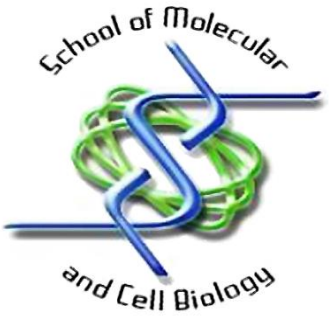
UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG



ROLE OF *FET-1* AS A POSSIBLE OVARIAN SEX DETERMINING GENE IN AVIANS

Chiron Loubser

A Dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science

Johannesburg, 2019

DECLARATION

I, Chiron Loubser (751733), am a student registered for the degree of MSc in the academic year 2019.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged. In this context, I understand that the use of editing services is considered aided work and must be declared.
- I have not submitted this work before for any other degree or examination at this or any other University.
- The information used in the Dissertation **HAS NOT** been obtained by me while employed by, or working under the aegis of, any person or organisation other than the University.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature  25 day of February

ABSTRACT

The mass culling of male chicks of egg-laying breeds is ethically and financially questionable, estimated to cost South African commercial breeders $10 million annually, and close to $1 billion worldwide. Understanding sex determination in birds is essential to the development of an all-female chicken breed, which would eliminate the need for mass culling. In birds, where females are ZW and males are ZZ, sex determination is hypothesised to be the result of either a W determinant or Z dosage. One of the few candidate W determinants is the Female Expressed Transcript -1 (*FET-1*), previously localized to the W chromosome with female-specific expression. The aim of this research was to further investigate the function of *FET-1* and its potential role in avian sex determination. RT-PCR on extracted RNA from E4.5 gonads confirmed *FET-1* expression. Unexpectedly, expression was significantly higher in females than in males. A local alignment search of *FET-1* against the chicken genome suggested multiple *FET-1* integrations in the Z chromosomes and autosomes, consistent with retroviral elements such as *FET-1*. Using RACE (Rapid Amplification of cDNA Ends) two long intergenic non-coding RNA (lincRNA) populations were identified that overlapped
d along the *FET-1* transcript. RNA expression was observed for both of the lincRNA populations as well as the FET-1 protein-coding sequence at E4.5 and E6.5 embryos; stronger gonadal and epidermal staining observed in females compared to males, suggesting a sex-specific role. Whole mount IHC using a rabbit polyclonal anti-FET-1 antibody showed the presence of *FET-1* protein in the epidermis but not the gonads. Collectively the data suggests that *FET-1* is not likely to be a W

determinant; however predominant expression in female suggests a role in the development of sex-specific characteristics.

DEDICATION

My parents

For their unconditional support, encouragement and sacrifice shaping me into the person I am today.

My sister

"Sometimes we need someone to simply be there. Not to fix anything, or to do anything in particular, but just to let us feel that we are cared for and supported"

- Unknown

ACKNOWLEDGEMENTS

A sincere thanks my supervisor, Professor Nikitina for her support and guidance during the project. I would also like to acknowledge the University if the Witwatersrand as well as the National Research Foundation for their financial support during the time period of this MSc.

RESEARCH OUPTUTS

1. Conference outputs: poster presentation, *FET-1* in avian sex determination, 77th Annual meeting of the Society for Developmental Biology, Portland, USA

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

♂	Male
♀	Female
2n	Diploid chromosome number
Ab	Antibody
AMH	Anit-Müllerian hormone
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BLAST	Basic Local Alignment Search Tool
BLASTp	Protein BLAST
CASI	Cell autonomous sex identity
CBX2	Chromobox 2
cDNA	complementary DNA
CDS	Coding sequence
DHH	Desert hedgehog
DMRT1	Doublesex and mab-3 related transcription factor 1
E	Embryonic day
E.coli	*Escherichia coli*
ERA	Estrogen receptor α
ESD	Environmental sex determination

FAF	Female-associated factor
FET-1	Female expressed transcript-1
FGF9	Fibroblast growth factor 9
FISH	Fluorescent *in situ* hybridization
FOG2	Friend of GATA protein 2
FOXL2	Forkhead box L2
GSD	Genetic sex determination
GSP	Gene-specific primers
HINTW	Histidine triad nucleotide binding protein W
HINTZ	Histidine triad nucleotide binding protein Z
hr	hour/s
KO	Knock out
LHX9	LIM Homeobox 9
LincRNA	Long intergenic non-coding RNA
lncRNA	long non-coding RNA
MHM	Male hypermethylated region
MTMR7	Myotubularin related protein 7
mya	million years ago
PBS	Phosphate buffered saline

NCBI	National Center for Biotechnology Information
NBT	Nitro blue tetrazolium
ncRNA	non-coding RNA
ORF	Open reading frame
PFA	Paraformaldehyde
PCR	Polymerase chain reaction
Pdgs	Prostaglandins
PGC	Primordial germ cells
RACE	Rapid amplification of cDNA ends
RSPON1	R-spondin 1
RT	Room temperature
RT-PCR	Reverse transcription PCR
SF1	Steriodogenic factor-1
SOX9	SRY-like box 9
SRA	steroid receptor RNA activator
SRY	Sex-determining region Y
Ta	Annealing temperature
UTR	Untranslated region
WNT4	Wnt family member 4

WT1	Wilms tumor-1
XIST	X inactive specific transcript
XX/XY	Mammalian sex-determining system
ZZ/ZW	Avian sex-determining system

CHAPTER 1 – INTRODUCTION

1. Literature review

1.1. Overview of vertebrate sex development systems

Most organisms which use sexual reproduction to create off spring have two sexes, male and female; the exception being hermaphrodites that poses both male and female qualities. In these organisms the majority of differences between males and females are determined during embryonic development. Two broad sex-determining mechanisms exists: 1) genetic sex determination (GSD), where sex is determined by inherited sex chromosomes and the presence of a sex-determining factor (usually a gene or a combination of genes) and; 2) environmental sex determination (ESD), where the male and female sex is determined by environmental factors experienced during early development regardless of the genotype[1]. The most studied GSD mechanism is the mammalian XY/XX system where males are heterogametic (XY) and females homogametic (XX). The sex-determining factor is the male Y-encoded Sex-determining region Y (*SRY*), acting as the primary male determinant to trigger male development[2,3]. The lesser known ZZ/ZW system is observed in birds where the females are heterogametic (ZW) and the males homogametic (ZZ)[4]. Amphibians, reptiles and fish have been shown to utilize both GSD and ESD mechanisms, alone or in combination, as well as both the XX/XY and ZZ/ZW systems[5–9]. Differentiation of the sex chromosomes usually occurs when one member of the ancestral autosomal pair (W/Y ancestor) acquires a sex-specific gene resulting in suppression of recombination with the other member (Z/X). In the region of decreased recombination, genes are easily inactivated, mutated and replaced by repetitive

sequences[10]. Sex development is split into two processes. The first is sex determination (primary sex), the process whereby undifferentiated gonads are directed (usually by transcription factors) to develop as testes or ovaries. The second is sex differentiation (secondary sex), when the products produced by the gonads (steroid hormones) induce the development of the phenotypic sex.

1.2. Sex development in mammals

The mammalian sex development system (Figure 1.1) is the most well-studied system (from mice and documented disorders of sex development in humans) and often used as a guideline for other animals. The gonads develop as bilateral, bipotential structures and differentiate into ovaries or testes based on expression of sex-specific genes. In humans, the genital ridge develops at 5 weeks after which the bipotential gonads develop and differentiate into ovaries or testes. In mice and humans, the development of the bipotential gonads is dependent on a gene network consisting of LIM Homeobox 9 (*LHX9),* Chromobox 2 *(CBX2),* Steriodogenic factor-1 *(SF1)* and Wilms tumor-1 (*WT1)*. Knock out (KO) phenotypes result in failure to develop discreet gonads along with male-female sex reversal of XY individuals[11].

Development into ovaries or testes requires antagonistic gene networks. At 7 weeks, the male-specific determinant *SRY* (Sex-determining region Y) is expressed in pre-Sertoli cells under the influence of *WT1, CBX2, GATA4* and its binding partner Friend of GATA protein 2 (*FOG2)*[11] to induce testes formation[2]. Synergistic binding of *SRY* and *SF1* to a SRY-like box 9 (*SOX9)* gonad specific enhancer up-regulates

SOX9 expression to initiate Sertoli cell differentiation[12]. One major function of *SOX9* is its antagonistic role toward the female developmental pathway, specifically Forkhead box L2 (*FOXL2)* and *β-catenin*[11]. Upon critical threshold expression of *Sox9*, several positive feedback mechanisms are initiated to maintain *SOX9* expression. This includes 1) auto-regulation; binding to its own gonad specific enhancer with *SF1* when *SRY* expression decreases; 2) a positive feedback loop between *SOX9* and Fibroblast growth factor 9 (*FGF9)* and 3) interactions with prostaglandin (Pdgs) signalling where Pdgs facilitates *SOX9* movement into the nucleus and *SOX9* in turn binds to the Pdgs promoter[13]. *FGF9* and Pdgs are further required for maintenance and sustained function of the Sertoli cells. *FGF9* and Wnt family member 4 (*WNT4)* of the female pathway are antagonistic and *FGF9* KO leads to male-to-female sex reversal[13]. Further, loss-of-function of both *SRY* and *SOX9* leads to male-to-female sex reversal. In *SRY* loss-of-function ectopic expression of *SOX9* is sufficient to rescue male development. Additionally, independent ectopic expression of both *SRY* and *SOX9* can induce testes formation in XX individuals. *SOX9*, *SF1* and *WT1* mediate the expression of the Anti-Müllerian hormone (Amh) from the Sertoli cells[14]. Amh is responsible for the degeneration of the Müllerian duct from which the female reproductive tract develops. Once differentiated, Sertoli cells induce the development of foetal Leydig cells via the expression of *DHH* signalling and the synthesis of androgens from these cells.

