, T. I. and R. A. Young (2000). "Transcription of eukaryotic protein-coding genes." Annu Rev Genet **34**: 77-137.

Lei, N. and L. L. Heckert (2004). "Gata4 regulates testis expression of Dmrt1." Mol Cell Biol **24**(1): 377-388.

Levallet, J., P. Koskimies, N. Rahman and I. Huhtaniemi (2001). "The promoter of murine follicle-stimulating hormone receptor: functional characterization and regulation by transcription factor steroidogenic factor 1." Mol Endocrinol **15**(1): 80-92.

Li, L., S. He, J. M. Sun and J. R. Davie (2004). "Gene regulation by Sp1 and Sp3." Biochem Cell Biol **82**(4): 460-471.

Li, Q., K. R. Peterson, X. Fang and G. Stamatoyannopoulos (2002). "Locus control regions." Blood **100**(9): 3077-3086.

Li, T., J. F. Hu, X. Qiu, J. Ling, H. Chen, S. Wang, A. Hou, T. H. Vu and A. R. Hoffman (2008). "CTCF regulates allelic expression of Igf2 by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop." Mol Cell Biol **28**(20): 6473-6482.

Li, Y., M. Choi, G. Cavey, J. Daugherty, K. Suino, A. Kovach, N. C. Bingham, S. A. Kliewer and H. E. Xu (2005). "Crystallographic identification and functional characterization of phospholipids as ligands for the orphan nuclear receptor steroidogenic factor-1." Mol Cell **17**(4): 491-502.

Lin, R., M. Roychowdhury-Saha, C. Black, A. T. Watt, E. G. Marcusson, S. M. Freier and T. S. Edgington (2011). "Control of RNA processing by a large non-coding RNA over-expressed in carcinomas." FEBS Lett **585**(4): 671-676.

Lin, Y. C., S. Jhunjhunwala, C. Benner, S. Heinz, E. Welinder, R. Mansson, M. Sigvardsson, J. Hagman, C. A. Espinoza, J. Dutkowski, T. Ideker, C. K. Glass and C. Murre (2010). "A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate." Nat Immunol **11**(7): 635-643.

Liska, F., P. Snajdr, L. Sedova, O. Seda, B. Chylikova, P. Slamova, E. Krejci, D. Sedmera, M. Grim, D. Krenova and V. Kren (2009). "Deletion of a conserved noncoding sequence in Plzf intron leads to Plzf down-regulation in limb bud and polydactyly in the rat." Dev Dyn **238**(3): 673-684.

Litt, M. D., M. Simpson, M. Gaszner, C. D. Allis and G. Felsenfeld (2001). "Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus." Science **293**(5539): 2453-2455.

Litt, M. D., M. Simpson, F. Recillas-Targa, M. N. Prioleau and G. Felsenfeld (2001). "Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci." EMBO J **20**(9): 2224-2235.

Liu, F., D. A. Austin, P. L. Mellon, J. M. Olefsky and N. J. Webster (2002). "GnRH activates ERK1/2 leading to the induction of c-fos and LHbeta protein expression in LbetaT2 cells." Mol Endocrinol **16**(3): 419-434.

Liu, F., M. S. Ruiz, D. A. Austin and N. J. Webster (2005). "Constitutively active Gq impairs gonadotropin-releasing hormone-induced intracellular signaling and luteinizing hormone secretion in LbetaT2 cells." Mol Endocrinol **19**(8): 2074-2085.

Liu, X., K. Fortin and Z. Mourelatos (2008). "MicroRNAs: biogenesis and molecular functions." Brain Pathol **18**(1): 113-121.

Lodder, E. M., B. H. Eussen, D. A. van Hassel, A. J. Hoogeboom, P. J. Poddighe, J. H. Coert, B. A. Oostra, A. de Klein and E. de Graaff (2009). "Implication of long-distance regulation of the HOXA cluster in a patient with postaxial polydactyly." Chromosome Res **17**(6): 737-744.

Ludwig, M. Z., N. H. Patel and M. Kreitman (1998). "Functional analysis of eve stripe 2 enhancer evolution in Drosophila: rules governing conservation and change." Development **125**(5): 949-958.

Lund, E. and J. E. Dahlberg (2006). "Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs." Cold Spring Harb Symp Quant Biol **71**: 59-66.
Luo, X., Y. Ikeda, D. S. Lala, L. A. Baity, J. C. Meade and K. L. Parker (1995). "A cell-specific nuclear receptor plays essential roles in adrenal and gonadal development." Endocr Res **21**(1-2): 517-524.

Luo, X., Y. Ikeda, D. S. Lala, L. A. Baity, J. C. Meade and K. L. Parker (1995). "A cell-specific nuclear receptor plays essential roles in adrenal and gonadal development." Endocrine research **21**(1-2): 517-524.

Luo, X., Y. Ikeda and K. L. Parker (1994). "A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation." Cell **77**(4): 481-490.

Luo, X., Y. Ikeda, D. A. Schlosser and K. L. Parker (1995). "Steroidogenic factor 1 is the essential transcript of the mouse Ftz-F1 gene." Molecular endocrinology **9**(9): 1233-1239.

Lustig, A. J. (2004). "Telomerase RNA: a flexible RNA scaffold for telomerase biosynthesis." Curr Biol **14**(14): R565-567.

Lyden, D., A. Z. Young, D. Zagzag, W. Yan, W. Gerald, R. O'Reilly, B. L. Bader, R. O. Hynes, Y. Zhuang, K. Manova and R. Benezra (1999). "Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts." Nature **401**(6754): 670-677.

Lyle, R., D. Watanabe, D. te Vruchte, W. Lerchner, O. W. Smrzka, A. Wutz, J. Schageman, L. Hahner, C. Davies and D. P. Barlow (2000). "The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1." Nat Genet **25**(1): 19-21.
Lyon, M. F. (1961). "Gene action in the X-chromosome of the mouse (Mus musculus L.)." Nature **190**: 372-373.

Ma, M. K., C. Heath, A. Hair and A. G. West (2011). "Histone crosstalk directed by H2B ubiquitination is required for chromatin boundary integrity." PLoS Genet **7**(7): e1002175.

Magram, J., K. Chada and F. Costantini (1985). "Developmental regulation of a cloned adult beta-globin gene in transgenic mice." Nature **315**(6017): 338-340.

Majdic, G., M. Young, E. Gomez-Sanchez, P. Anderson, L. S. Szczepaniak, R. L. Dobbins, J. D. McGarry and K. L. Parker (2002). "Knockout mice lacking steroidogenic factor 1 are a novel genetic model of hypothalamic obesity." Endocrinology **143**(2): 607-614.

Mariani, S., L. Salvatori, S. Basciani, M. Arizzi, G. Franco, E. Petrangeli, G. Spera and L. Gnessi (2006). "Expression and cellular localization of follicle-stimulating hormone receptor in normal human prostate, benign prostatic hyperplasia and prostate cancer." J Urol **175**(6): 2072-2077; discussion 2077.

Marson, A., S. S. Levine, M. F. Cole, G. M. Frampton, T. Brambrink, S. Johnstone, M. G. Guenther, W. K. Johnston, M. Wernig, J. Newman, J. M. Calabrese, L. M. Dennis, T. L. Volkert, S. Gupta, J. Love, N. Hannett, P. A. Sharp, D. P. Bartel, R. Jaenisch and R. A. Young (2008). "Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells." Cell **134**(3): 521-533.

Martianov, I., A. Ramadass, A. Serra Barros, N. Chow and A. Akoulitchev (2007). "Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript." Nature **445**(7128): 666-670.

Martinez-Jimenez, C. P., M. J. Gomez-Lechon, J. V. Castell and R. Jover (2005). "Transcriptional regulation of the human hepatic CYP3A4: identification of a new distal enhancer region responsive to CCAAT/enhancer-binding protein beta isoforms (liver activating protein and liver inhibitory protein)." Mol Pharmacol **67**(6): 2088-2101.

Mason, A. J., J. S. Hayflick, R. T. Zoeller, W. S. Young, 3rd, H. S. Phillips, K. Nikolics and P. H. Seeburg (1986). "A deletion truncating the gonadotropin-releasing hormone gene is responsible for hypogonadism in the hpg mouse." Science **234**(4782): 1366-1371.

Massari, M. E. and C. Murre (2000). "Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms." Mol Cell Biol **20**(2): 429-440.

Maston, G. A., S. K. Evans and M. R. Green (2006). "Transcriptional regulatory elements in the human genome." Annu Rev Genomics Hum Genet **7**: 29-59.

Matsuo, H., Y. Baba, R. M. Nair, A. Arimura and A. V. Schally (1971). "Structure of the porcine LH- and FSH-releasing hormone. I. The proposed amino acid sequence." Biochem Biophys Res Commun **43**(6): 1334-1339.

Mattick, J. S. (2005). "The functional genomics of noncoding RNA." Science **309**(5740): 1527-1528.
Mattick, J. S. and I. V. Makunin (2005). "Small regulatory RNAs in mammals." Hum Mol Genet **14 Spec No 1**: R121-132.

McCaffrey, A. P., L. Meuse, T. T. Pham, D. S. Conklin, G. J. Hannon and M. A. Kay (2002). "RNA interference in adult mice." Nature **418**(6893): 38-39.

McGaughey, D. M., Z. E. Stine, J. L. Huynh, R. M. Vinton and A. S. McCallion (2009). "Asymmetrical distribution of non-conserved regulatory sequences at PHOX2B is reflected at the ENCODE loci and illuminates a possible genome-wide trend." BMC Genomics **10**: 8.

McNeilly, A. S., C. J. Souza, D. T. Baird, I. A. Swanston, J. McVerry, J. Crawford, M. Cranfield and G. A. Lincoln (2002). "Production of inhibin A not B in rams: changes in plasma inhibin A during testis growth, and expression of inhibin/activin subunit mRNA and protein in adult testis." Reproduction **123**(6): 827-835.

Meachem, S. J., R. I. McLachlan, D. M. de Kretser, D. M. Robertson and N. G. Wreford (1996). "Neonatal exposure of rats to recombinant follicle stimulating hormone increases adult Sertoli and spermatogenic cell numbers." Biol Reprod **54**(1): 36-44.
Meachem, S. J., S. M. Ruwanpura, J. Ziolkowski, J. M. Ague, M. K. Skinner and K. L. Loveland (2005). "Developmentally distinct in vivo effects of FSH on proliferation and apoptosis during testis maturation." J Endocrinol **186**(3): 429-446.

Mello, C. C. and D. Conte, Jr. (2004). "Revealing the world of RNA interference." Nature **431**(7006): 338-342.

Melton, C., R. L. Judson and R. Blelloch (2010). "Opposing microRNA families regulate self-renewal in mouse embryonic stem cells." Nature **463**(7281): 621-626.

Merika, M. and D. Thanos (2001). "Enhanceosomes." Curr Opin Genet Dev **11**(2): 205-208.

Miele, A. and J. Dekker (2009). "Mapping cis- and trans- chromatin interaction networks using chromosome conformation capture (3C)." Methods Mol Biol **464**: 105-121.

Mikkelsen, T. S., M. Ku, D. B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T. K. Kim, R. P. Koche, W. Lee, E. Mendenhall, A. O'Donovan, A. Presser, C. Russ, X. Xie, A. Meissner, M. Wernig, R. Jaenisch, C. Nusbaum, E. S. Lander and B. E. Bernstein (2007). "Genome-wide maps of chromatin state in pluripotent and lineage-committed cells." Nature **448**(7153): 553-560.

Minegishi, T., M. Tano, K. Nakamura, S. Karino, K. Miyamoto and Y. Ibuki (1995). "Regulation of follicle-stimulating hormone receptor messenger ribonucleic acid levels in cultured rat granulosa cells." Mol Cell Endocrinol **108**(1-2): 67-73.

Moguilevsky, J. A., C. Libertun and V. G. Foglia (1970). "Metabolic sensitivity of different hypothalamic areas to luteinizing hormone (LH), follicle stimulating hormone (FSH), and testosterone." Neuroendocrinology **6**(3): 153-159.

Mohn, F. and D. Schubeler (2009). "Genetics and epigenetics: stability and plasticity during cellular differentiation." Trends Genet **25**(3): 129-136.

Monaco, L., N. S. Foulkes and P. Sassone-Corsi (1995). "Pituitary follicle-stimulating hormone (FSH) induces CREM gene expression in Sertoli cells: involvement in long-term desensitization of the FSH receptor." Proc Natl Acad Sci U S A **92**(23): 10673-10677.

Montgomery, G. W., M. L. Tate, H. M. Henry, J. M. Penty and R. M. Rohan (1995). "The follicle-stimulating hormone receptor and luteinizing hormone receptor genes are closely linked in sheep and deer." J Mol Endocrinol **15**(3): 259-265.

Moore, C. R. and D. Price (1930). "The Question of Sex Hormone Antagonism." Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.) **28**(1): 38-40.

Moore, C. R. and D. Price (1932). "Gonad hormone functions, and the reciprocal influence between gonads and hypophysis with its bearing on the problem of sex hormone antagonism." American Journal of Anatomy **50**(1): 13-71.

Morohashi, K. (1999). "Gonadal and Extragonadal Functions of Ad4BP/SF-1: Developmental Aspects." Trends Endocrinol Metab **10**(5): 169-173.

Morohashi, K. (1999). "Gonadal and Extragonadal Functions of Ad4BP/SF-1: Developmental Aspects." Trends in endocrinology and metabolism: TEM **10**(5): 169-173.

Morohashi, K., O. Hatano, M. Nomura, K. Takayama, M. Hara, H. Yoshii, A. Takakusu and T. Omura (1995). "Function and distribution of a steroidogenic cell-specific transcription factor, Ad4BP." J Steroid Biochem Mol Biol **53**(1-6): 81-88.

Morohashi, K., O. Hatano, M. Nomura, K. Takayama, M. Hara, H. Yoshii, A. Takakusu and T. Omura (1995). "Function and distribution of a steroidogenic cell-specific transcription factor, Ad4BP." The Journal of steroid biochemistry and molecular biology **53**(1-6): 81-88.

Morohashi, K., O. Hatano, M. Nomura, K. Takayama, M. Hara, H. Yoshii, A. Takakusu and T. Omura (1995). "Function and distribution of a steroidogenic cell-specific transcription factor, Ad4BP." J Steroid Biochem Mol Biol **53**(1-6): 81-88.
Morohashi, K., S. Honda, Y. Inomata, H. Handa and T. Omura (1992). "A common trans-acting factor, Ad4-binding protein, to the promoters of steroidogenic P-450s." J Biol Chem **267**(25): 17913-17919.
Morohashi, K., S. Honda, Y. Inomata, H. Handa and T. Omura (1992). "A common trans-acting factor, Ad4-binding protein, to the promoters of steroidogenic P-450s." The Journal of biological chemistry **267**(25): 17913-17919.

Morohashi, K., H. Iida, M. Nomura, O. Hatano, S. Honda, T. Tsukiyama, O. Niwa, T. Hara, A. Takakusu, Y. Shibata and et al. (1994). "Functional difference between Ad4BP and ELP, and their distributions in steroidogenic tissues." Molecular endocrinology **8**(5): 643-653.

Morohashi, K., H. Iida, M. Nomura, O. Hatano, S. Honda, T. Tsukiyama, O. Niwa, T. Hara, A. Takakusu, Y. Shibata and et al. (1994). "Functional difference between Ad4BP and ELP, and their distributions in steroidogenic tissues." Mol Endocrinol **8**(5): 643-653.

Morohashi, K., H. Tsuboi-Asai, S. Matsushita, M. Suda, M. Nakashima, H. Sasano, Y. Hataba, C. L. Li, J. Fukata, J. Irie, T. Watanabe, H. Nagura and E. Li (1999). "Structural and functional abnormalities in the spleen of an mFtz-F1 gene-disrupted mouse." Blood **93**(5): 1586-1594.

Morohashi, K. I. and T. Omura (1996). "Ad4BP/SF-1, a transcription factor essential for the transcription of steroidogenic cytochrome P450 genes and for the establishment of the reproductive function." FASEB J **10**(14): 1569-1577.

Mullen, A. C., D. A. Orlando, J. J. Newman, J. Loven, R. M. Kumar, S. Bilodeau, J. Reddy, M. G. Guenther, R. P. DeKoter and R. A. Young (2011). "Master transcription factors determine cell-type-specific responses to TGF-beta signaling." Cell **147**(3): 565-576.

Mural, R. J., M. D. Adams, E. W. Myers, H. O. Smith, G. L. Miklos, R. Wides, A. Halpern, P. W. Li, G. G. Sutton, J. Nadeau, S. L. Salzberg, R. A. Holt, C. D. Kodira, F. Lu, L. Chen, Z. Deng, C. C. Evangelista, W. Gan, T. J. Heiman, J. Li, Z. Li, G. V. Merkulov, N. V. Milshina, A. K. Naik, R. Qi, B. C. Shue, A. Wang, J. Wang, X. Wang, X. Yan, J. Ye, S. Yooseph, Q. Zhao, L. Zheng, S. C. Zhu, K. Biddick, R. Bolanos, A. L. Delcher, I. M. Dew, D. Fasulo, M. J. Flanigan, D. H. Huson, S. A. Kravitz, J. R. Miller, C. M. Mobarry, K. Reinert, K. A. Remington, Q. Zhang, X. H. Zheng, D. R. Nusskern, Z. Lai, Y. Lei, W. Zhong, A. Yao, P. Guan, R. R. Ji, Z. Gu, Z. Y. Wang, F. Zhong, C. Xiao, C. C. Chiang, M. Yandell, J. R. Wortman, P. G. Amanatides, S. L. Hladun, E. C. Pratts, J. E. Johnson, K. L. Dodson, K. J. Woodford, C. A. Evans, B. Gropman, D. B. Rusch, E. Venter, M. Wang, T. J. Smith, J. T. Houck, D. E. Tompkins, C. Haynes, D. Jacob, S. H. Chin, D. R. Allen, C. E. Dahlke, R. Sanders, K. Li, X. Liu, A. A. Levitsky, W. H. Majoros, Q. Chen, A. C. Xia, J. R. Lopez, M. T. Donnelly, M. H. Newman, A. Glodek, C. L. Kraft, M. Nodell, F. Ali, H. J. An, D. Baldwin-Pitts, K. Y. Beeson, S. Cai, M. Carnes, A. Carver, P. M. Caulk, A. Center, Y. H. Chen, M. L. Cheng, M. D. Coyne, M. Crowder, S. Danaher, L. B. Davenport, R. Desilets, S. M. Dietz, L. Doup, P. Dullaghan, S. Ferriera, C. R. Fosler, H. C. Gire, A. Gluecksmann, J. D. Gocayne, J. Gray, B. Hart, J. Haynes, J. Hoover, T. Howland, C. Ibegwam, M. Jalali, D. Johns, L. Kline, D. S. Ma, S. MacCawley, A. Magoon, F. Mann, D. May, T. C. McIntosh, S. Mehta, L. Moy, M. C. Moy, B. J. Murphy, S. D. Murphy, K. A. Nelson, Z. Nuri, K. A. Parker, A. C. Prudhomme, V. N. Puri, H. Qureshi, J. C. Raley, M. S. Reardon, M. A. Regier, Y. H. Rogers, D. L. Romblad, J. Schutz, J. L. Scott, R. Scott, C. D. Sitter, M. Smallwood, A. C. Sprague, E. Stewart, R. V. Strong, E. Suh, K. Sylvester, R. Thomas, N. N. Tint, C. Tsonis, G. Wang, M. S. Williams, S. M. Williams, S. M. Windsor, K. Wolfe, M. M. Wu, J. Zaveri, K. Chaturvedi, A. E. Gabrielian, Z. Ke, J. Sun, G. Subramanian, J. C. Venter, C. M. Pfannkoch,

M. Barnstead and L. D. Stephenson (2002). "A comparison of whole-genome shotgun-derived mouse chromosome 16 and the human genome." Science **296**(5573): 1661-1671.

Nagano, T., J. A. Mitchell, L. A. Sanz, F. M. Pauler, A. C. Ferguson-Smith, R. Feil and P. Fraser (2008). "The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin." Science **322**(5908): 1717-1720.

Nakamura, M., T. Minegishi, Y. Hasegawa, K. Nakamura, S. Igarashi, I. Ito, H. Shinozaki, K. Miyamoto, Y. Eto and Y. Ibuki (1993). "Effect of an Activin-a on Follicle-Stimulating-Hormone (Fsh) Receptor Messenger-Ribonucleic-Acid Levels and Fsh Receptor Expressions in Cultured Rat Granulosa-Cells." Endocrinology **133**(2): 538-544.