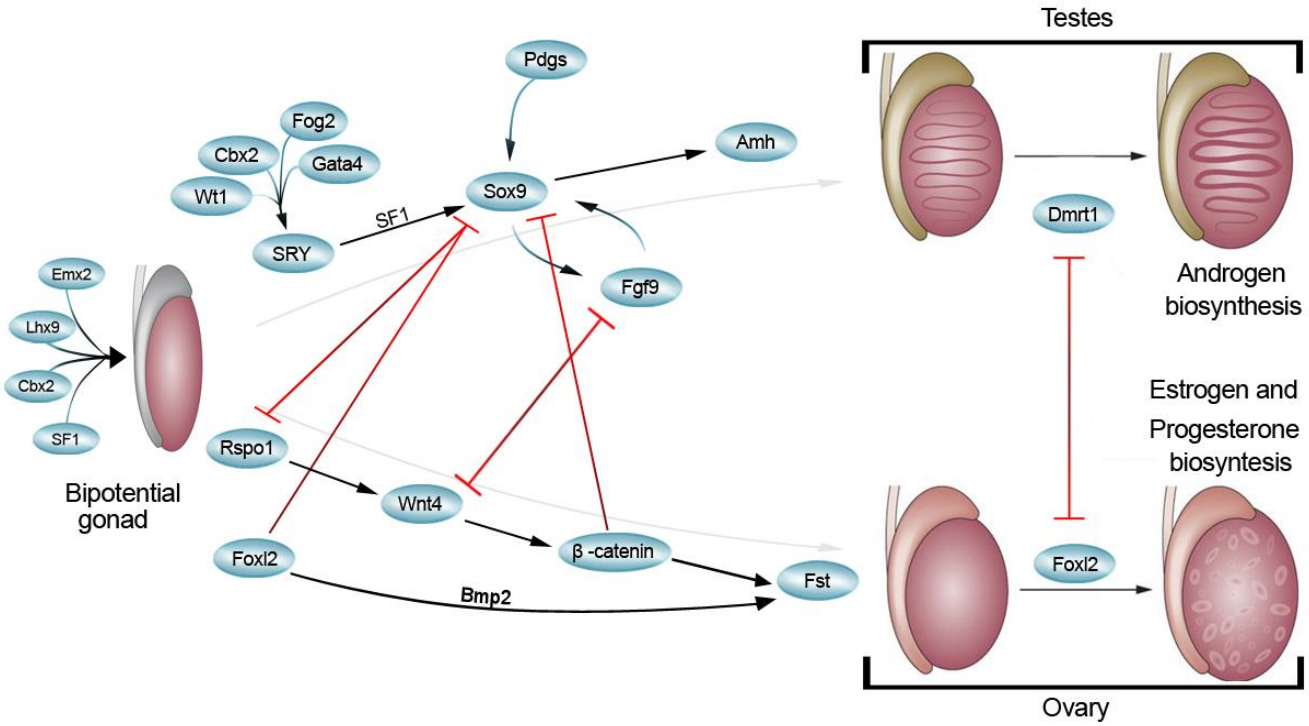


Figure 1. 1. Overview of mammalian sex determination. *SRY* acts as the male sex determinant that activates the male pathway, allowing *SOX9* expression to reach a critical threshold and induce expression of a range of factors, including Amh, required for male determination. *SOX9* further maintains its own expression via positive feedback loops with *FGF9* and Pdgs. *DMRT1* is expressed later in the male determining pathway and maintains the differentiated state of the testes. Female determination is dependent on two separate pathways; 1)*RSPON1* that mediates the expression of β-catenin via *WNT4* signalling to active *FST* and 2) *FOXL2* that also induces *FST* expression and also plays a role in steroid hormone synthesis in the ovary. Since *SRY* is absent, *SOX9* never reaches its critical threshold to further activate the male pathway. Image adapted from Eggers et al., 2014.

DMRT1 (Doublesex and mab-3 related transcription factor 1) expression occurs in the Sertoli cells from weeks 8-20 of development and is most abundant in the spermatogonia during childhood and post-puberty. Expression in the oocytes of early ovaries was also observed; however upon meiotic entry of the germ cells, *DMRT1* expression is completely down regulated. *DMRT1* is needed for maintenance of both the Sertoli and germ cells[13]. Loss-of-function of *DMRT1* results in male-female sex reversal in humans. *In vivo* studies have demonstrated the critical nature of *DMRT1* expression in the developing and adult testes: 1) deletion of *DMRT1* in adult testes

showed *FOXL2* female-like cells with increased estrogen synthesizing genes, aromatase and *HSD17β1*[15]; 2) KO of *DMRT1* in adult ovaries silenced *FOXL2* expression, activated multiple male-specific genes and changes the ovary-specific cells to Sertoli-like cells[16]. *DMRT1* additionally demonstrated to be responsible for dosage sensitive sex reversal in humans. This is shown in a complete male-female sex reversal due to a point mutation in an allele of *DMRT1* that disrupted wild type *DMRT1*-DNA binding[17].

Female development is mediated by two parallel pathways: 1) *RSPON1/ WNT4/β-catenin* canonical pathway and 2) *FOXL2* pathway. In the absence of *SRY*, *SOX9* expression does not reach its critical threshold and is suppressed by female-specific genes that are expressed by the time *SRY* and *SOX9* reach their maximal levels in the male gonad. *RSPO1* is a positive regulator of *WNT4*, which acts as a suppressor of *FGF9* and initiates the expression of *β-catenin*[11]. *β-catenin* facilitates the initial expression of *Fst* in the differentiating stroma (supportive tissue of the ovary consisting of connective tissues and blood vessels) to induce female coelomic blood vessel formation and inhibit that of the male[11]. *FOXL2* with *BMP2* maintains *FST* expression during later stages of follicle formation. *β-catenin* and *FOXL2* independently repress *SOX9*. Additionally, stabilized *β-catenin* is sufficient for a male-to-female sex reversal[18] while conditional KO of *FOXL2* results in ectopic expression of male-specific genes with the ovary specialised cells (granulosa cells and theca cells) differentiating into male specialised cells (Sertoli cells and Leydig cells)[13]. Both the *RSPO1/ WNT4/β-catenin* and *FOXL2* pathways are responsible for early female gonad differentiation and repression of the male pathway while *FOXL2* is responsible for tissue maintenance and follicle maturation.

After gonad differentiation, steroid hormones direct the phenotypic sex of the individual. In males *SF1* regulates the genes required for androgen synthesis, including testosterone from the Leydig cells. Testosterone is converted to dihydrotestosterone and both bind to androgen receptors to mediate the differentiation of the Wolffian duct into male internal genitalia. In females aromatase is expressed under the direction of *FOXL2*. Aromatase in turn mediated the synthesis of estrogen. *SF1* and *DAX1* act to inhibit *AMH* temporarily, allowing for development of the Müllerian duct into female internal genitalia under the influence of estrogen. In females, the Wolffian duct regresses due to the absence of testosterone[19].

1.3. Avian sex determination

1.3.1. ZW/ZZ chromosomes

Extant birds are divided into two clades that split approximately 120 million years ago: 1) the Palaeognathae, comprising of ratites (flightless birds: ostrich, emus and allies) and tinamous; and 2) the Neognathae that comprise all the rest of modern birds[20].

Most birds have diploid chromosomes numbers ranging between 2n= 76-80[21], with *Gallus gallus* 2n = 78. Unlike mammals, birds use ZZ/ZW sex chromosomes. Among neognathae the Z and W vary in size and degree of homology. In *Gallus gallus* the Z chromosome comprises 7% of the haploid genome and the W chromosome 1.5%. Among the Palaeognathae the Z and W chromosomes are generally homomorphic[22];

distinguishable only by their centromere positions and are genetically homologous[23]. Several genes mapped to the W chromosomes have been shown to be present and closely conserved on the Z chromosome, indicating evolution from an ancestral pair of autosomes[24]. Additionally, several genes located on the Z chromosome has been annotated to several human autosomes with a large part of the Z chromosome homologous to human chromosome 9[25,26]. This suggests that the ZW and XY chromosome systems evolved from separate autosome pairs.

The W chromosome is a degenerate version of the Z chromosome. Differentiation has made the W chromosome female specific and contributed to the loss of the majority of functional genes. This has occurred over two strata. Stratum 1: a single inversion on the W chromosome 102-170 million years ago (mya), disrupting homology to the long arm of the Z chromosome (Zq) and; Stratum 2: the independent termination among different avian lineages of recombination between the short arm of the Z (Zp) and W chromosome, 58-85 mya[23]. The independent termination accounts for the varying degree of Z-W homology between the Neognathae and Palaeognathae.

1.3.2. Established gene regulatory network involved in avian sex determination

Several sex-determining genes are conserved between mammals and avians; however, they may not be similar in function. In mammals, the Y-liked gene *SRY* is the key trigger to male development. Birds, however, neither have a Y chromosome

with an *SRY* nor an orthologue of the gene. *DMRT1* is also a conserved gene in mammalian and avian male gonad differentiation[27–29]. In birds it is Z-linked and expressed higher in males than females in the genital ridge prior to gonad differentiation[29]. Upon sexual differentiation expression becomes male specific[30]. Chicken KO of *DMRT1* was shown to cause partial feminization of the testes, marked with a decrease in *SOX9* expression and an ectopic increase of the female marker aromatase[28]. Over-expression of *DMRT1* induces expression of male-specific genes and antagonises that of females[31]. In the chicken, *DMRT1* clearly acts as an upstream sex-determining factor compared to higher vertebrates, where it is placed lower in the signalling pathway. Regardless, reduced *DMRT1* expression in birds appears similar to haplo-insufficiency in mammals which is associated with failure of testes differentiation and XY sex reversal, suggesting a highly conserved function across species. *SOX9* and AMH are two additional genes conserved in male gonad determination. *SOX9* regulates seminiferous cord formation and Sertoli cell differentiation[32]. AMH is responsible for the regression of the Müllerian ducts in males. In mammals *SOX9* precedes and triggers AMH expression. In birds, however, AMH precedes *SOX9*[33]. *SOX9* can therefore not trigger AMH, but can possibly maintain AMH expression since its promoter does contain *Sox9* binding site. It has additionally been shown that the AMH promoter contains an *SF1* binding site and that *SF1* can bind to and activate the AMH promoter[34]. A recently identified gene, *cHEMGN* was observed to play a significant role during male gonad development. Expressed in pre-Sertoli cells after sex determination and before the expression of *SOX9*, over expression of *cHEMGN* induced masculinisation of ZW gonads. The masculinisation of female gonads as a result of aromatase KO also induced ectopic *cHEMGN* expression[35].

There are two distinct canonical pathways in ovarian development. The first pathway involves *FOXL2*, *17βHSD* (17-β-hydroxysteroid dehydrogenase) and aromatase and mediates the steroidogenic action of estrogen.

In the chicken *FOXL2* is autosomal and is expressed before, during and after gonad differentiation[36]. Initial expression is observed in both gonads and a subsequent decrease in the right ovary reflects the asymmetric[37–39] ovary development observed in birds where the left ovary develops while the right ovary degenerates. In contrast to mammals, avian gonad development is sensitive to steroid hormones. Estrogen is a key factor regulating ovary development by controlling proliferation of the cortex[40]. The enzymes responsible for the production of estrogen from circulating androgens are aromatase and *17βHDS*, both only expressed in females[41]. Aromatase expression has been shown to colocalise with *FOXL2*[36] and its knockdown reduces *FOXL2* and estrogen expression. This shows that *FOXL2* acts upstream of aromatase in the synthesis of estrogen[42] and suggests the presence of a positive feedback loop mediated by estrogen acting upon *FOXL2* to maintain its expression. Estrogen receptor α (*ERα*) is required to mediate the function of estrogen in the developing ovary. *ERα* is expressed in the left gonadal cortex of both males and females[41,43]. Male expression is, however, transient and since neither aromatase nor *17βHDS* is expressed in males, estrogen has no developmental effects during testes differentiation.

The second pathway in females involves the autosomal *WNT4, RSPON1* and *β-catenin*. *RSPON1* is expressed dimorphically and upregulated in the female genital ridge. The *RSPON1* protein is expressed in the outer cortical zone of the left ovary, where germ cell meiosis and the differentiation of somatic supporting cells into follicles occur[44]. *WNT4* is expressed in both males and females while during gonadal differentiation expression becomes dimorphic; upregulated in females but down regulated in males[44]. *RSPON1* acts though *WNT4* signalling to stabilize intracellular *β-catenin*. *β-catenin* in turn enters the nucleus to regulate gene expression[44] associated with primordial germ cell proliferation[45] and folliculogenesis. *RSPON1* is also involved in female-specific upregulation and maintains expression of *WNT4*[46], possibly through a positive feedback loop of *β-catenin* acting on *WNT4*[40].

Little work has been performed on the knock down of *RSPON1, WNT4 and β-catenin* in chickens; inferences are therefore drawn from loss-of-function experiments performed in other model organisms. In mice, *Rspon1* knock down has resulted in depleted germ cells, masculinisation of gonads, down regulation of *Wnt4* and ectopic testosterone expression[47]. Through *Rspon1*, *Wnt4* is regulated to mediate germ cell development. In *Wnt4* KO mice, germ cells fail to accumulate in the cortex, fail to enter meiosis and appear to develop male characteristics[40]. The observation that *Wnt4* KO mice fail to develop follicles and subsequent ovarian tissue suggests that germ cells have an important role in maintaining ovarian morphology[40].

While the same factors function in the canonical pathways across vertebrates (e.g. *SOX9,* AMH and *DMRT1*) their relationship to one another within their pathways are

not conserved. In birds this is likely due to the absence of *SRY* or an *SRY*-orthologue. The male pathway likely evolved differently, utilizing the same male factors located on autosomes (*SOX9* and AMH) while elevating the sex-linked factor *DMRT1* to play the role of the avian male determinant. *DMRT1* in mammals acts to maintain the gonadal sex and prevent female reprogramming[48].

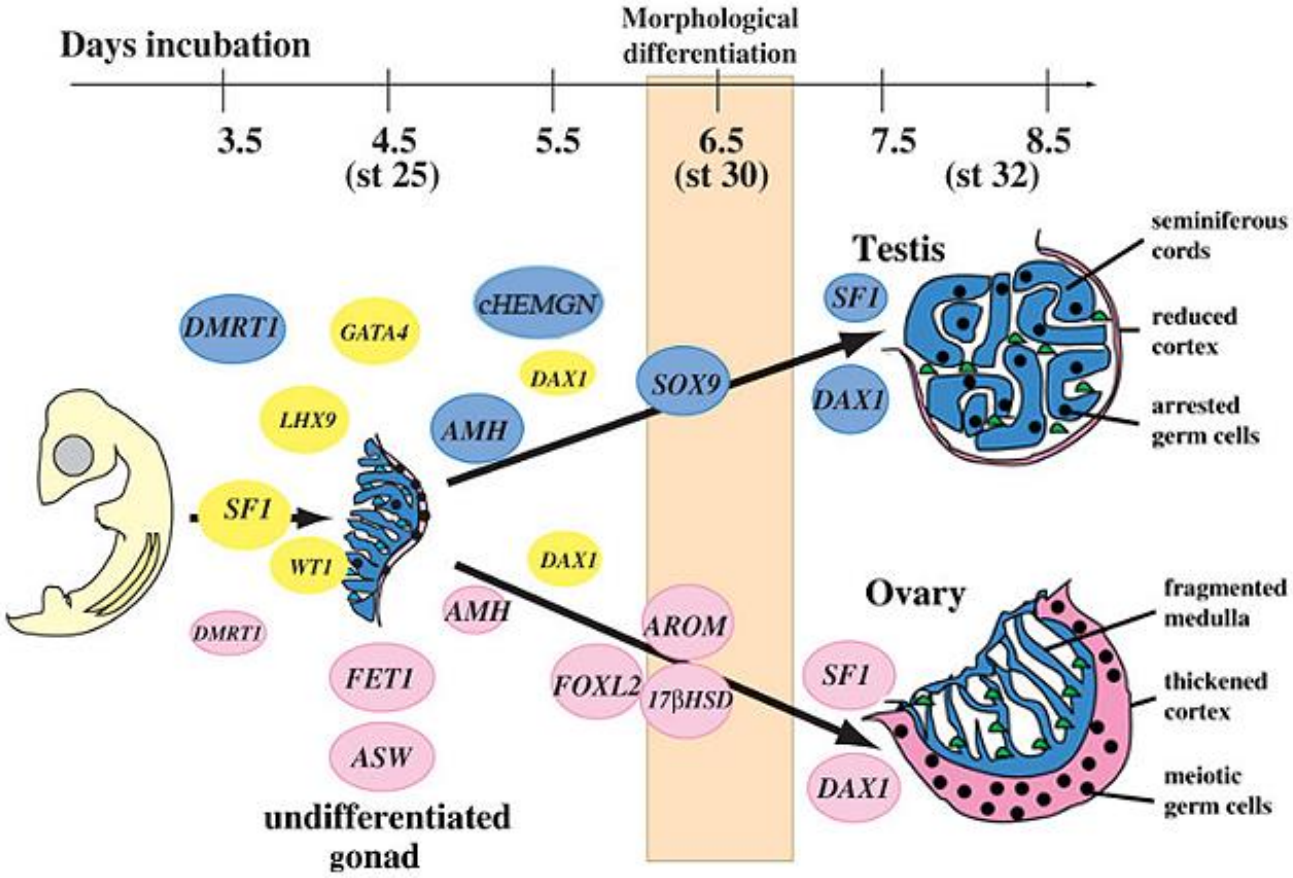


Figure 1. 2. Avian sex differentiation and determination. Similar factors to those seen in mammals are involved in the development of the undifferentiated gonads; *LHX9, WT1* and *SF1*. *DMRT1* is expressed dimorphically; higher in males than females. In contrast to mammals, expression of AMH and SOX9 is reversed. In the ovary, estrogen mediates the differentiation of the cortex; *FOXL2*, aromatase and *12βHSD* mediate the synthesis of estrogen, which in turn binds to *ERα* expressed in the cortex. Blue and pink circles indicate factors involved in the male and female pathways respectively. The size of the circles represents relative expression compared to the opposite sex. Yellow circles represent factors expressed at similar levels in both sexes. E6.5 indicates the start of gonad differentiation, with development of male and female

1.4. Unanswered questions in avian sex determination

Even though the downstream factors in avian male and female sex determination is largely known, the main factor conferring primary sex identity is still unknown. There are three hypotheses currently being considered: 1) Z dosage, 2) the presence of a W determinant or 3) a combination of both.

1.4.1. Z dosage vs. W determinant vs. Combination

Z dosage. The first hypothesis is based on the idea that male (ZZ) development is triggered by a double dose of Z-linked gene/s whereas female (ZW) development is determined by a lower dose. Unlike mammals where one of the X chromosomes is silenced to equalise gene dosage between males and females[49], no global dosage compensation has been observed in birds. Transcriptional activity from the Z chromosome has indicated uncompensated expression from the Z chromosomes[50]; however equalized Z-linked mRNA has suggested dosage compensation post-transcriptionally[51]. Additionally, dosage compensation has been suggested to occur on a local gene-to-gene basis that can vary by tissue and age[52]. Nevertheless, dosage compensation in birds does seem to be less effective than in mammals[53]. One gene that conforms to dosage compensation is *DMRT1*. As discussed previously, expression is higher in males than females[54]. *DMRT1* has no gametologue on the W chromosome, and with only one copy in females, a lower dose is expressed. One possible mechanism contributing to the lower *DMRT1* expression level in females is the presence of the Male Hyper Methylated region (*MHM*). *MHM* is located on the Z chromosome and consists of a *Bam*HI tandem repeat sequence. In males the *MHM* is highly methylated and not expressed, while in

females *MHM* is hypomethylated and transcribed into high molecular weight non-coding RNA (ncRNA)[55]. The ncRNA accumulates upstream, at the site of transcription, of the *DMRT1* klocus and is suggested to reduce *DMRT1* expression in females during development. *HMH* expression is present not only in the gonads but in the heart, brain, limb and brachial arches in the female embryo[56]. *HMH* therefore does play a bigger role in sexually dimorphic development, as shown during mis-expression where most notably altered asymmetric ovarian development is seen[56].

W determinant. The second hypothesis is the presence of a W-linked female determinant, similar to the *SRY* in mammals. In birds there is no *SRY* orthologue and all established components of the gene regulatory network are autosomal. None of these can therefore be the primary determinant. To date 26 bona fide genes have been annotated to the W chromosome, however, all except for *FAF* and *FET-1* have gametologues on the Z chromosome and are unlikely to act as a primary determinant[57]. Candidate genes have been identified and debunked, including *HINTW*. The Histidine triad Nucleotide binding protein W (*HINTW*) does have a Z gametologue, however, it lacks the histidine triad motif that confers enzymatic function that is present in the Z gametologue (*HINTZ*)[58,59]. Expression was observed throughout gonad development; from gonad formation (E3), gonad differentiation (E6-7.5) and up to hatching[59]. It is hypothesized that *HINTW* forms a heterodimer with *HINTZ*, thereby inhibiting the biological function of *HINTZ* to possibly contribute to testes development[59]. Mis-expression of *HINTW*, however, did not affect testes development nor did it affect the normal asymmetric ovarian development[60]. *HINTW* therefore does not seem to play a dominant role in gonad differentiation while the role of *HINTW* in overall female development is yet to be investigated.

Combination of Z dosage and a W determinant. Evidence for this is shown in triploid ZZW birds. ZZW chickens develop as intersexes with right testis, left ovotestis and a female phenotype at hatching. During maturation the ovotestis degenerates to develop into a functional testis. This suggests the presence of a W-linked factor that contributes to female development, however this factor is likely recessive and can be over ridden by the double dosage of two Z chromosomes[61]. Aneuploids could also be insightful to the possibility of a combination Z and W factors influencing sex determination, however no ZO chicken has been documented. It has been hypothesized that these genotypes are lethal due to incompatibilities of the resulting dosage of one Z chromosome in the absence of a W chromosome[62], further lending support to the importance of the W chromosome and the potential W-linked factor during sex determination.

1.4.2. Cell autonomous sex identity

Cell autonomous sex identity (CASI) refers to the inherent sexual identity of somatic cells. The true extent of CASI can be observed in gynantromorphs, male/female chimeras. These birds are bilaterally asymmetrical; one side of the bird is male and the other female with the genotype reflecting the phenotype of each side respectively. Unlike mammals, circulating hormones secreted from the developing gonads do not have an effect on the somatic cells or the secondary sex structures[63]. This is apparent from gynantromorphs, where both sides of the bird are exposed to the same circulating hormone composition. Evidence for this is shown in transplant experiments between opposite sex and same sex donor/recipient embryos. Same

sex transplant tissue displays the ability to interpret and respond to developmental signals, developing secondary sex structures of the recipient embryo. Opposite sex transplant tissue retains its original sexual identity and develops secondary structures associated with the donor sex[63].

It is unclear whether CASI is a result of single gene action or a network of interacting genes. These gene/s are likely to be dimorphically expressed during development and so far up to 300 genes have been hypothesized to be CASI genes[64]. The relationship between CASI and the genes regulating gonad differentiation remains unknown. The accepted male gonadal determinant *DMRT1* cannot act as a CASI gene since expression its restricted to the urogential system and it is not expressed prior to the initiation of gonad differentiation[64]. It has been observed that CASI can be temporarily overridden, at least in the gonads, where knock down of *DMRT1* and aromatase will feminise[28] and masculinise[65] the gonads respectively. These tissues, however, revert back to their genotypic form and the majority of somatic tissues do not exhibit phenotypic changes during sex-reversal.

1.5. *FET-1* as a candidate female determinant

Female expressed transcript-1 (*FET-1*) is an ancient retroviral integration that is predicted to encode a 434 amino acid protein with a trans-membrane domain, containing a conserved domain of the subfamily ENVV1-like HR1-HR2 that forms part of the Ebola HIV-1-like HR1-HR2 superfamily. This domain spans the transmembrane subunit of various endogenous retroviruses[66].