Nakatani, A., S. Shimasaki, L. V. Depaolo, G. F. Erickson and N. Ling (1991). "Cyclic changes in follistatin messenger ribonucleic acid and its protein in the rat ovary during the estrous cycle." Endocrinology **129**(2): 603-611.

Nakaya, H. I., P. P. Amaral, R. Louro, A. Lopes, A. A. Fachel, Y. B. Moreira, T. A. El-Jundi, A. M. da Silva, E. M. Reis and S. Verjovski-Almeida (2007). "Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription." Genome Biol **8**(3): R43.

Nativio, R., K. S. Wendt, Y. Ito, J. E. Huddleston, S. Uribe-Lewis, K. Woodfine, C. Krueger, W. Reik, J. M. Peters and A. Murrell (2009). "Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus." PLoS Genet **5**(11): e1000739.

Ngan, E. S., P. K. Cheng, P. C. Leung and B. K. Chow (1999). "Steroidogenic factor-1 interacts with a gonadotrope-specific element within the first exon of the human gonadotropin-releasing hormone receptor gene to mediate gonadotrope-specific expression." Endocrinology **140**(6): 2452-2462.

Nimrod, A., G. F. Erickson and K. J. Ryan (1976). "A specific FSH receptor in rat granulosa cells: properties of binding in vitro." Endocrinology **98**(1): 56-64.

Ninomiya, Y., M. Okada, N. Kotomura, K. Suzuki, T. Tsukiyama and O. Niwa (1995). "Genomic organization and isoforms of the mouse ELP gene." J Biochem **118**(2): 380-389.

Nobrega, M. A., I. Ovcharenko, V. Afzal and E. M. Rubin (2003). "Scanning human gene deserts for long-range enhancers." Science **302**(5644): 413.

Nomura, M., S. Bartsch, H. Nawata, T. Omura and K. Morohashi (1995). "An E box element is required for the expression of the ad4bp gene, a mammalian homologue of ftz-f1 gene, which is essential for adrenal and gonadal development." J Biol Chem **270**(13): 7453-7461.

Nordhoff, V., J. Gromoll, L. Foppiani, C. M. Luetjens, S. Schlatt, E. Kostova, I. Huhtaniemi, E. Nieschlag and M. Simoni (2003). "Targeted expression of human FSH receptor Asp567Gly mutant mRNA in testis of transgenic mice: role of human FSH receptor promoter." Asian J Androl **5**(4): 267-275.

Nothnick, W. B. (2012). "The role of micro-RNAs in the female reproductive tract." Reproduction **143**(5): 559-576.

Numata, K., A. Kanai, R. Saito, S. Kondo, J. Adachi, L. G. Wilming, D. A. Hume, Y. Hayashizaki and M. Tomita (2003). "Identification of putative noncoding RNAs among the RIKEN mouse full-length cDNA collection." Genome Res **13**(6B): 1301-1306.

O'Shaughnessy, P. J., P. Marsh and K. Dudley (1994). "Follicle-stimulating hormone receptor mRNA in the mouse ovary during post-natal development in the normal mouse and in the adult hypogonadal (hpg) mouse: structure of alternate transcripts." Mol Cell Endocrinol **101**(1-2): 197-201.

Ogbourne, S. and T. M. Antalis (1998). "Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes." Biochem J **331 (Pt 1)**: 1-14.

Ohhata, T., Y. Hoki, H. Sasaki and T. Sado (2008). "Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification." Development **135**(2): 227-235.

Ohlsson, R., M. Bartkuhn and R. Renkawitz (2010). "CTCF shapes chromatin by multiple mechanisms: the impact of 20 years of CTCF research on understanding the workings of chromatin." Chromosoma **119**(4): 351-360.

Okazaki, Y., M. Furuno, T. Kasukawa, J. Adachi, H. Bono, S. Kondo, I. Nikaido, N. Osato, R. Saito, H. Suzuki, I. Yamanaka, H. Kiyosawa, K. Yagi, Y. Tomaru, Y. Hasegawa, A. Nogami, C. Schonbach, T. Gojobori, R. Baldarelli, D. P. Hill, C. Bult, D. A. Hume, J. Quackenbush, L. M. Schriml, A. Kanapin, H. Matsuda, S. Batalov, K. W. Beisel, J. A. Blake, D. Bradt, V. Brusic, C. Chothia, L. E. Corbani, S. Cousins, E. Dalla, T. A. Dragani, C. F. Fletcher, A. Forrest, K. S. Frazer, T. Gaasterland, M. Gariboldi, C. Gissi, A. Godzik, J. Gough, S. Grimmond, S. Gustincich, N. Hirokawa, I. J. Jackson, E. D. Jarvis, A. Kanai, H. Kawaji, Y. Kawasawa, R. M. Kedzierski, B. L. King, A. Konagaya, I. V. Kurochkin, Y. Lee, B. Lenhard, P. A. Lyons, D. R. Maglott, L. Maltais, L. Marchionni, L. McKenzie, H. Miki, T. Nagashima, K. Numata, T. Okido, W. J. Pavan, G. Pertea, G. Pesole, N. Petrovsky, R. Pillai, J. U. Pontius, D. Qi, S. Ramachandran, T. Ravasi, J. C. Reed, D. J. Reed, J. Reid, B. Z. Ring, M. Ringwald, A. Sandelin, C. Schneider, C. A. Semple, M. Setou, K. Shimada, R. Sultana, Y. Takenaka, M. S. Taylor, R. D. Teasdale, M. Tomita, R. Verardo, L. Wagner, C. Wahlestedt, Y. Wang, Y. Watanabe, C. Wells, L. G. Wilming, A. Wynshaw-Boris, M. Yanagisawa, I. Yang, L. Yang, Z. Yuan, M. Zavolan, Y. Zhu, A. Zimmer, P. Carninci, N. Hayatsu, T. Hirozane-Kishikawa, H. Konno, M. Nakamura, N. Sakazume, K. Sato, T. Shiraki, K. Waki, J. Kawai, K. Aizawa, T. Arakawa, S. Fukuda, A. Hara, W. Hashizume, K. Imotani, Y. Ishii, M. Itoh, I. Kagawa, A. Miyazaki, K. Sakai, D. Sasaki, K. Shibata, A. Shinagawa, A. Yasunishi, M. Yoshino, R. Waterston, E. S. Lander, J. Rogers, E. Birney and Y. Hayashizaki (2002). "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs." Nature **420**(6915): 563-573.

Orom, U. A., T. Derrien, M. Beringer, K. Gumireddy, A. Gardini, G. Bussotti, F. Lai, M. Zytnicki, C. Notredame, Q. Huang, R. Guigo and R. Shiekhattar (2010). "Long noncoding RNAs with enhancer-like function in human cells." Cell **143**(1): 46-58.

Orth, J. and A. K. Christensen (1977). "Localization of 125I-labeled FSH in the testes of hypophysectomized rats by autoradiography at the light and electron microscope levels." Endocrinology **101**(1): 262-278.

Orth, J. M. (1984). "The role of follicle-stimulating hormone in controlling Sertoli cell proliferation in testes of fetal rats." Endocrinology **115**(4): 1248-1255.

Orth, J. M., G. L. Gunsalus and A. A. Lamperti (1988). "Evidence from Sertoli cell-depleted rats indicates that spermatid number in adults depends on numbers of Sertoli cells produced during perinatal development." Endocrinology **122**(3): 787-794.

Ovcharenko, I., G. G. Loots, M. A. Nobrega, R. C. Hardison, W. Miller and L. Stubbs (2005). "Evolution and functional classification of vertebrate gene deserts." Genome Res **15**(1): 137-145.

Ovcharenko, I., M. A. Nobrega, G. G. Loots and L. Stubbs (2004). "ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes." Nucleic Acids Res **32**(Web Server issue): W280-286.

Palstra, R. J., B. Tolhuis, E. Splinter, R. Nijmeijer, F. Grosveld and W. de Laat (2003). "The beta-globin nuclear compartment in development and erythroid differentiation." Nat Genet **35**(2): 190-194.

Pan, L., S. Sato, J. P. Frederick, X. H. Sun and Y. Zhuang (1999). "Impaired immune responses and B-cell proliferation in mice lacking the Id3 gene." Mol Cell Biol **19**(9): 5969-5980.

Pandey, R. R., T. Mondal, F. Mohammad, S. Enroth, L. Redrup, J. Komorowski, T. Nagano, D. Mancini-Dinardo and C. Kanduri (2008). "Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation." Mol Cell **32**(2): 232-246.

Parelho, V., S. Hadjur, M. Spivakov, M. Leleu, S. Sauer, H. C. Gregson, A. Jarmuz, C. Canzonetta, Z. Webster, T. Nesterova, B. S. Cobb, K. Yokomori, N. Dillon, L. Aragon, A. G. Fisher and M. Merkenschlager (2008). "Cohesins functionally associate with CTCF on mammalian chromosome arms." Cell **132**(3): 422-433.

Parker, K. L. and B. P. Schimmer (1997). "Steroidogenic factor 1: a key determinant of endocrine development and function." Endocr Rev **18**(3): 361-377.

Parker, K. L. and B. P. Schimmer (1997). "Steroidogenic factor 1: a key determinant of endocrine development and function." Endocrine reviews **18**(3): 361-377.

Pasquinelli, A. E. (2012). "MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship." Nat Rev Genet **13**(4): 271-282.

Pasquinelli, A. E. and G. Ruvkun (2002). "Control of developmental timing by micrornas and their targets." Annu Rev Cell Dev Biol **18**: 495-513.

Pelusi, C., Y. Ikeda, M. Zubair and K. L. Parker (2008). "Impaired Follicle Development and Infertility in Female Mice Lacking Steroidogenic Factor 1 in Ovarian Granulosa Cells." Biology of Reproduction **79**(6): 1074-1083.

Peluso, J. J. and R. W. Steger (1978). "Role of FSH in regulating granulosa cell division and follicular atresia in rats." J Reprod Fertil **54**(2): 275-278.

Peng, X. R., A. J. Hsueh, P. S. LaPolt, L. Bjersing and T. Ny (1991). "Localization of luteinizing hormone receptor messenger ribonucleic acid expression in ovarian cell types during follicle development and ovulation." Endocrinology **129**(6): 3200-3207.

Pennacchio, L. A. (2003). "Insights from human/mouse genome comparisons." Mamm Genome **14**(7): 429-436.

Pennacchio, L. A., N. Ahituv, A. M. Moses, S. Prabhakar, M. A. Nobrega, M. Shoukry, S. Minovitsky, I. Dubchak, A. Holt, K. D. Lewis, I. Plajzer-Frick, J. Akiyama, S. De Val, V. Afzal, B. L. Black, O. Couronne, M. B. Eisen, A. Visel and E. M. Rubin (2006). "In vivo enhancer analysis of human conserved non-coding sequences." Nature **444**(7118): 499-502.

Perk, J., A. Iavarone and R. Benezra (2005). "Id family of helix-loop-helix proteins in cancer." Nat Rev Cancer **5**(8): 603-614.

Phillips, J. E. and V. G. Corces (2009). "CTCF: master weaver of the genome." Cell **137**(7): 1194-1211.
Pikaart, M. J., F. Recillas-Targa and G. Felsenfeld (1998). "Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators." Genes Dev **12**(18): 2852-2862.

Pillai, R. S. (2005). "MicroRNA function: multiple mechanisms for a tiny RNA?" RNA **11**(12): 1753-1761.
Pink, R. C., K. Wicks, D. P. Caley, E. K. Punch, L. Jacobs and D. R. Carter (2011). "Pseudogenes: pseudo-functional or key regulators in health and disease?" RNA **17**(5): 792-798.

Plant, T. M. (2008). "Hypothalamic control of the pituitary-gonadal axis in higher primates: key advances over the last two decades." J Neuroendocrinol **20**(6): 719-726.

Poliseno, L., L. Salmena, J. Zhang, B. Carver, W. J. Haveman and P. P. Pandolfi (2010). "A coding-independent function of gene and pseudogene mRNAs regulates tumour biology." Nature **465**(7301): 1033-1038.

Pollard, K. S., M. J. Hubisz, K. R. Rosenbloom and A. Siepel (2010). "Detection of nonneutral substitution rates on mammalian phylogenies." Genome Res **20**(1): 110-121.

Ponjavic, J., C. P. Ponting and G. Lunter (2007). "Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs." Genome Res **17**(5): 556-565.

Ponting, C. P., P. L. Oliver and W. Reik (2009). "Evolution and functions of long noncoding RNAs." Cell **136**(4): 629-641.

Prabhakar, S., F. Poulin, M. Shoukry, V. Afzal, E. M. Rubin, O. Couronne and L. A. Pennacchio (2006). "Close sequence comparisons are sufficient to identify human cis-regulatory elements." Genome Res **16**(7): 855-863.

Prior, J. C. (2007). "FSH and bone--important physiology or not?" Trends Mol Med **13**(1): 1-3.
Privalsky, M. L. (2004). "The role of corepressors in transcriptional regulation by nuclear hormone receptors." Annu Rev Physiol **66**: 315-360.

Putowski, L. T., W. J. Schillings, C. M. Lee, E. P. Reddy and J. A. Jakowicki (2004). "Human follicle-stimulating hormone receptor (FSH-R) promoter/enhancer activity is inhibited by transcriptional factors, from the upstream stimulating factors family, via E-box and newly identified initiator element (Inr) in FSH-R non-expressing cells." Gynecol Endocrinol **19**(1): 9-17.

Rajasethupathy, P., I. Antonov, R. Sheridan, S. Frey, C. Sander, T. Tuschl and E. R. Kandel (2012). "A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity." Cell **149**(3): 693-707.

Ramayya, M. S., J. Zhou, T. Kino, J. H. Segars, C. A. Bondy and G. P. Chrousos (1997). "Steroidogenic factor 1 messenger ribonucleic acid expression in steroidogenic and nonsteroidogenic human tissues: Northern blot and in situ hybridization studies." The Journal of clinical endocrinology and metabolism **82**(6): 1799-1806.

Rannikki, A. S., F. P. Zhang and I. T. Huhtaniemi (1995). "Ontogeny of follicle-stimulating hormone receptor gene expression in the rat testis and ovary." Mol Cell Endocrinol **107**(2): 199-208.

Rannikko, A., T. L. Penttila, F. P. Zhang, J. Toppari, M. Parvinen and I. Huhtaniemi (1996). "Stage-specific expression of the FSH receptor gene in the prepubertal and adult rat seminiferous epithelium." J Endocrinol **151**(1): 29-35.

Ravasi, T., H. Suzuki, K. C. Pang, S. Katayama, M. Furuno, R. Okunishi, S. Fukuda, K. Ru, M. C. Frith, M. M. Gongora, S. M. Grimmond, D. A. Hume, Y. Hayashizaki and J. S. Mattick (2006). "Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome." Genome Res **16**(1): 11-19.

Recillas-Targa, F., M. J. Pikaart, B. Burgess-Beusse, A. C. Bell, M. D. Litt, A. G. West, M. Gaszner and G. Felsenfeld (2002). "Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities." Proc Natl Acad Sci U S A **99**(10): 6883-6888.

Ren, B., F. Robert, J. J. Wyrick, O. Aparicio, E. G. Jennings, I. Simon, J. Zeitlinger, J. Schreiber, N. Hannett, E. Kanin, T. L. Volkert, C. J. Wilson, S. P. Bell and R. A. Young (2000). "Genome-wide location and function of DNA binding proteins." Science **290**(5500): 2306-2309.

Renaud, S., D. Loukinov, F. T. Bosman, V. Lobanenkov and J. Benhattar (2005). "CTCF binds the proximal exonic region of hTERT and inhibits its transcription." Nucleic Acids Res **33**(21): 6850-6860.

Rhead, B., D. Karolchik, R. M. Kuhn, A. S. Hinrichs, A. S. Zweig, P. A. Fujita, M. Diekhans, K. E. Smith, K. R. Rosenbloom, B. J. Raney, A. Pohl, M. Pheasant, L. R. Meyer, K. Learned, F. Hsu, J. Hillman-Jackson, R. A. Harte, B. Giardine, T. R. Dreszer, H. Clawson, G. P. Barber, D. Haussler and W. J. Kent (2010). "The UCSC Genome Browser database: update 2010." Nucleic Acids Res **38**(Database issue): D613-619.

Rice, D. A., A. R. Mouw, A. M. Bogerd and K. L. Parker (1991). "A shared promoter element regulates the expression of three steroidogenic enzymes." Molecular endocrinology **5**(10): 1552-1561.

Rice, D. A., A. R. Mouw, A. M. Bogerd and K. L. Parker (1991). "A shared promoter element regulates the expression of three steroidogenic enzymes." Mol Endocrinol **5**(10): 1552-1561.

Richards, J. S., J. J. Ireland, M. C. Rao, G. A. Bernath, A. R. Midgley, Jr. and L. E. Reichert, Jr. (1976). "Ovarian follicular development in the rat: hormone receptor regulation by estradiol, follicle stimulating hormone and luteinizing hormone." Endocrinology **99**(6): 1562-1570.

Richards, J. S. and A. R. Midgley, Jr. (1976). "Protein hormone action: a key to understanding ovarian follicular and luteal cell development." Biol Reprod **14**(1): 82-94.

Rinn, J. L., M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H. Goodnough, J. A. Helms, P. J. Farnham, E. Segal and H. Y. Chang (2007). "Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs." Cell **129**(7): 1311-1323.

Ritter, V., B. Thuering, P. Saint Mezard, N. H. Luong-Nguyen, Y. Seltenmeyer, U. Junker, B. Fournier, M. Susa and F. Morvan (2008). "Follicle-stimulating hormone does not impact male bone mass in vivo or human male osteoclasts in vitro." Calcif Tissue Int **82**(5): 383-391.

Rizzolio, F., S. Bione, C. Sala, C. Tribioli, R. Ciccone, O. Zuffardi, N. di Iorgi, M. Maghnie and D. Toniolo (2008). "Highly conserved non-coding sequences and the 18q critical region for short stature: a common mechanism of disease?" PLoS One **3**(1): e1460.

Roberts, V. J., R. Steenbergen and C. Murre (1993). "Localization of E2A mRNA expression in developing and adult rat tissues." Proc Natl Acad Sci U S A **90**(16): 7583-7587.

Robinson, L. J., I. Tourkova, Y. Wang, A. C. Sharrow, M. S. Landau, B. B. Yaroslavskiy, S. Li, M. Zaidi and H. C. Blair (2010). "FSH-Receptor Isoforms and FSH-dependent Gene Transcription in Human Monocytes and Osteoclasts." Biochem Biophys Res Commun.
Rodriguez, C. I., N. Girones and M. Fresno (2003). "Cha, a basic helix-loop-helix transcription factor involved in the regulation of upstream stimulatory factor activity." J Biol Chem **278**(44): 43135-43145.

Roh, T. Y., G. Wei, C. M. Farrell and K. Zhao (2007). "Genome-wide prediction of conserved and nonconserved enhancers by histone acetylation patterns." Genome Res **17**(1): 74-81.

Rousseau-Merck, M. F., M. Atger, H. Loosfelt, E. Milgrom and R. Berger (1993). "The chromosomal localization of the human follicle-stimulating hormone receptor gene (FSHR) on 2p21-p16 is similar to that of the luteinizing hormone receptor gene." Genomics **15**(1): 222-224.

Russell, L. D. and M. D. Griswold (1993). The Sertoli cell. Clearwater, FL, Cache River Press.

Ruwanpura, S. M., R. I. McLachlan, P. G. Stanton and S. J. Meachem (2008). "Follicle-stimulating hormone affects spermatogonial survival by regulating the intrinsic apoptotic pathway in adult rats." Biol Reprod **78**(4): 705-713.

Ruzinova, M. B. and R. Benezra (2003). "Id proteins in development, cell cycle and cancer." Trends Cell Biol **13**(8): 410-418.

Sabo, P. J., R. Humbert, M. Hawrylycz, J. C. Wallace, M. O. Dorschner, M. McArthur and J. A. Stamatoyannopoulos (2004). "Genome-wide identification of DNaseI hypersensitive sites using active chromatin sequence libraries." Proc Natl Acad Sci U S A **101**(13): 4537-4542.

Sado, T., Y. Hoki and H. Sasaki (2005). "Tsix silences Xist through modification of chromatin structure." Dev Cell **9**(1): 159-165.

Sadovsky, Y. and P. A. Crawford (1998). "Developmental and physiologic roles of the nuclear receptor steroidogenic factor-1 in the reproductive system." J Soc Gynecol Investig **5**(1): 6-12.