The *FET-1* was first isolated from a differential screen of male and female embryonic urogential ridge cDNA library. Reed and Sinclair, 2002 has shown that *FET-1* expression is female specific and observed in the gonads, waste collection system and the nervous system. Using *FET-1* cDNA as a probe an integration on the W chromosome was detected using fluorescent in situ hybridisation (FISH)[67]. Digoxigenin (DIG) whole mount *in situ* hybridisation showed initial expression at E3 when the gonads are first seen as discreet structures to be similar between males and females[67]. Asymmetrical expression is seen from E4.5 till E6 with higher expression observed in the left ovary, after which expression sharply decreases at E6.5. Within the ovary, *FET-1* expression was observed to be higher in the cortex where folliculogenesis and oogenesis take place.

1.6. Significance of this work

This work carries significance in two respects: 1) it provides insight into the evolution of avian sex determination and 2) can contribute to a solution for the ethical and financial difficulties associated with egg-laying chicken breeds in the poultry sector.

Gallus gallus (chicken) are warm-blooded vertebrates that share the majority of their biology with mammals; they develop as amniotes and their development resembles that of mammals. This implies the presence of conserved key developmental pathways. One of the major differences between mammals and chickens is sex determination and development. Birds represent a “transitional state” in which they

use GSD, sharing some genes with mammals, while retaining some qualities of lower vertebrates. One of these qualities is the use of hormones (e.g. estrogens) during gonad development[4], whereas developing mammalian gonads are resistant to exogenous hormones. Clarification of the avian sex-determining pathway provides valuable information on the evolution of sex-determining mechanisms and the transition between lower vertebrate mechanism utilizing hormones and those of higher vertebrates.

Two types of chicken breeds are used in the poultry sector: the meat breed (broilers) and the egg-laying breed. Within the former both males and females are bred and raised for meat supplies, while culling (maceration or asphyxiation) of the male hatchlings is employed within the latter since only females are needed for egg production. This raises both ethical and financial problems, highlighted in several countries during the last few years. Genetic engineering of an egg-laying breed to produce only females, with subsequent increases in egg production, would contribute to eliminate the ethical and financial issues associated with this practice. This however requires a detailed understanding of the avian sex-determining pathway.

1.7. Hypothesis, aim and objectives

1.7.1. Hypothesis

It is hypothesised that *FET-1* plays a female-specific role during sex determination and gonad differentiation and may therefore act as a possible primary female determinant.

1.7.2. Aim and objectives

The aim of this project was to determine if *FET-1* acts as an ovarian determinant in *Gallus gallus.*

Objective 1: Confirm the female-specific expression of *FET-1* by RT-PCR and RACE.

Objective 2: Determine the expression patterns of various *FET-1* associated RNAs by *in situ* hybridization

Objective 3: Investigate the presence and expression pattern of FET-1 protein.

CHAPTER 2 - MATERIALS AND METHODS

2.1. Bioinformatics

FET-1 analysis

The *FET-1* cDNA sequence (AY113681.1) was obtained from the National Center for Biotechnology Information (NCBI), as part of the Galgal5.0 genome assembly. Basic Local Alignment Search (BLAST)[68] was performed in Ensembl against the nucleotide database, since the chicken genome seems to be better annotated in this database.

Open reading frames were predicted in the annotated *FET-1* genomic location using NCBI Open reading frame (ORF) finder (https://www.ncbi.nlm.nih.gov/orffinder/).

The non-coding RNA (ncRNA) populations were identified using BLAST against the Ensembl ncRNA database.

Predicted protein analysis

The FET-1 predicted protein was analysed using NCBI Smartblast (https://ncbi.nlm.nih.gov/smartblast/), a local alignment tool that compares a protein query against a landmark database consisting of 27 genomes across a wide taxonomic range.

Sequence alignments

Pairwise sequence alignments were performed using the Wilbur-Lipman method in MegAlign (DNAstar, Lasergene). This is a word method; splitting the sequence/s into k-mers which are used to find region of similarity between two or more DNA sequences.

Primer design

All primers were designed using PrimerQuest from Integrated DNA technologies (eu.idtdna.com/PrimerQuest/). Additionally, all secondary structures and their ΔG (Gibbs free energy) were checked to ensure linearization of any secondary structures (hetero- and self- dimerisation) that could impede binding during the annealing stage. Specificity of the primers were checked using PrimerBlast (NCBI).

2.2. Embryo collection and dissection

Embryo collection and incubation conditions

Fertilized, unincubated eggs were ordered from the Agricultural Research Council Farm (Irene, South Africa) and specific pathogen-free White Leghorn embryos were supplied by Avi-Farms (Centurion, South Africa). Once collected, the eggs were equilibrated at room temperature (~25˚C) overnight and placed in the incubator the next day. This is due to the refrigeration of the eggs upon selection/before delivery to deter development of the embryos. From experience, immediate incubation upon delivery drastically decreases the viability of the eggs; development not taking place

or stunted. The embryos were incubated at 37°C until desired developmental stages. To note is the natural slight variation between the development of the individual embryos, resulting in some embryos being slightly younger or older that the desired developmental stages.

All animal experiments were carried out according to the rules and regulation of the Wits animal ethics committee (WITS Ethics permit number: 2017/01/01/0). Approval for the use of genetically modified organisms was obtained from the Department of agriculture, forestry and fisheries (DAFF) (Registration number: 39.2/Witwatersrand – 16/120).

Embryo dissection

After appropriate incubation, the eggs were windowed to visualise and determine the developmental stage of the embryo according to Hamburger Hamilton staging[69]. Embryos developed to the approximate desired stage were removed from the eggs and dissected in phosphate buffered saline (PBS). All embryos younger than the desired stage were placed back in the incubator; PBS was added to the egg to prevent the embryo from drying out and the windowed egg sealed with sellotape. These embryos were dissected when appropriate developmental stage was reached. All dissection instruments were sterilized between embryo dissections using 70% ethanol and then flamed over a Bunsen burner. Additionally, the PBS was replaced between individual dissections to prevent DNA cross contamination. All instruments

used for RNA work were treated with RNaseZAP (Sigma-Aldrich) to inhibit RNAses and prevent RNA degradation.

Dissection for RNA and DNA extraction

The embryos were incubated to embryonic day (E) 4.5. The vitelline membrane surrounding the embryo was removed to increase accessibility after which any unnecessary structures, such as the allantois, gut and heart were removed. The head (anterior) was removed and stored in PBS for DNA extraction (2.3) and PCR sexing (2.4). The gonads/kidneys, located on either side of the pre-vertebra was removed, homogenized and stored in the RA1 buffer (NucleoSpin RNA, Macherey-Nagel) at -80˚C for subsequent RNA extraction (2.3). All female and male gonads/kidneys were pooled respectively to obtain sufficient RNA working concentrations.

Since the gonads are microscopically small, fragile at the desired stages required for the project (E3-E6.5) and attached to the developing kidneys, they were dissected together and used for DNA/RNA extraction.

Dissection for In Situ Hybridization (ISH) and Immunohisto Chemistry (IHC)

The embryos were incubated to E3.5, E4.5 and E6.5. Males and females were identified using PCR sexing (2.4). Embryos were dissected as described above;

however, the gonads were not removed and were processed according to the needs of the method used (see below for ISH and IHC methods)

Dissection for crude protein extraction

Embryos were incubated to E6.5. PCR sexing was performed as described in 2.4. Four tissue groups were dissected: the limb, prectoral muscle, gonad and kidney. The tissue was pooled and used for crude protein extraction (2.8).

2.3. DNA and RNA extraction

DNA used for all PCR work excluding PCR sexing was extracted using the NucleoSpin® Tissue kit (Macherey-Nagel) according to the manufacturer's protocol. PCR sexing was performed on crude DNA extractions. Briefly, tissue, approximately 3mm $_x$ 3mm was collected during embryo dissection. A crude extraction buffer was made according to Li et al. 2011 consisting of Tris-HCl, EDTA, NaCl and proteinase K (See Chapter 7 part A for detailed recipe). The tissue was incubated in the buffer for 30 min at 55℃ and 5 min at 95˚C to inhibit the proteinase K. The samples were vortexed before, after as well as 15 min into the 55˚C incubation. The samples were centrifuged at maximum speed for 1 minute, after which the supernatant was used for PCR sexing (2.4).

RNA extraction was performed using the NucleoSpin RNA kit (Macherey-Nagel). All tissue was homogenized and stored at -80˚C in the RA1 buffer supplied in the kit if extraction was performed at a later time.

2.4. PCR sexing

PCR sexing was performed according to Clinton *et al*. 2001. This method utilises the *Xho*I repeat sequence on the W chromosome to distinguish between males and females, as well as the 18S ribosomal sequence as an internal control. Two primer sets were used (Table 2.1), amplifying a 415 bp product of the *Xho*I repeat and a 256 bp product of the 18S ribosomal gene. Females presented with two PCR products, at 415 (*Xho*I) and 256 (18S) and males with one at 256 (18S).

Table 2. 1. Clinton *et al*. sexing primers. Two primer pairs designed in the W *Xho*I repeat sequence and the 18S sequence to distinguish between females and males; females with two PCR products and males' one.

	Forward	**Reverse**
***Xho*I**	5´CCCAAATATAACACGCTTCACT 3´	5´ GAAATGAATTATTTTCTGGCGAC 3´
18S	5´ AGCTCTTTCTCGATTCCGTG 3´	5´ GGGTAGACACAAGCTGAGCC 3´

2.5. PCR, RT PCR and Rapid Amplification of cDNA Ends (RACE)

Sequences for all primers used in this work are listed in Table 2.3. These include: 1) primers for semi-nested touchdown genomic PCR (Table 2.2) performed to confirm

the location of *FET-1*. Using semi-nested PCR increased specificity for *FET-1*, with the 5' primer overlapping both *FET-1* and the overlapping gene *MTMR7* on chromosome 4 during both PCR reactions. The 3' primer was only specific to *MTMR7*, ensuring amplification form chromosome 4. 2) Primers for standard PCR of the predicted ORFs in the proposed *FET-1* location. 3) Primers for cloning of the DNA templates for *in situ* hybridization riboprobe synthesis

Table 2. 2. MTMR7 Touchdown PCR and *FET-1* semi-nested PCR cycle conditions

MTMR7 Touchdown**		
Phase I	Temperature (˚C)	Time
Denaturation	95	5 sec
11 Cycles		
Denaturation	95	30 sec
Annealing	72 - 61*	45 sec
Elongation	72	4.5 min
Phase II		
21 cycles		
Denaturation	95	30 sec
Annealing	60	45 sec
Elongation	72	4.5 min
Final elongation	72	10 min
Hold	4	∞

FET-1 Semi-nested		
	Temperature (˚C)	Time
Denaturation	95	2 min
Annealing	60	45 sec
Elongation	72	2.5 min
Final elongation	72	10 min
Hold	4	∞

* Annealing temperature 10˚C above calculated Ta, decreasing by 1 ˚C with each cycle.

** Touchdown PCR occurs over two phases. During the first phase the annealing temperature (Ta) is set 10˚C above the calculated Ta and decreased by 1˚C every cycle until the desired Ta is achieved. This ensures specific binding of the primers to the target sequence as the temperature decreases and preferential amplification of the target sequence. The second phase consists of standard PCR, during which the target sequence will be preferentially amplified, being in greater volume to non-specific products

RT-PCR was performed using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) on RNA extracted from E4.5 male and female embryos. Random hexamers were used for reverse transcription after which gene-specific *FET-1* primers (Table 2.3) were used for PCR amplification of the cDNA.