Sadovsky, Y., P. A. Crawford, K. G. Woodson, J. A. Polish, M. A. Clements, L. M. Tourtellotte, K. Simburger and J. Milbrandt (1995). "Mice deficient in the orphan receptor steroidogenic factor 1 lack adrenal glands and gonads but express P450 side-chain-cleavage enzyme in the placenta and have normal embryonic serum levels of corticosteroids." Proceedings of the National Academy of Sciences of the United States of America **92**(24): 10939-10943.

Sadovsky, Y., P. A. Crawford, K. G. Woodson, J. A. Polish, M. A. Clements, L. M. Tourtellotte, K. Simburger and J. Milbrandt (1995). "Mice deficient in the orphan receptor steroidogenic factor 1 lack

adrenal glands and gonads but express P450 side-chain-cleavage enzyme in the placenta and have normal embryonic serum levels of corticosteroids." Proc Natl Acad Sci U S A **92**(24): 10939-10943.

Sairam, M. R., L. G. Jiang, T. A. Yarney and H. Khan (1996). "Follitropin signal transduction: alternative splicing of the FSH receptor gene produces a dominant negative form of receptor which inhibits hormone action." Biochem Biophys Res Commun **226**(3): 717-722.

Sairam, M. R. and V. S. Subbarayan (1997). "Characterization of the 5' flanking region and potential control elements of the ovine follitropin receptor gene." Mol Reprod Dev **48**(4): 480-487.

Sambrook, J. and D. W. Russell (2006). "Selection of Poly(A)+ RNA by Oligo(dT)-Cellulose Chromatography." CSH Protoc **2006**(1).

Sandmann, T., C. Girardot, M. Brehme, W. Tongprasit, V. Stolc and E. E. Furlong (2007). "A core transcriptional network for early mesoderm development in Drosophila melanogaster." Genes Dev **21**(4): 436-449.

Sandmann, T., L. J. Jensen, J. S. Jakobsen, M. M. Karzynski, M. P. Eichenlaub, P. Bork and E. E. Furlong (2006). "A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development." Dev Cell **10**(6): 797-807.

Sanford, J. C. and B. E. Batten (1989). "Endocytosis of follicle-stimulating hormone by ovarian granulosa cells: analysis of hormone processing and receptor dynamics." J Cell Physiol **138**(1): 154-164.

Scherrer, S. P., D. A. Rice and L. L. Heckert (2002). "Expression of steroidogenic factor 1 in the testis requires an interactive array of elements within its proximal promoter." Biol Reprod **67**(5): 1509-1521.

Schimmer, B. P. and P. C. White (2010). "Minireview: steroidogenic factor 1: its roles in differentiation, development, and disease." Mol Endocrinol **24**(7): 1322-1337.
Schneider, R. and R. Grosschedl (2007). "Dynamics and interplay of nuclear architecture, genome organization, and gene expression." Genes Dev **21**(23): 3027-3043.

Schoenborn, J. R., M. O. Dorschner, M. Sekimata, D. M. Santer, M. Shnyreva, D. R. Fitzpatrick, J. A. Stamatoyannopoulos and C. B. Wilson (2007). "Comprehensive epigenetic profiling identifies multiple distal regulatory elements directing transcription of the gene encoding interferon-gamma." Nat Immunol **8**(7): 732-742.

Schwartz, N. B. (1982). "Role of ovarian inhibin (folliculostatin) in regulating FSH secretion in the female rat." Adv Exp Med Biol **147**: 15-36.

Scobey, M. J., C. A. Fix and W. H. Walker (2004). "The Id2 transcriptional repressor is induced by follicle-stimulating hormone and cAMP." J Biol Chem **279**(16): 16064-16070.

Scott, M. P. and A. J. Weiner (1984). "Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila." Proc Natl Acad Sci U S A **81**(13): 4115-4119.

Sealfon, S. C., H. Weinstein and R. P. Millar (1997). "Molecular mechanisms of ligand interaction with the gonadotropin-releasing hormone receptor." Endocr Rev **18**(2): 180-205.

Seger, R., T. Hanoch, R. Rosenberg, A. Dantes, W. E. Merz, J. F. Strauss, 3rd and A. Amsterdam (2001). "The ERK signaling cascade inhibits gonadotropin-stimulated steroidogenesis." J Biol Chem **276**(17): 13957-13964.

Seitz, H., H. Royo, M. L. Bortolin, S. P. Lin, A. C. Ferguson-Smith and J. Cavaille (2004). "A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain." Genome Res **14**(9): 1741-1748.

Seitz, H., N. Youngson, S. P. Lin, S. Dalbert, M. Paulsen, J. P. Bachellerie, A. C. Ferguson-Smith and J. Cavaille (2003). "Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene." Nat Genet **34**(3): 261-262.

Senger, K., G. W. Armstrong, W. J. Rowell, J. M. Kwan, M. Markstein and M. Levine (2004). "Immunity regulatory DNAs share common organizational features in Drosophila." Mol Cell **13**(1): 19-32.

Shahmanesh, M., M. Sedigh, B. Azedeh, M. H. Sheikholeslami and N. K. Nair (1980). "Feedback control of FSH secretion in the male rat." Horm Res **12**(5): 266-276.

Shetty, J., G. K. Marathe and R. R. Dighe (1996). "Specific immunoneutralization of FSH leads to apoptotic cell death of the pachytene spermatocytes and spermatogonial cells in the rat." Endocrinology **137**(5): 2179-2182.

Shieh, B. H., R. S. Sparkes, R. B. Gaynor and A. J. Lusis (1993). "Localization of the gene-encoding upstream stimulatory factor (USF) to human chromosome 1q22-q23." Genomics **16**(1): 266-268.

Shima, Y., K. Miyabayashi, T. Baba, H. Otake, Y. Katsura, S. Oka, M. Zubair and K. Morohashi (2012). "Identification of an enhancer in the Ad4BP/SF-1 gene specific for fetal Leydig cells." Endocrinology **153**(1): 417-425.

Shima, Y., M. Zubair, S. Ishihara, Y. Shinohara, S. Oka, S. Kimura, S. Okamoto, Y. Minokoshi, S. Suita and K. Morohashi (2005). "Ventromedial hypothalamic nucleus-specific enhancer of Ad4BP/SF-1 gene." Mol Endocrinol **19**(11): 2812-2823.

Shima, Y., M. Zubair, T. Komatsu, S. Oka, C. Yokoyama, T. Tachibana, T. A. Hjalt, J. Drouin and K. Morohashi (2008). "Pituitary homeobox 2 regulates adrenal4 binding protein/steroidogenic factor-1 gene transcription in the pituitary gonadotrope through interaction with the intronic enhancer." Mol Endocrinol **22**(7): 1633-1646.

Shimizu, A., K. Tsutsui and S. Kawashima (1987). "Autoradiographic study of binding and internalization of follicle-stimulating hormone in the mouse testis minces in vitro." Endocrinol Jpn **34**(3): 431-442.

Shin, J. T., J. R. Priest, I. Ovcharenko, A. Ronco, R. K. Moore, C. G. Burns and C. A. MacRae (2005). "Human-zebrafish non-coding conserved elements act in vivo to regulate transcription." Nucleic Acids Res **33**(17): 5437-5445.

Shinoda, K., H. Lei, H. Yoshii, M. Nomura, M. Nagano, H. Shiba, H. Sasaki, Y. Osawa, Y. Ninomiya, O. Niwa and et al. (1995). "Developmental defects of the ventromedial hypothalamic nucleus and pituitary gonadotroph in the Ftz-F1 disrupted mice." Dev Dyn **204**(1): 22-29.

Shinoda, K., H. Lei, H. Yoshii, M. Nomura, M. Nagano, H. Shiba, H. Sasaki, Y. Osawa, Y. Ninomiya, O. Niwa and et al. (1995). "Developmental defects of the ventromedial hypothalamic nucleus and pituitary

gonadotroph in the Ftz-F1 disrupted mice." Developmental dynamics : an official publication of the American Association of Anatomists **204**(1): 22-29.

Shoemaker, D. D., E. E. Schadt, C. D. Armour, Y. D. He, P. Garrett-Engele, P. D. McDonagh, P. M. Loerch, A. Leonardson, P. Y. Lum, G. Cavet, L. F. Wu, S. J. Altschuler, S. Edwards, J. King, J. S. Tsang, G. Schimmack, J. M. Schelter, J. Koch, M. Ziman, M. J. Marton, B. Li, P. Cundiff, T. Ward, J. Castle, M. Krolewski, M. R. Meyer, M. Mao, J. Burchard, M. J. Kidd, H. Dai, J. W. Phillips, P. S. Linsley, R. Stoughton, S. Scherer and M. S. Boguski (2001). "Experimental annotation of the human genome using microarray technology." Nature **409**(6822): 922-927.

Siepel, A., G. Bejerano, J. S. Pedersen, A. S. Hinrichs, M. Hou, K. Rosenbloom, H. Clawson, J. Spieth, L. W. Hillier, S. Richards, G. M. Weinstock, R. K. Wilson, R. A. Gibbs, W. J. Kent, W. Miller and D. Haussler (2005). "Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes." Genome Res **15**(8): 1034-1050.

Simoni, M., J. Gromoll, W. Hoppner and E. Nieschlag (1997). "Molecular pathophysiology of the pituitary-gonadal axis." Adv Exp Med Biol **424**: 89-97.

Simoni, M., J. Gromoll and E. Nieschlag (1997). "The follicle-stimulating hormone receptor: biochemistry, molecular biology, physiology, and pathophysiology." Endocr Rev **18**(6): 739-773.

Siomi, M. C., K. Sato, D. Pezic and A. A. Aravin (2011). "PIWI-interacting small RNAs: the vanguard of genome defence." Nat Rev Mol Cell Biol **12**(4): 246-258.

Sites, C. K., K. Patterson, C. S. Jamison, S. J. F. Degen and A. R. Labarbera (1994). "Follicle-Stimulating-Hormone (Fsh) Increases Fsh Receptor Messenger-Ribonucleic-Acid While Decreasing Fsh Binding in Cultured Porcine Granulosa-Cells." Endocrinology **134**(1): 411-417.

Smale, S. T. and J. T. Kadonaga (2003). "The RNA polymerase II core promoter." Annu Rev Biochem **72**: 449-479.

Sokka, T. and I. Huhtaniemi (1990). "Ontogeny of gonadotrophin receptors and gonadotrophin-stimulated cyclic AMP production in the neonatal rat ovary." J Endocrinol **127**(2): 297-303.