The GeneRacer Kit with SuperScript III RT and TOPO TA Cloning Kit for Sequencing (Thermo Fisher Scientific) was used for RACE. Reverse transcription was performed with Oligo-dT primers supplied with the RACE kit and PCR performed with gene-specific primers (GSP) in combination with the standard primers supplied with the RACE kit. *FET-1* GSP was designed according to the RACE manual (Table 2.3). Briefly, primers were designed with a 50% - 70% GC content to obtain high annealing temperatures. The high annealing temperature in combination with primers 23-28 nucleotides in length increases binding specificity.

Table 2. 3. Primers for all genomic PCR, RT-PCR and RACE applications

Primer	**Forward (F 5' → 3')** **Reverse (R 5' → 3')**	**Product size (bp)**
Semi-nested genomic PCR primers		
MTMR7	F 5' AGTCTAGGGGATCACCGACC 3' R 5' AGCTTCCAGCCTTTCTCTCGTTC 3'	4500
FET-1 semi-nested	F 5' AGTCTAGGGGATCACCGACC 3' R 5' TCACTGCGAGAGCAGACAAG 3'	2487
***FET-1* 5' and 3' UTR RT-PCR primers**		
FET-1 UTR primers	F 5' CTGATTTCCGCACGTCAGTTT 3' R 5' CCCTTTCTGTGCTGTGCATAAG 3'	432 or 2436*

Table 2.3 *Continued*

Predicted ORF primers			
ORF1		F 3' GAAGTACTGTAATCAGTGGTGGTC 5' R 3' GCCGCTCCTTCTCTAAAGTTAATA 5'	245
ORF2		F 3' GATAAGTTAAGGGACGCCATGAG 5' R 3' TGCTTCCTGTAGGACTGCTATTA 5'	252
ORF3		F 3' CATGTGTCAGGATACCCACTATATC 5' R 3' CTGAGATTTCTCGGAAGGACTC 5'	217
RNA control primer set			
18S		F 5´ AGCTCTTTCTCGATTCCGTG 3´ R 5´ GGGTAGACACAAGCTGAGCC 3´	256
ISH riboprobe primers			
S1	240-727	F 5' GGACCCTACCAAGTTCTCCT 3 R 5' TGGTCCCTTTTGTGGTTTGTC 3'	400
S2	728-1640	F 5' CGAAACTTGACGGGCTTAGG 3' R 5' GCACTTGATCATCACACAGGT 3'	867
S3	1641-2293	F 5' AAGGCAAAAGGTGGAAAGCG 3' R 5' CCACCAAACTTACACACCGT 3'	610
RT-PCR and RACE primers			
FET-1 GSP		5' TGGCTATGTGGAGACGGCCAGGCACGAA 3' 5' TCGGCACAGGCAGGCAGAAGCACTTGA 3'	658
RACE primers**		5' CGACTGGAGCACGAGGACACTGA 3' 5' GCTGTCAACGATACGCTACGTAACG 3'	-

*The length of the amplicon depends on whether *FET-1* is expressed from chromosome 4 and the annotated intron is spliced to produce a 432 bp transcript. The 2463bp transcript will be amplified if *FET-1* is not spliced and expressed from elsewhere in the genome

**RACE primers are used in combination with *FET-1* gene-specific primers (GSP). The forward RACE primer is binds to the RNA oligo that replaces the mRNA 5' cap. The RACE reverse primer binds to a priming sequence added to the 3' end during reverse transcription. See Figure 3.6 for RACE mechanism.

2.6. PCR product cloning

The RACE PCR product was cleaned using the GeneJET PCR Purification kit (Thermo Fischer Scientific) to remove primer dimers in preparation for ligation. Primer dimers are small enough to be preferentially ligated above any larger PCR products present. The purified RACE product was ligated into the pGEM®-T Easy vector system (Promega). The recombinant plasmid was subsequently used to transform One Shot TOP10 Chemically Competent *Escherichia coli* (E.coli) cells (Invitrogen), which were grown overnight at 37˚C on LB Agar plates with 100 µg/ml ampicillin.

Multiple colonies were selected and individual 10 ml liquid cultures were grown in LB media with 100 µg/ml ampicillin overnight on a shaker incubator at 37˚C. From the liquid culture 6 ml was used for plasmid extraction using the Zymo Plasmid miniprep kit (Zymo Research). The extracted plasmids were sent for sequencing (Inqaba biotec^{TM}) using the T7 and SP6 universal primers.

2.7. *In situ* hybridization (ISH)

Embryo collection

Two male and female embryos (for the sense and anti-sense probes) at E3.4, E4.5 and E6.5 were dissected to expose the genital ridge (E3.5)/gonads (E4.5 and E6.4). The head was used for DNA extraction and subsequent PCR sexing as detailed above. The embryos were fixed in 4% paraformaldehyde (PFA) overnight (O/N) at

4˚C and treated with 10 µg/ml proteinase K for 30 min at RT (See Chapter 7 part C for the detailed protocol).

Probe design and synthesis

DIG-labelled RNA probes (riboprobes) were designed against 3 segments corresponding to the three different *FET-1*-like lincRNA populations, referred to as S1, S2 and S3 (Chapter 3). Primers against the three segments were designed using the IDT primer design tool (eu.idtdna.com/PrimerQuest/) (Table 2.3). Standard PCR was performed and the PCR products of the three segments were cloned into pGEM®-T Easy vectors (Promega). The recombinant vectors were used to transform One Shot TOP10 chemically competent E.coli cells (Invitrogen). The transformed cells (100 µl) were spread on agar plates containing 100 µg/ml AmpIcillin and grown at 37˚C, overnight (~16 hours). Three to five single colonies were selected for each transformant and colony PCR was used to confirm the presence of the desired PCR (S1, S2 and S3) insert. The colonies that contained the desired insert were used to make a liquid culture in LB media containing 100 µg/ml Ampicillin. The cultures were grown overnight at 37˚C on a shaker incubator. The plasmids were isolated using the Zymo Plasmid Miniprep kit and sequenced (Inqaba biotec). The sequencing results were used to 1) confirm the sequence via a local alignment using BLAST against the NCBI refseq_RNA database to ensure specificity to *FET-1*; and 2) to confirm the sequence and identify the direction of ligation. All inserts were ligated into pGEM®-T Easy in a forward orientation in relation to *LacZα*. All sense probes were therefore synthesised using T7 RNA polymerase and all anti-sense probes using SP6 RNA polymerase. The plasmid DNA (10µg) was linearised; plasmids containing the sense and anti-sense probes with *Nde*I and *Nco*I respectively (Figure

2.1). Linearised plasmid (1µg) and T7/SP6 RNA polymerase probe synthesis mix was used for the RNA probe synthesis, performed as per the protocol in Chapter 7, part B. DTT (0.1M) was added to the SP6 RNA polymerase to increase activity. RNAse-free DNAse was used to remove the plasmid DNA after which the RNA probes were purified and concentrated using the Zymo RNA clean and concentrator kit (Zymo research). The probes were stored 1:1 in hybridization mix at – 20˚C.

Controls

cSOX9 sense and anti-sense probe templates were prepared from pBS-*cSOX9*, where *cSOX9* was ligated in reverse orientation with regards to *LacZα*. The sense probe template was prepared using *Xba*I and the anti-sense probe template using *Xho*I. This was used as positive and negative controls respectively. A “no probe” control was included as a negative control for any background staining and trapping of the staining solution.

Staining

NBT/BCIP (Roche) was used for colour development. The *cSOX9* sense probe was incubated for 45 minutes. The S3 sense and anti-sense was stained for 6 hours and the remaining probes for 48 hours. Staining was terminated when the no probe control started to show colour development.

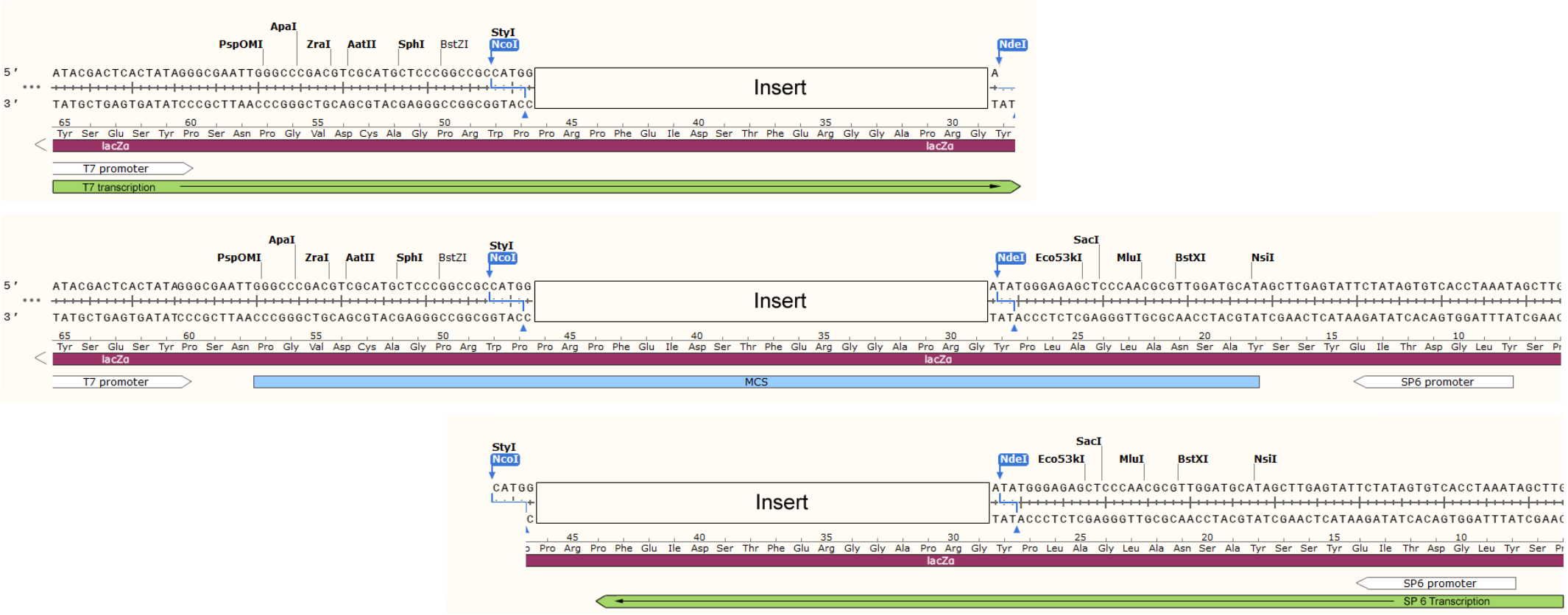


Figure 2. 1. Plasmid linearization for riboprobe synthesis. Middle panel, pGEM®-T Easy plasmid multiple cloning site (MCS) with the *Nde*I and *Nco*I restriction sites flanking S1/S2 and S3 *FET-1* inserts. Top panel, *Nde*I plasmid linearization for sense probe transcription. Bottom panel, *Nco*I plasmid linearization for anti-sense probe transcription.

2.8. Immunohistochemistry (IHC)

IHC was performed on both whole mount embryos and tissue sections, due to uncertainty of staining patterns observed during initial attempts of IHC.

E4.5 male and female embryos were dissected (as detailed in 2.2) and PCR sexing performed (2.4). The embryos were fixed O/N in 4% PFA at 4˚C.

Whole mount

The embryos were processed from 4% PFA to PBS and stored at 4˚C in preparation for IHC.

Tissue sections

The embryos were processed from 4% PFA to $TBST_xD$. Next the embryos were cryoprotected via a sucrose gradient: 5% for 6 hours and 15% overnight, washed and processed to Cryomatrix (Thermo Fischer scientific). They were incubated in Cryomatrix (Thermo Fischer scientific) overnight at 4˚C and flash frozen in liquid nitrogen in the appropriate molds. Sections (10 µm) were obtained using a Bright cryostat/microtome 3050.

Whole mount and tissue sections

Embryos/ tissue sections were washed in $TBST_xD$ and endogenous peroxidise activity inhibited with incubation in 0.3% H_2O_2 in methanol for 20-30 minutes at room temperature. Indirect IHC was performed (See Chapter 7, part D) using a custom-

made polyclonal rabbit anti-FET-1 primary antibody synthesised by GenScript. The amino acid sequence (Uniprot: Q8JGM1) used for antibody synthesis:

MYRSLIIIGMIGTVVTAWQENSYLRMTEKIGQTLNRSCWICTALPRGEGRVPVYGIV
VLNWTIPEWVEHGKCKFGIEDKPQKGPKFDLLVINSSRSIFVDRKVNGKDGVGVGW
RNLTGLGCSWEYGNFRAISEVSEAGRWNQGEGCDPRIDKESYWASDKNILGIVGM
GASFTHQENMARPCKLPHGYWWLCGDGQARKALPLNWTGTCTTGYLIPQTTVHE
EIPGGLLRTPWIRTKHTYNPLAIYNKGYHSFLRALIPALGVAQLEKAILNISATLEIVEN
ATTDALRAIQEEVSSLSKVVLQNRMALDLLTAKEGGVCTIINQSCCAYINKDLRIETD
LRKIWEQTKVLHRTTQDDTN

The antibody was used at a concentration of 10 µl/ml and incubated at 4˚C overnight. A secondary polyclonal Goat anti-rabbit HRP conjugated antibody (Dako) was used at a 1:2000 dilution for detection. A pre-immune serum control was included as well as a no primary antibody negative control along with a Pax7[72] positive control. All embryos were incubated for 20 minutes for colour detection using 3, 3'-Diaminobenzidine (DAB, Sigma-Aldrich).

2.9. Crude protein extraction

A crude protein extract was prepared from tissue dissected from E6.5 males and females (2.2) in 2X Laemmli sample buffer (65.8 mM TRIS pH 6.8, 26.3% glycerol, 2.1% SDS, 10 mM DTT and β-mercaptoethanol). The tissue was freeze/thawed; flash frozen in liquid nitrogen and immediately thawed (approximately 2 min) at 95˚C, to fragment the cells. The samples were vortexed and centrifuged at max speed and the supernatant used as the crude extract. The crude extract was stored at -80˚C for later use.

2.10. Dot Blot

The dot blot was performed using 0.45 µm nitrocellulose membrane. Ten micrograms of the crude protein extract was blotted onto the membrane and allowed to dry. Additionally, a FET-1 recombinant protein (Genscript) was blotted as a positive control as well as a Pax7 positive control. A pre-immune serum negative control was also included. The 0.45 µm nitrocellulose membrane was washed in PBS_xT and blocked in 3% BSA for 2 hours at 37˚C in a shaker incubator. The crude extract and recombinant FET-1 protein was incubated in the custom-made FET-1 polyclonal primary antibody, 0.2 µg/ml (Genscript), and the pre-immune sample incubated in 0.2 µg/ml pre-immune serum at 37˚C for 2 hour in a shaker incubator. The samples were incubated in a secondary polyclonal Goat anti-rabbit HRP conjugated antibody (Dako), used at a 1:2000 dilution, at 37˚C for 2 hours in a shaker incubator. The membranes were incubated for 10 minutes in DAB and 3% H_2O_2 for colour detection.

2.11. SDS Page and Western blot

The crude protein extract (2.8) was separated on two 12% SDS-PAGE gels, as per the protocol in Chapter 7, part E. The running voltage used was 120V for the stacking gel and 180V for the separating gel.

Coomassie blue staining

The first SDS-PAGE gel was stained with Coomassie blue (0.1% Coommasie R250, 10% glacial acetic acid, 40% methanol) for 1 hour at room temperature with gentle

agitation; until gel was indistinguishable from staining solution. De-staining solution (10% glacial acetic acid, 20% methanol) was added to the gel container and microwaved for 10 second intervals, taking care not to boil the de-staining solution; this was continued until the gel was transparent. The de-staining solution was replaced once saturated with the Coomassie stain.

Western blot

The second SDS-PAGE gel was used for capillary blotting[73] (as per the protocol in Chapter 7, part E) onto a 0.45 µm nitrocellulose membrane. Briefly, filter paper was wetted in transfer buffer (25mM Tris, 192mM glycine and 10% methanol) and placed onto a raised platform with the ends dipped in buffer reservoirs on both sides of the platform to ensure saturation throughout the blotting period. The nitrocellulose membrane and SDS-PAGE gel was equilibrated in the transfer buffer for 15 min. The gel was placed onto the filter paper bed. The nitrocellulose membrane was place on top and bubbles between the layers removed. Multiple layers (4-7) dry 3MM filter papers were placed onto the nitrocellulose membrane and a heavy weight placed on top. Blotting was performed overnight.

The blotted membrane was washed in PBS_xT (PBS with 0.1% Tween20) and blocked with 3% BSA in PBS_xT for 1 hour at room temperature with gentle agitation. The membrane was washed and incubated in the custom made FET-1 antibody (2.7) overnight at 4˚C. Excess antibody was washed off and incubated in the secondary HRP-linked goat anti-rabbit antibody (Dako, 1:2000 dilution) for 1 hour at room temperature with agitation. Staining was performed with 0.5mg/ml DAB in PBS for 15 minutes.

CHAPTER 3 – RESULTS

3.1. Partial *FET-1* integration present on chromosome 4

At the start of this project, bioinformatics investigation of *FET-1* revealed multiple partial integrations in the Ensembl Galgal4.0 genome assembly in both autosomal and sex chromosomes. Shortly after, the Galgal5.0 genome assembly was released showing a different set of integrations. In NCBI, *FET-1* was annotated to chromosome 4 as a gene consisting of two exons. However, this record was recently withdrawn due to insufficient evidence to define a distinct gene. The initial report of *FET-1* by Reed and Sinclair, 2002 localized *FET-1* to the W chromosome using fluorescent DNA *in situ* hybridization, using a probe synthesized against the full length *FET-1* cDNA[67]. Given the multiple integrations and the *FET-1* localization of *FET-1* by Reed and Sinclair, 2002, it was decided to reinvestigate the localization of *FET-1* in the genome.

FET-1 localization on chromosome 4

The *FET-1* cDNA contains a 1305 bp coding sequence (CDS), predicted to encode a protein of 434 amino acids (Figure 3.1a and b). A pairwise alignment using the Wilbur-Lipman method (MegAlign, DNAstar) of the *FET-1* cDNA to the proposed chromosome 4 location (NC_006091.4, 63515418 – 63518160, Galgal5.0, NCBI) aligned the 5’ and 3’ UTR of the *FET-1* cDNA to the two annotated exons. The *FET*-1 CDS, however, was absent from chromosome 4 (Figure 3.1c, Appendix 3 Figure 6A2.1 for sequence alignment).

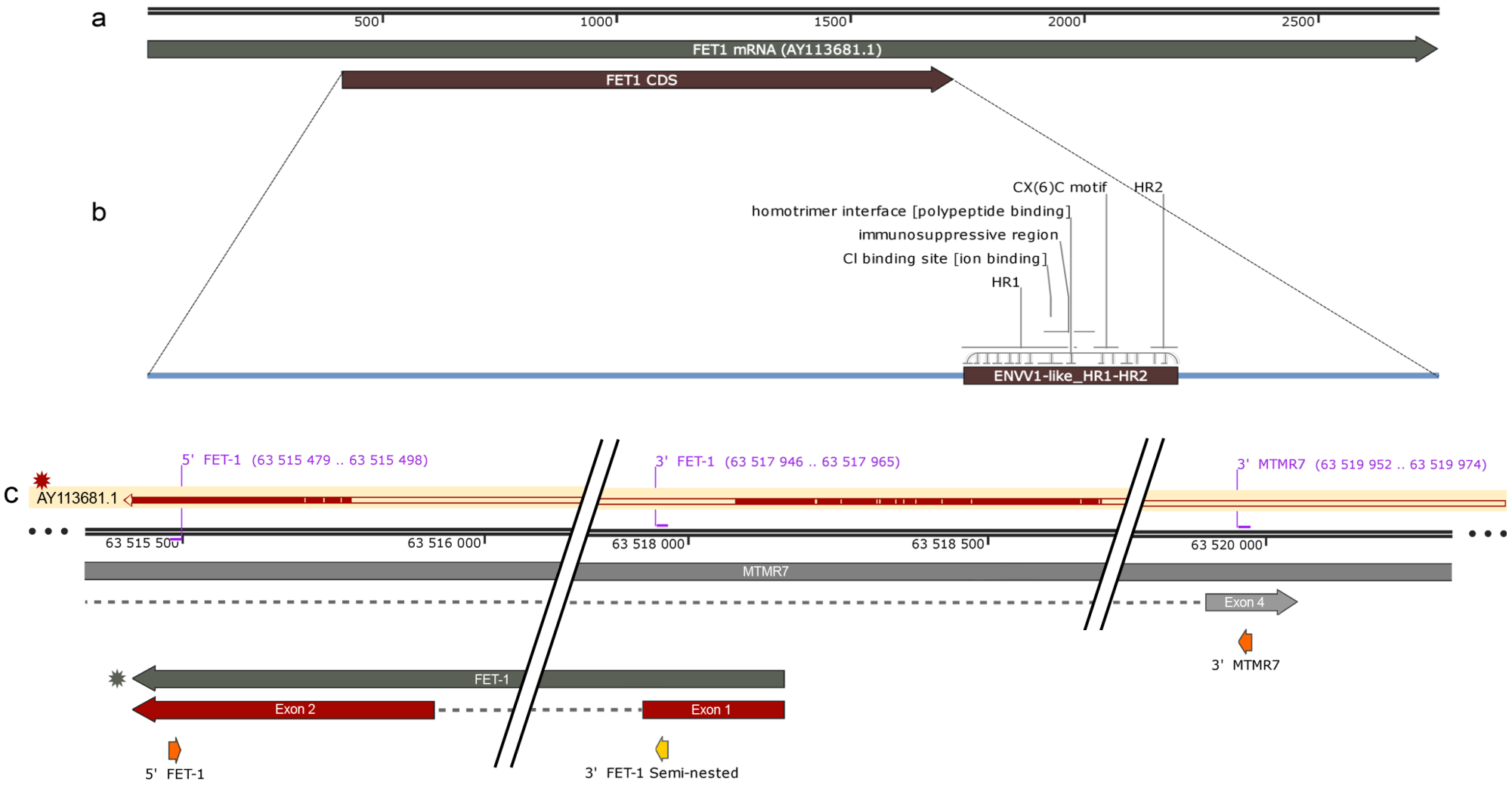


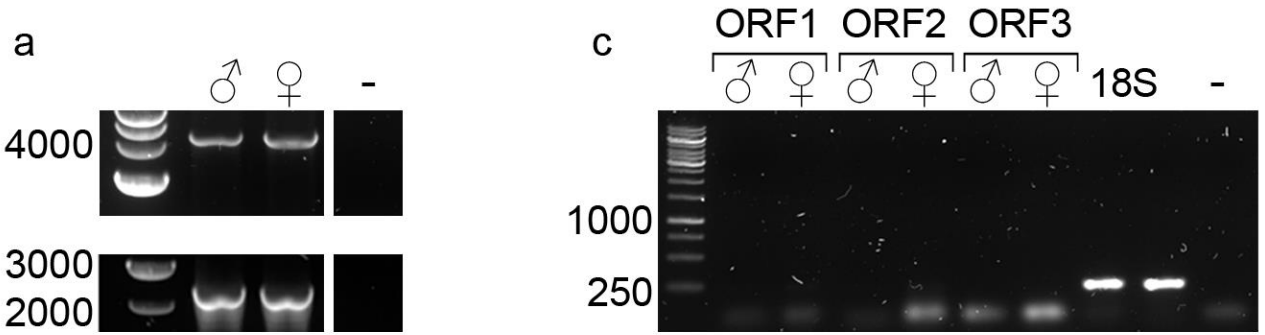
Figure 3. 1. Bioinformatics analysis of *FET-1*. a) The *FET-1* cDNA (AY11368.1) isolated by Reed and Sinclair, 2002. b) The predicted protein translated from the *FET-1* cDNA, containing a retroviral transmembrane domain. c) Pairwise alignment of the *FET-1* cDNA against the proposed chromosome 4 location. The topmost band: AY113681.1 (red star), represents the *FET-1* cDNA alignment to the chromosome 4 *FET-1* (grey star). The solid red regions are aligned to the proposed exons on chromosome 4 and the clear region represents the unaligned regions, including the *FET-1* CDS. The orange and yellow arrows are the primers used for semi-nested PCR, with the primer nucleotide coordinates in the Galgal5.0 genome in purple. MTMR7: the gene overlapping the proposed *FET-1* location on chromosome 4.