Sparago, A., F. Cerrato, M. Vernucci, G. B. Ferrero, M. C. Silengo and A. Riccio (2004). "Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith-Wiedemann syndrome." Nat Genet **36**(9): 958-960.

Spicuglia, S., D. Payet, R. K. Tripathi, P. Rameil, C. Verthuy, J. Imbert, P. Ferrier and W. M. Hempel (2000). "TCRalpha enhancer activation occurs via a conformational change of a pre-assembled nucleo-protein complex." EMBO J **19**(9): 2034-2045.

Spilianakis, C. G., M. D. Lalioti, T. Town, G. R. Lee and R. A. Flavell (2005). "Interchromosomal associations between alternatively expressed loci." Nature **435**(7042): 637-645.

Spitz, F. and E. E. Furlong (2012). "Transcription factors: from enhancer binding to developmental control." Nat Rev Genet **13**(9): 613-626.

Splinter, E., H. Heath, J. Kooren, R. J. Palstra, P. Klous, F. Grosveld, N. Galjart and W. de Laat (2006). "CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus." Genes Dev **20**(17): 2349-2354.

Sprengel, R., T. Braun, K. Nikolics, D. L. Segaloff and P. H. Seeburg (1990). "The testicular receptor for follicle stimulating hormone: structure and functional expression of cloned cDNA." Mol Endocrinol **4**(4): 525-530.

Srinivasan, L. and M. L. Atchison (2004). "YY1 DNA binding and PcG recruitment requires CtBP." Genes Dev **18**(21): 2596-2601.

Stallings, N. R., N. A. Hanley, G. Majdic, L. Zhao, M. Bakke and K. L. Parker (2002). "Development of a transgenic green fluorescent protein lineage marker for steroidogenic factor 1." Endocr Res **28**(4): 497-504.

Stallings, N. R., N. A. Hanley, G. Majdic, L. Zhao, M. Bakke and K. L. Parker (2002). "Development of a transgenic green fluorescent protein lineage marker for steroidogenic factor 1." Mol Endocrinol **16**(10): 2360-2370.

Stanislaus, D., J. A. Janovick, S. Brothers and P. M. Conn (1997). "Regulation of G(q/11)alpha by the gonadotropin-releasing hormone receptor." Mol Endocrinol **11**(6): 738-746.

Steinberger, A., K. H. Thanki and B. Siegal (1974). "FSH binding in rat testes during maturation and following hypophysectomy. Cellular localization of FSH receptors." Curr Top Mol Endocrinol **1**: 177-191.

Stojilkovic, S. S., J. Reinhart and K. J. Catt (1994). "Gonadotropin-releasing hormone receptors: structure and signal transduction pathways." Endocr Rev **15**(4): 462-499.

Struhl, K. (2007). "Transcriptional noise and the fidelity of initiation by RNA polymerase II." Nat Struct Mol Biol **14**(2): 103-105.

Suh, M. R., Y. Lee, J. Y. Kim, S. K. Kim, S. H. Moon, J. Y. Lee, K. Y. Cha, H. M. Chung, H. S. Yoon, S. Y. Moon, V. N. Kim and K. S. Kim (2004). "Human embryonic stem cells express a unique set of microRNAs." Dev Biol **270**(2): 488-498.

Suh, N., L. Baehner, F. Moltzahn, C. Melton, A. Shenoy, J. Chen and R. Blelloch (2010). "MicroRNA function is globally suppressed in mouse oocytes and early embryos." Curr Biol **20**(3): 271-277.

Sun, B. K., A. M. Deaton and J. T. Lee (2006). "A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization." Mol Cell **21**(5): 617-628.

Sun, F. L. and S. C. Elgin (1999). "Putting boundaries on silence." Cell **99**(5): 459-462.

Sun, L., Y. Peng, A. C. Sharrow, J. Iqbal, Z. Zhang, D. J. Papachristou, S. Zaidi, L. L. Zhu, B. B. Yaroslavskiy, H. Zhou, A. Zallone, M. R. Sairam, T. R. Kumar, W. Bo, J. Braun, L. Cardoso-Landa, M. B. Schaffler, B. S. Moonga, H. C. Blair and M. Zaidi (2006). "FSH directly regulates bone mass." Cell **125**(2): 247-260.

Szabo, P., S. H. Tang, A. Rentsendorj, G. P. Pfeifer and J. R. Mann (2000). "Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function." Curr Biol **10**(10): 607-610.

Takayama, K., H. Sasano, T. Fukaya, K. Morohashi, T. Suzuki, M. Tamura, M. J. Costa and A. Yajima (1995). "Immunohistochemical localization of Ad4-binding protein with correlation to steroidogenic

enzyme expression in cycling human ovaries and sex cord stromal tumors." J Clin Endocrinol Metab **80**(9): 2815-2821.

Tan, J. A., S. H. Hall, K. G. Hamil, G. Grossman, P. Petrusz and F. S. French (2002). "Protein inhibitors of activated STAT resemble scaffold attachment factors and function as interacting nuclear receptor coregulators." J Biol Chem **277**(19): 16993-17001.

Tano, K., R. Mizuno, T. Okada, R. Rakwal, J. Shibato, Y. Masuo, K. Ijiri and N. Akimitsu (2010). "MALAT-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes." FEBS Lett **584**(22): 4575-4580.

Tano, M., T. Minegishi, K. Nakamura, S. Karino, Y. Ibuki and K. Miyamoto (1997). "Transcriptional and post-transcriptional regulation of FSH receptor in rat granulosa cells by cyclic AMP and activin." J Endocrinol **153**(3): 465-473.

Tano, M., T. Minegishi, K. Nakamura, M. Nakamura, S. Karino, K. Miyamoto and Y. Ibuki (1995). "Regulation of follistatin messenger ribonucleic acid in cultured rat granulosa cells." Mol Cell Endocrinol **109**(2): 167-174.

Tena-Sempere, M., P. R. Manna and I. Huhtaniemi (1999). "Molecular cloning of the mouse follicle-stimulating hormone receptor complementary deoxyribonucleic acid: functional expression of alternatively spliced variants and receptor inactivation by a C566T transition in exon 7 of the coding sequence." Biol Reprod **60**(6): 1515-1527.

Thanos, D. and T. Maniatis (1995). "Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome." Cell **83**(7): 1091-1100.

Themmen, A. P., L. J. Blok, M. Post, W. M. Baarends, J. W. Hoogerbrugge, M. Parmentier, G. Vassart and J. A. Grootegoed (1991). "Follitropin receptor down-regulation involves a cAMP-dependent post-transcriptional decrease of receptor mRNA expression." Mol Cell Endocrinol **78**(3): R7-13.

Thomas, J. W., J. W. Touchman, R. W. Blakesley, G. G. Bouffard, S. M. Beckstrom-Sternberg, E. H. Margulies, M. Blanchette, A. C. Siepel, P. J. Thomas, J. C. McDowell, B. Maskeri, N. F. Hansen, M. S. Schwartz, R. J. Weber, W. J. Kent, D. Karolchik, T. C. Bruen, R. Bevan, D. J. Cutler, S. Schwartz, L. Elnitski, J. R. Idol, A. B. Prasad, S. Q. Lee-Lin, V. V. Maduro, T. J. Summers, M. E. Portnoy, N. L. Dietrich, N. Akhter, K. Ayele, B. Benjamin, K. Cariaga, C. P. Brinkley, S. Y. Brooks, S. Granite, X. Guan, J. Gupta, P. Haghighi, S. L. Ho, M. C. Huang, E. Karlins, P. L. Laric, R. Legaspi, M. J. Lim, Q. L. Maduro, C. A. Masiello, S. D. Mastrian, J. C. McCloskey, R. Pearson, S. Stantripop, E. E. Tiongson, J. T. Tran, C. Tsurgeon, J. L. Vogt, M. A. Walker, K. D. Wetherby, L. S. Wiggins, A. C. Young, L. H. Zhang, K. Osoegawa, B. Zhu, B. Zhao, C. L. Shu, P. J. De Jong, C. E. Lawrence, A. F. Smit, A. Chakravarti, D. Haussler, P. Green, W. Miller and E. D. Green (2003). "Comparative analyses of multi-species sequences from targeted genomic regions." Nature **424**(6950): 788-793.

Tian, D., S. Sun and J. T. Lee (2010). "The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation." Cell **143**(3): 390-403.

Tolhuis, B., R. J. Palstra, E. Splinter, F. Grosveld and W. de Laat (2002). "Looping and interaction between hypersensitive sites in the active beta-globin locus." Mol Cell **10**(6): 1453-1465.

Topilko, P., S. Schneider-Maunoury, G. Levi, A. Trembleau, D. Gourdji, M. A. Driancourt, C. V. Rao and P. Charnay (1998). "Multiple pituitary and ovarian defects in Krox-24 (NGFI-A, Egr-1)-targeted mice." Mol Endocrinol **12**(1): 107-122.

Townes, T. M., J. B. Lingrel, H. Y. Chen, R. L. Brinster and R. D. Palmiter (1985). "Erythroid-specific expression of human beta-globin genes in transgenic mice." EMBO J **4**(7): 1715-1723.

Tremblay, J. J. and J. Drouin (1999). "Egr-1 is a downstream effector of GnRH and synergizes by direct interaction with Ptx1 and SF-1 to enhance luteinizing hormone beta gene transcription." Mol Cell Biol **19**(4): 2567-2576.

Tripathi, V., J. D. Ellis, Z. Shen, D. Y. Song, Q. Pan, A. T. Watt, S. M. Freier, C. F. Bennett, A. Sharma, P. A. Bubulya, B. J. Blencowe, S. G. Prasanth and K. V. Prasanth (2010). "The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation." Mol Cell **39**(6): 925-938.

Trompouki, E., T. V. Bowman, L. N. Lawton, Z. P. Fan, D. C. Wu, A. DiBiase, C. S. Martin, J. N. Cech, A. K. Sessa, J. L. Leblanc, P. Li, E. M. Durand, C. Mosimann, G. C. Heffner, G. Q. Daley, R. F. Paulson, R. A. Young and L. I. Zon (2011). "Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration." Cell **147**(3): 577-589.

Tsai-Morris, C. H., E. Buczko, W. Wang, X. Z. Xie and M. L. Dufau (1991). "Structural organization of the rat luteinizing hormone (LH) receptor gene." J Biol Chem **266**(17): 11355-11359.

Tsai-Morris, C. H., Y. Geng, E. Buczko and M. L. Dufau (1995). "Characterization of diverse functional elements in the upstream Sp1 domain of the rat luteinizing hormone receptor gene promoter." J Biol Chem **270**(13): 7487-7494.

Tsai-Morris, C. H., Y. Geng, X. Z. Xie, E. Buczko and M. L. Dufau (1994). "Transcriptional protein binding domains governing basal expression of the rat luteinizing hormone receptor gene." J Biol Chem **269**(22): 15868-15875.

Tsai-Morris, C. H., X. Xie, W. Wang, E. Buczko and M. L. Dufau (1993). "Promoter and regulatory regions of the rat luteinizing hormone receptor gene." J Biol Chem **268**(6): 4447-4452.
Tsai, M. C., O. Manor, Y. Wan, N. Mosammaparast, J. K. Wang, F. Lan, Y. Shi, E. Segal and H. Y. Chang (2010). "Long noncoding RNA as modular scaffold of histone modification complexes." Science **329**(5992): 689-693.

Tsutsumi, M., W. Zhou, R. P. Millar, P. L. Mellon, J. L. Roberts, C. A. Flanagan, K. Dong, B. Gillo and S. C. Sealfon (1992). "Cloning and functional expression of a mouse gonadotropin-releasing hormone receptor." Mol Endocrinol **6**(7): 1163-1169.

Ulitsky, I., A. Shkumatava, C. H. Jan, H. Sive and D. P. Bartel (2011). "Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution." Cell **147**(7): 1537-1550.

Urs, A. N., E. Dammer and M. B. Sewer (2006). "Sphingosine regulates the transcription of CYP17 by binding to steroidogenic factor-1." Endocrinology **147**(11): 5249-5258.

Vagin, V. V., M. S. Klenov, A. I. Kalmykova, A. D. Stolyarenko, R. N. Kotelnikov and V. A. Gvozdev (2004). "The RNA interference proteins and vasa locus are involved in the silencing of retrotransposons in the female germline of Drosophila melanogaster." RNA Biol **1**(1): 54-58.

Valen, E., P. Preker, P. R. Andersen, X. Zhao, Y. Chen, C. Ender, A. Dueck, G. Meister, A. Sandelin and T. H. Jensen (2011). "Biogenic mechanisms and utilization of small RNAs derived from human protein-coding genes." Nat Struct Mol Biol **18**(9): 1075-1082.

Valenzuela, L. and R. T. Kamakaka (2006). "Chromatin insulators." Annu Rev Genet **40**: 107-138.
Vasilyev, V. V., M. A. Lawson, D. Dipaolo, N. J. Webster and P. L. Mellon (2002). "Different signaling pathways control acute induction versus long-term repression of LHbeta transcription by GnRH." Endocrinology **143**(9): 3414-3426.

Vassart, G., L. Pardo and S. Costagliola (2004). "A molecular dissection of the glycoprotein hormone receptors." Trends Biochem Sci **29**(3): 119-126.

Vasudevan, S., Y. Tong and J. A. Steitz (2007). "Switching from repression to activation: microRNAs can up-regulate translation." Science **318**(5858): 1931-1934.

Vazquez, B. N., T. Laguna, J. Carabana, M. S. Krangel and P. Lauzurica (2009). "CD69 gene is differentially regulated in T and B cells by evolutionarily conserved promoter-distal elements." J Immunol **183**(10): 6513-6521.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. Zhang, G. L. Gabor Miklos, C. Nelson, S. Broder, A. G. Clark, J. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A. E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T. J. Heiman, M. E. Higgins, R. R. Ji, Z. Ke, K. A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G. V. Merkulov, N. Milshina, H. M. Moore, A. K. Naik, V. A. Narayan, B. Neelam, D. Nusskern, D. B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M. L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y. H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N. N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J. F. Abril, R. Guigo, M. J. Campbell, K. V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y. H. Chiang, M. Coyne, C. Dahlke, A. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh and X. Zhu (2001). "The sequence of the human genome." Science **291**(5507): 1304-1351.

Viollet, B., A. M. Lefrancois-Martinez, A. Henrion, A. Kahn, M. Raymondjean and A. Martinez (1996). "Immunochemical characterization and transacting properties of upstream stimulatory factor isoforms." J Biol Chem **271**(3): 1405-1415.

Visel, A., M. J. Blow, Z. Li, T. Zhang, J. A. Akiyama, A. Holt, I. Plajzer-Frick, M. Shoukry, C. Wright, F. Chen, V. Afzal, B. Ren, E. M. Rubin and L. A. Pennacchio (2009). "ChIP-seq accurately predicts tissue-specific activity of enhancers." Nature **457**(7231): 854-858.

Visel, A., S. Prabhakar, J. A. Akiyama, M. Shoukry, K. D. Lewis, A. Holt, I. Plajzer-Frick, V. Afzal, E. M. Rubin and L. A. Pennacchio (2008). "Ultraconservation identifies a small subset of extremely constrained developmental enhancers." Nat Genet **40**(2): 158-160.

Visel, A., Y. Zhu, D. May, V. Afzal, E. Gong, C. Attanasio, M. J. Blow, J. C. Cohen, E. M. Rubin and L. A. Pennacchio (2010). "Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice." Nature **464**(7287): 409-412.

Viswanathan, P., M. A. Wood and W. H. Walker (2009). "Follicle-stimulating hormone (FSH) transiently blocks FSH receptor transcription by increasing inhibitor of deoxyribonucleic acid binding/differentiation-2 and decreasing upstream stimulatory factor expression in rat Sertoli cells." Endocrinology **150**(8): 3783-3791.

Wan, L. B., H. Pan, S. Hannenhalli, Y. Cheng, J. Ma, A. Fedoriw, V. Lobanenkov, K. E. Latham, R. M. Schultz and M. S. Bartolomei (2008). "Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development." Development **135**(16): 2729-2738.

Wang, K. C. and H. Y. Chang (2011). "Molecular mechanisms of long noncoding RNAs." Mol Cell **43**(6): 904-914.

Wang, W., C. Zhang, A. Marimuthu, H. I. Krupka, M. Tabrizizad, R. Shelloe, U. Mehra, K. Eng, H. Nguyen, C. Settachatgul, B. Powell, M. V. Milburn and B. L. West (2005). "The crystal structures of human steroidogenic factor-1 and liver receptor homologue-1." Proc Natl Acad Sci U S A **102**(21): 7505-7510.

Waterston, R. H., K. Lindblad-Toh, E. Birney, J. Rogers, J. F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson, P. An, S. E. Antonarakis, J. Attwood, R. Baertsch, J. Bailey, K. Barlow, S. Beck, E. Berry, B. Birren, T. Bloom, P. Bork, M. Botcherby, N. Bray, M. R. Brent, D. G. Brown, S. D. Brown, C. Bult, J. Burton, J. Butler, R. D. Campbell, P. Carninci, S. Cawley, F. Chiaromonte, A. T. Chinwalla, D. M. Church, M. Clamp, C. Clee, F. S. Collins, L. L. Cook, R. R. Copley, A. Coulson, O. Couronne, J. Cuff, V. Curwen, T. Cutts, M. Daly, R. David, J. Davies, K. D. Delehaunty, J. Deri, E. T. Dermitzakis, C. Dewey, N. J. Dickens, M. Diekhans, S. Dodge, I. Dubchak, D. M. Dunn, S. R. Eddy, L. Elnitski, R. D. Emes, P. Eswara, E. Eyras, A. Felsenfeld, G. A. Fewell, P. Flicek, K. Foley, W. N. Frankel, L. A. Fulton, R. S. Fulton, T. S. Furey, D. Gage, R. A. Gibbs, G. Glusman, S. Gnerre, N. Goldman, L. Goodstadt, D. Grafham, T. A. Graves, E. D. Green, S. Gregory, R. Guigo, M. Guyer, R. C. Hardison, D. Haussler, Y. Hayashizaki, L. W. Hillier, A. Hinrichs, W. Hlavina, T. Holzer, F. Hsu, A. Hua, T. Hubbard, A. Hunt, I. Jackson, D. B. Jaffe, L. S. Johnson, M. Jones, T. A. Jones, A. Joy, M. Kamal, E. K. Karlsson, D. Karolchik, A. Kasprzyk, J. Kawai, E. Keibler, C. Kells, W. J. Kent, A. Kirby, D. L. Kolbe, I. Korf, R. S. Kucherlapati, E. J. Kulbokas, D. Kulp, T. Landers, J. P. Leger, S. Leonard, I. Letunic, R. Levine, J. Li, M. Li, C. Lloyd, S. Lucas, B. Ma, D. R. Maglott, E. R. Mardis, L. Matthews, E. Mauceli, J. H. Mayer, M. McCarthy, W. R. McCombie, S. McLaren, K. McLay, J. D. McPherson, J. Meldrim, B. Meredith, J. P. Mesirov, W. Miller, T. L. Miner, E. Mongin, K. T. Montgomery, M. Morgan,

R. Mott, J. C. Mullikin, D. M. Muzny, W. E. Nash, J. O. Nelson, M. N. Nhan, R. Nicol, Z. Ning, C. Nusbaum, M. J. O'Connor, Y. Okazaki, K. Oliver, E. Overton-Larty, L. Pachter, G. Parra, K. H. Pepin, J. Peterson, P. Pevzner, R. Plumb, C. S. Pohl, A. Poliakov, T. C. Ponce, C. P. Ponting, S. Potter, M. Quail, A. Reymond, B. A. Roe, K. M. Roskin, E. M. Rubin, A. G. Rust, R. Santos, V. Sapojnikov, B. Schultz, J. Schultz, M. S. Schwartz, S. Schwartz, C. Scott, S. Seaman, S. Searle, T. Sharpe, A. Sheridan, R. Shownkeen, S. Sims, J. B. Singer, G. Slater, A. Smit, D. R. Smith, B. Spencer, A. Stabenau, N. Stange-Thomann, C. Sugnet, M. Suyama, G. Tesler, J. Thompson, D. Torrents, E. Trevaskis, J. Tromp, C. Ucla, A. Ureta-Vidal, J. P. Vinson, A. C. Von Niederhausern, C. M. Wade, M. Wall, R. J. Weber, R. B. Weiss, M. C. Wendl, A. P. West, K. Wetterstrand, R. Wheeler, S. Whelan, J. Wierzbowski, D. Willey, S. Williams, R. K. Wilson, E. Winter, K. C. Worley, D. Wyman, S. Yang, S. P. Yang, E. M. Zdobnov, M. C. Zody and E. S. Lander (2002). "Initial sequencing and comparative analysis of the mouse genome." Nature **420**(6915): 520-562.