Experimental confirmation of the absence of FET-1 CDS on chromosome 4

To confirm the pairwise alignment and the absence of the CDS from chromosome 4, semi-nested PCR was performed on genomic DNA from male and female embryos. Primers were designed in *MTMR7*, the gene overlapping the proposed chromosome 4 location for *FET-1* on the opposite strand (Figure 3.1c; orange arrow and yellow arrows, Table 3.1). The primers were designed in such a manner that the 5' primer overlapped both the MTMR7 and *FET-1* locus while the reverse primers were specific to *MTMR7* and *FET-1* respectively; that of *FET-1* used for semi-nested PCR from the *MTMR7* PCR product. Initial PCR on *MTMR7* produced excessive background amplification with a faint amplicon at the expected size of 4500 bp. This is the transcript size expected based on the chromosome 4 genome data where MTMR7 is located. The expected product size of 4500 bp is large, shorter non-specific products were therefore preferentially amplified. To overcome this, touchdown PCR was performed[74] (Table 2.2). Genomic PCR produced the expected product size (4500bp) for the touchdown MTMR7 PCR (Figure 3.2a, top panel). After gel extraction of the MTMR7 PCR product, semi-nested PCR was performed using the *FET-1* semi-nested primer producing the expected 2487 bp product (Figure 3.2a bottom panel). Sequencing of the PCR products further confirmed the absence of the *FET-1* CDS on chromosome 4 (Appendix 2, Figures 6A2.2a and 6A2.2b).

Out of interest and to determine if 1) a possible unannotated larger, overlapping protein coding transcript is transcribed from the annotated *FET-1* sequence and; 2) if a possible variation of *FET-1* not containing the *FET*-1 CDS is transcribed, open

reading frames were predicted within the chromosome 4 annotated sequence. Three major ORFs, overlapping smaller predicted ORFs, were identified using the NCBI Open Reading Frame finder (Figure 3.2b). Primers were designed against the three identified ORFs (Table 2.3) and RT-PCR performed on E4.5 male and female extracted RNA. To keep in mind is that this will not amplify any *MTMR7* transcripts, since *FET-1* lies within an intronic region with regards to *MTMR7*. No significant PCR amplification was observed (Figure 3.3c), indicating that 1) there is no protein coding transcript transcribed and 2) no variation of *FET*-1 is transcribed from this region.



b

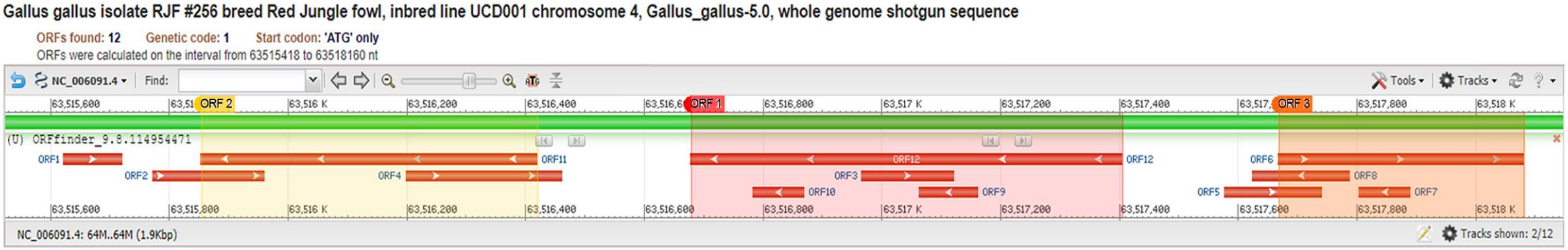


Figure 3. 2. *FET-1* placement on chromosome 4. a) Semi-nested PCR of the chromosomal 4 *FET-1* region. Top panel, MTMR7 touchdown PCR with a amplicon of 4500 bp. Bottom panel, *FET-1* semi-nested PCR performed on the MTMR7 PCR product with an amplicon at ~2500 bp. b) NCBI ORF prediction of ORFs in the *FET-1* chromosomal 4 region. The largest ORFs (highlighted regions) were selected and RT-PCR performed to determine if alternative transcripts are made from the region. c) RT-PCR of selected ORF showing non-significant amplification below the expected sizes of 245 (ORF1), 252 (ORF2) and 217 (ORF3). 18S, integrity control for the RNA used. -, No template control.

As a confirmation of the absence of *FET-1* on chromosome 4, RT-PCR primers were designed in the 5' and 3' UTRs of the *FET-1* cDNA; the regions aligning to the annotated exons on chromosome 4 (Figure 3.3a). Amplification of a *FET-1* transcript from chromosome 4 would produce a 432 bp product, since the intron would be spliced, while amplification of a 2436 bp product indicates *FET-1* amplification product across the CDS, expressed from elsewhere in the genome. Amplification of a 2436 bp transcript confirmed the absence of *FET-1* from chromosome 4 (Figure 3.3b).

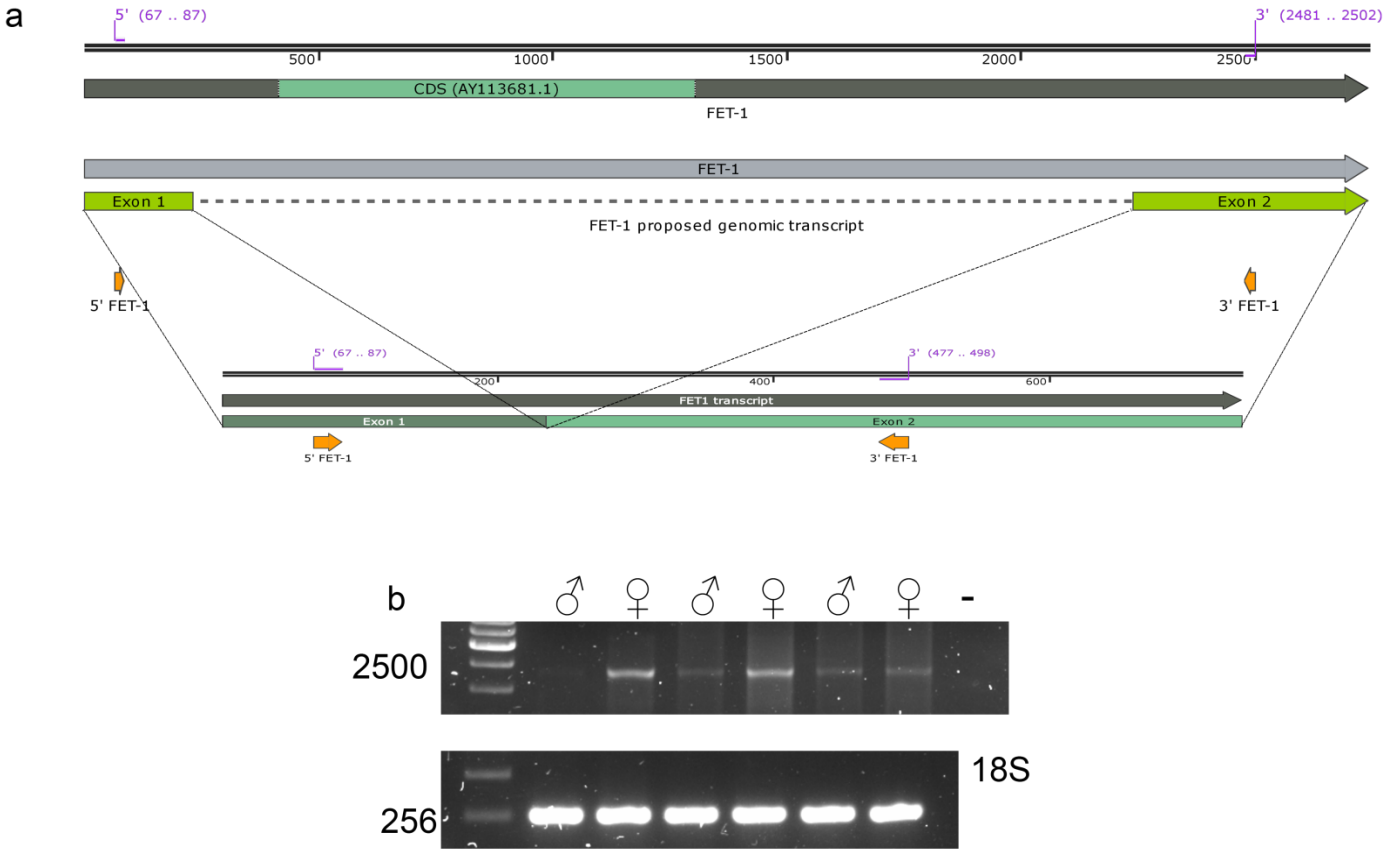


Figure 3. 3. Confirmation of absence of *FET-1* expression from chromosome 4 using RT-PCR. a) Primers (orange arrows) designed in the 5' and 3' UTRs of the *FET-1* cDNA corresponds to that of the annotated exons (green) in the chromosomal 4 sequence. Using RT-PCR, distinction between the expression of the *FET-1* cDNA (including the *FET-1* CDS) and the annotated *FET-1*(excluding the *FET-1* CDS) on chromosome 4 can be made. Removal of the intron between the two annotated exons on chromosome 4 would produce a 432 bp amplicon. Amplification from the *FET-1* cDNA (AY113681.1) would produce an amplicon of 2436 bp. This serves to confirm the absence of a full *FET-1* integration being expressed from chromosome 4. b) RT-PCR performed on E4.5 male and female embryo gonads with a single PCR amplicon at 2436 bp. This confirms the absence of *FET-1* expression from the chromosome 4, indicating that *FET-1* is likely expressed from elsewhere in the genome where the *FET-1* CDS sequence is present. ♂ - male, ♀ - female. -, negative no template control. 18S, positive internal control serving as 1) a RNA quality control and 2) a PCR positive control.