Watson, J. D. and F. H. Crick (1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid." Nature **171**(4356): 737-738.

Wendt, K. S., K. Yoshida, T. Itoh, M. Bando, B. Koch, E. Schirghuber, S. Tsutsumi, G. Nagae, K. Ishihara, T. Mishiro, K. Yahata, F. Imamoto, H. Aburatani, M. Nakao, N. Imamoto, K. Maeshima, K. Shirahige and J. M. Peters (2008). "Cohesin mediates transcriptional insulation by CCCTC-binding factor." Nature **451**(7180): 796-801.

Wilczynski, B. and E. E. Furlong (2010). "Dynamic CRM occupancy reflects a temporal map of developmental progression." Mol Syst Biol **6**: 383.

Wilhelm, D. and C. Englert (2002). "The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1." Genes Dev **16**(14): 1839-1851.

Willingham, A. T., A. P. Orth, S. Batalov, E. C. Peters, B. G. Wen, P. Aza-Blanc, J. B. Hogenesch and P. G. Schultz (2005). "A strategy for probing the function of noncoding RNAs finds a repressor of NFAT." Science **309**(5740): 1570-1573.

Windle, J. J., R. I. Weiner and P. L. Mellon (1990). "Cell lines of the pituitary gonadotrope lineage derived by targeted oncogenesis in transgenic mice." Mol Endocrinol **4**(4): 597-603.

Wolfe, M. W. and G. B. Call (1999). "Early growth response protein 1 binds to the luteinizing hormone-beta promoter and mediates gonadotropin-releasing hormone-stimulated gene expression." Mol Endocrinol **13**(5): 752-763.

Wood, M. A. and W. H. Walker (2009). "USF1/2 transcription factor DNA-binding activity is induced during rat Sertoli cell differentiation." Biol Reprod **80**(1): 24-33.

Woodruff, T. K., J. D'Agostino, N. B. Schwartz and K. E. Mayo (1988). "Dynamic changes in inhibin messenger RNAs in rat ovarian follicles during the reproductive cycle." Science **239**(4845): 1296-1299.

Woodson, K. G., P. A. Crawford, Y. Sadovsky and J. Milbrandt (1997). "Characterization of the promoter of SF-1, an orphan nuclear receptor required for adrenal and gonadal development." Mol Endocrinol **11**(2): 117-126.

Woolfe, A., M. Goodson, D. K. Goode, P. Snell, G. K. McEwen, T. Vavouri, S. F. Smith, P. North, H. Callaway, K. Kelly, K. Walter, I. Abnizova, W. Gilks, Y. J. Edwards, J. E. Cooke and G. Elgar (2005).

"Highly conserved non-coding sequences are associated with vertebrate development." PLoS Biol **3**(1): e7.
Wu, J., L. T. Smith, C. Plass and T. H. Huang (2006). "ChIP-chip comes of age for genome-wide functional analysis." Cancer Res **66**(14): 6899-6902.

Wurmbach, E., T. Yuen, B. J. Ebersole and S. C. Sealfon (2001). "Gonadotropin-releasing hormone receptor-coupled gene network organization." J Biol Chem **276**(50): 47195-47201.

Wutz, A., T. P. Rasmussen and R. Jaenisch (2002). "Chromosomal silencing and localization are mediated by different domains of Xist RNA." Nat Genet **30**(2): 167-174.
Xiao, T., J. Wallace and G. Felsenfeld (2011). "Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity." Mol Cell Biol **31**(11): 2174-2183.

Xing, W., N. Danilovich and M. R. Sairam (2002). "Orphan receptor chicken ovalbumin upstream promoter transcription factors inhibit steroid factor-1, upstream stimulatory factor, and activator protein-1 activation of ovine follicle-stimulating hormone receptor expression via composite cis-elements." Biol Reprod **66**(6): 1656-1666.

Xing, W. and M. R. Sairam (2001). "Characterization of regulatory elements of ovine follicle-stimulating hormone (FSH) receptor gene: the role of E-box in the regulation of ovine FSHreceptor expression." Biol Reprod **64**(2): 579-589.

Xing, W. and M. R. Sairam (2002). "Cross talk of two Krupple transcription factors regulates expression of the ovine FSH receptor gene." Biochem Biophys Res Commun **295**(5): 1096-1101.

Xing, W. and M. R. Sairam (2002). "Retinoic acid mediates transcriptional repression of ovine follicle-stimulating hormone receptor gene via a pleiotropic nuclear receptor response element." Biol Reprod **67**(1): 204-211.

Xu, X., K. Tsumagari, J. Sowden, R. Tawil, A. P. Boyle, L. Song, T. S. Furey, G. E. Crawford and M. Ehrlich (2009). "DNaseI hypersensitivity at gene-poor, FSH dystrophy-linked 4q35.2." Nucleic Acids Res **37**(22): 7381-7393.

Xue, Q., Z. Lin, P. Yin, M. P. Milad, Y. H. Cheng, E. Confino, S. Reierstad and S. E. Bulun (2007). "Transcriptional activation of steroidogenic factor-1 by hypomethylation of the 5' CpG island in endometriosis." J Clin Endocrinol Metab **92**(8): 3261-3267.

Yamada, Y., H. Yamamoto, T. Yonehara, H. Kanasaki, H. Nakanishi, E. Miyamoto and K. Miyazaki (2004). "Differential activation of the luteinizing hormone beta-subunit promoter by activin and gonadotropin-releasing hormone: a role for the mitogen-activated protein kinase signaling pathway in LbetaT2 gonadotrophs." Biol Reprod **70**(1): 236-243.

Yanez-Cuna, J. O., H. Q. Dinh, E. Z. Kvon, D. Shlyueva and A. Stark (2012). "Uncovering cis-regulatory sequence requirements for context-specific transcription factor binding." Genome Res **22**(10): 2018-2030.

Yao, H., K. Brick, Y. Evrard, T. Xiao, R. D. Camerini-Otero and G. Felsenfeld (2010). "Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA." Genes Dev **24**(22): 2543-2555.

Yen, S. S. and C. C. Tsai (1971). "The biphasic pattern in the feedback action of ethinyl estradiol on the release of pituitary FSH and LH." J Clin Endocrinol Metab **33**(6): 882-887.

Yokoi, T., M. Ohmichi, K. Tasaka, A. Kimura, Y. Kanda, J. Hayakawa, M. Tahara, K. Hisamoto, H. Kurachi and Y. Murata (2000). "Activation of the luteinizing hormone beta promoter by gonadotropin-releasing hormone requires c-Jun NH2-terminal protein kinase." J Biol Chem **275**(28): 21639-21647.

Yokota, Y., S. Mori, O. Narumi and K. Kitajima (2001). "In vivo function of a differentiation inhibitor, Id2." IUBMB Life **51**(4): 207-214.

Yoon, Y. S., S. Jeong, Q. Rong, K. Y. Park, J. H. Chung and K. Pfeifer (2007). "Analysis of the H19ICR insulator." Mol Cell Biol **27**(9): 3499-3510.

Young, E. A. (1995). "The role of gonadal steroids in hypothalamic-pituitary-adrenal axis regulation." Crit Rev Neurobiol **9**(4): 371-381.

Zaidi, M., H. C. Blair, J. Iqbal, L. L. Zhu, T. R. Kumar, A. Zallone and L. Sun (2007). "Proresorptive actions of FSH and bone loss." Ann N Y Acad Sci **1116**: 376-382.

Zeitlinger, J., I. Simon, C. T. Harbison, N. M. Hannett, T. L. Volkert, G. R. Fink and R. A. Young (2003). "Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling." Cell **113**(3): 395-404.
Zhang, M., H. Shi, D. L. Segaloff and B. J. Van Voorhis (2001). "Expression and localization of luteinizing hormone receptor in the female mouse reproductive tract." Biol Reprod **64**(1): 179-187.

Zhang, X. Y., J. Chen, Y. F. Zheng, X. L. Gao, Y. Kang, J. C. Liu, M. J. Cheng, H. Sun and C. J. Xu (2009). "Follicle-stimulating hormone peptide can facilitate paclitaxel nanoparticles to target ovarian carcinoma in vivo." Cancer Res **69**(16): 6506-6514.

Zhao, J., B. K. Sun, J. A. Erwin, J. J. Song and J. T. Lee (2008). "Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome." Science **322**(5902): 750-756.

Zhao, L., M. Bakke, Y. Krimkevich, L. J. Cushman, A. F. Parlow, S. A. Camper and K. L. Parker (2001). "Steroidogenic factor 1 (SF1) is essential for pituitary gonadotrope function." Development **128**(2): 147-154.
Zhao, L., M. Bakke and K. L. Parker (2001). "Pituitary-specific knockout of steroidogenic factor 1." Mol Cell Endocrinol **185**(1-2): 27-32.

Zhao, L., K. W. Kim, Y. Ikeda, K. K. Anderson, L. Beck, S. Chase, S. A. Tobet and K. L. Parker (2008). "Central nervous system-specific knockout of steroidogenic factor 1 results in increased anxiety-like behavior." Molecular endocrinology **22**(6): 1403-1415.

Zheng, W., M. S. Magid, E. E. Kramer and Y. T. Chen (1996). "Follicle-stimulating hormone receptor is expressed in human ovarian surface epithelium and fallopian tube." Am J Pathol **148**(1): 47-53.
Zheng, X., Y. Wang, Q. Yao, Z. Yang and K. Chen (2009). "A genome-wide survey on basic helix-loop-helix transcription factors in rat and mouse." Mamm Genome **20**(4): 236-246.

Zhou, G. L., L. Xin, W. Song, L. J. Di, G. Liu, X. S. Wu, D. P. Liu and C. C. Liang (2006). "Active chromatin hub of the mouse alpha-globin locus forms in a transcription factory of clustered housekeeping genes." Mol Cell Biol **26**(13): 5096-5105.

Ziebarth, J. D., A. Bhattacharya and Y. Cui (2013). "CTCFBSDB 2.0: a database for CTCF-binding sites and genome organization." Nucleic Acids Res **41**(Database issue): D188-194.

Zubair, M., S. Oka, K. L. Parker and K. Morohashi (2009). "Transgenic expression of Ad4BP/SF-1 in fetal adrenal progenitor cells leads to ectopic adrenal formation." Mol Endocrinol **23**(10): 1657-1667.