FET-1 integrations in Galgal5.0 genome

Once the absence of *FET-1* on chromosome 4 was confirmed and that no other transcript was transcribed from the same region, other possible *FET-1* integrations were investigated. In Ensembl94, a local alignment against the nucleotide database using BLAST was used to align *FET-1* against the entire genome (Galga5.0). Multiple integrations were observed with partial integrations on chromosomes 5, 10 and 12. Several partial CDS integrations were observed on the autosomes and full CDS integrations were observed on the W and Z chromosomes (Figure 3.4a) Further investigation showed a full *FET-1* CDS integration on the reverse strand directly downstream of HINTW (ENSGALG00000035489, scaffold AADN04000843.1), confirming the FISH localization of Reed and Sinclair, 2002 (Figure 3.4b).

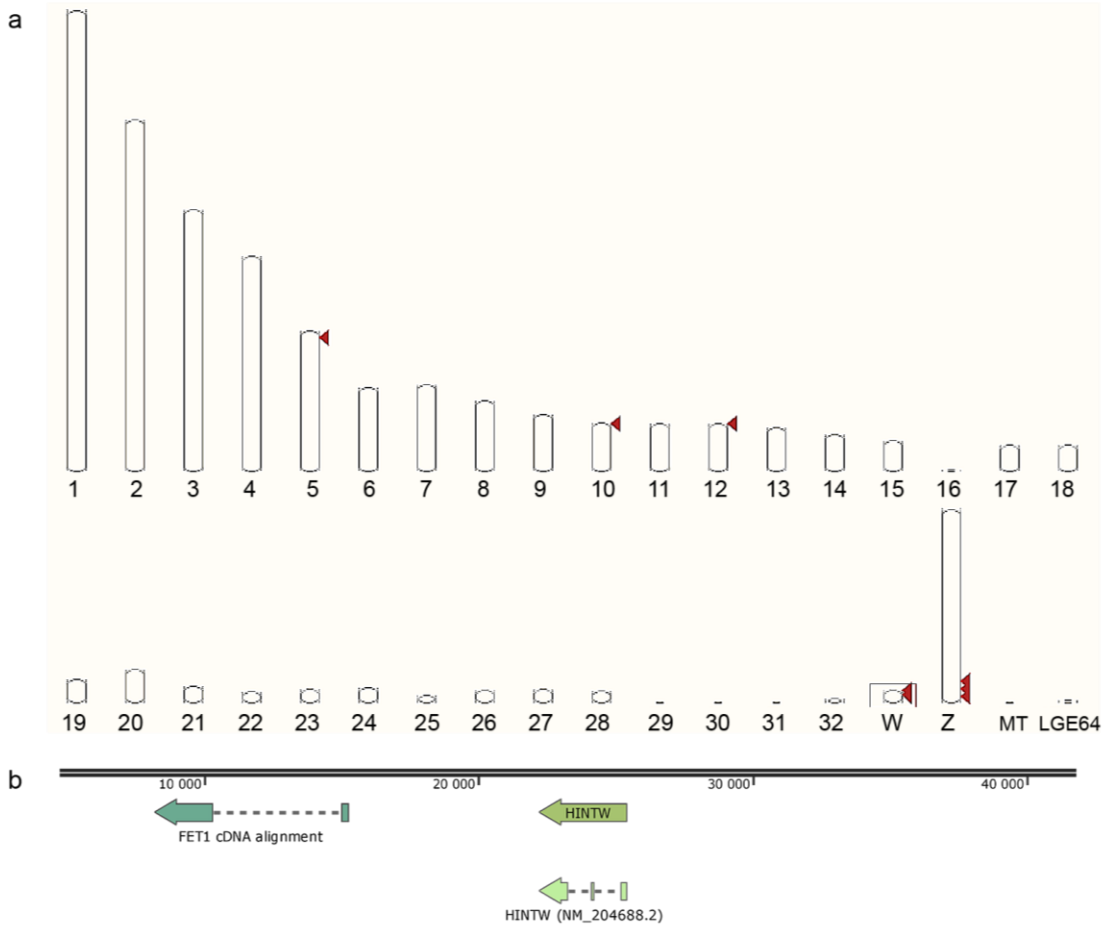


Figure 3. 4. *FET-1* integrations in *Gallus gallus* genome. a) *FET-1* (AY113681.1) local alignment search against the Galgal5.0 genome in Ensembl 90. Only partial CDS integrations were observed on the 5, 10 and 12 autosomes. Full *FET-1* CDS was observed on the Z and W sex chromosomes. b) Bioinformatics confirmation of Reed and Sinclair, 2002 localization of *FET-1* to Wp, showing an alignment of the full *FET-1* cDNA downstream of *HINTW*.

The full potential protein-coding *FET-1* is not located on chromosome 4, as annotated in NCBI. As shown by Reed and Sinclair, 2002, there is a full *FET-1* CDS integration on Wp, however it is downstream of HINTW and not within the HINTW repeat sequence. This observation was however most likely due to the low resolution of FISH. Full length *FET-1* CDS integrations, similar to that of the Reed and Sinclair, 2002 on the Wp, were observed on both the W and Z chromosomes. This raised the question why the probe used by Reed and Sinclair, 2002 did not hybridize to the other predicted integrations. It is also a possibility that the stringency washes during

the *in situ* protocol washed away loosely bound probes; given that some of the integrations identified are partial and the possibility of natural sequence variation, the binding efficiency of the probe would have been lower and therefore it would be easier to remove. The conformational state of the DNA could also have impeded the binding of the probes. Performing FISH in fibroblasts, which do not necessarily express *FET-1*, may have the *FET-1* integrations in regions where the conformational state of the DNA is inaccessible to the probe. However, since the chicken genome is still being annotated, especially the W chromosome, consideration must be taken of the constant revision and changes of genome annotations.

3.2. Multiple *FET-1* related transcripts are expressed in both males and females

A local alignment search using BLAST against the newly released Galgal5.0 genome assembly indicated the presence of multiple previously unidentified *FET-1*-like autosomal integrations. It is possible that as least some of these could be expressed. The expression data reported by Reed and Sinclair, 2002 seemed to indicate that only the W-integration is expressed, or that all of the autosomal integrations are expressed in a sex-specific manner. I decided to re-examine the expression profile of the *FET-1*-like transcripts

Reverse transcription PCR of FET-1

Initially standard RT-PCR was performed to confirm the sex-specific expression of *FET-1* reported by Reed and Sinclair, 2002. This was performed on extracted RNA from E4.5 male and female gonads of three chicken breeds. Primers were designed within the *FET-1* CDS (Figure3.5 a, Table 2.3). In contradiction to what was shown previously, *FET-1* expression was observed in both males and females, with expression significantly lower in males (Figure 3.5b).

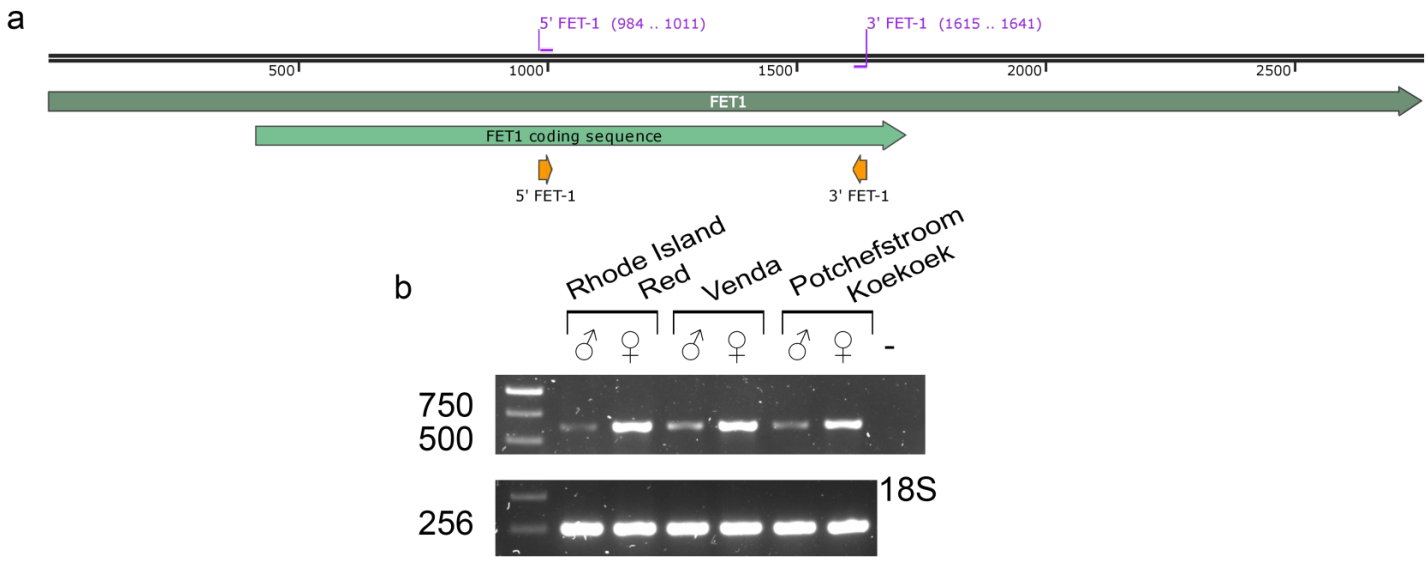


Figure 3. 5. Confirmation of Reed and Sinclair, 2002 ISH expression pattern using RT-PCR. a) Primers in the CDS of *FET-1* (orange arrows). b) RT-PCR on E4.5 male and female RNA extracted from three chicken breeds: Rhode Island Red, Venda and Potchefstroom Koekoek. Amplification observed in both sexes at 658 bp, consistently higher in females than in males for all breeds. ♂ - male, ♀ - female. -, no template control. 18S, integrity control for RNA used.

From the RT-PCR results, two questions arose: 1) what is the population of transcripts expressed? Given the multiple partial integrations identified from the genome database, there is a possibility that more than one of the *FET-1* integrations

is expressed; and 2) is the expression sex specific? If multiple related transcripts are expressed, it is clear from the RT-PCR that these transcripts are either down-regulated in the male or there is sex-specific expression, thus accounting for the higher expression in the female; a bigger population of transcripts being present.

Since RT-PCR is not indicative of the lengths or population of transcripts expressed, it was decided to perform a Rapid Amplification of cDNA Ends (RACE), a modified RT-PCR, which enables one to identify full length mRNA transcripts as well as variations of the mRNA of interest. Briefly, the 5' cap structure of the mRNA is removed and replaced with an RNA oligo of known sequence. The RNA oligo ligated to the 5' end of the mRNA would facilitate the addition of a known priming site on the 3' end of the cDNA. During reverse transcription using RACE specific Oligo-dT primers, the Oligo-dT primers will facilitate cDNA synthesis while at the same time extending the 5' end of the cDNA with a known priming site located within the Oligo-dT primers (Table 2.3). In combination with gene-specific primers (Table 2.3) full length *FET-1* mRNA and alternative splice forms or truncated versions can be identified (Figure 3.6a). 5' and 3' RACE was performed on RNA extracted from E4.5 male and female embryos. Both males and females showed a number of different transcript sizes (Figure 3.6b), suggesting different populations of transcripts and variations of *FET-1*. This coincides with the genome BLAST results showing multiple truncated *FET-1* integrations throughout the chicken genome (Figure 3.4a). The presence of different sized transcripts observed from RACE could be indicative of the various truncated *FET-1* integrations, possibly also indicating the sexually dimorphic activity of selected *FET-1* integrations.

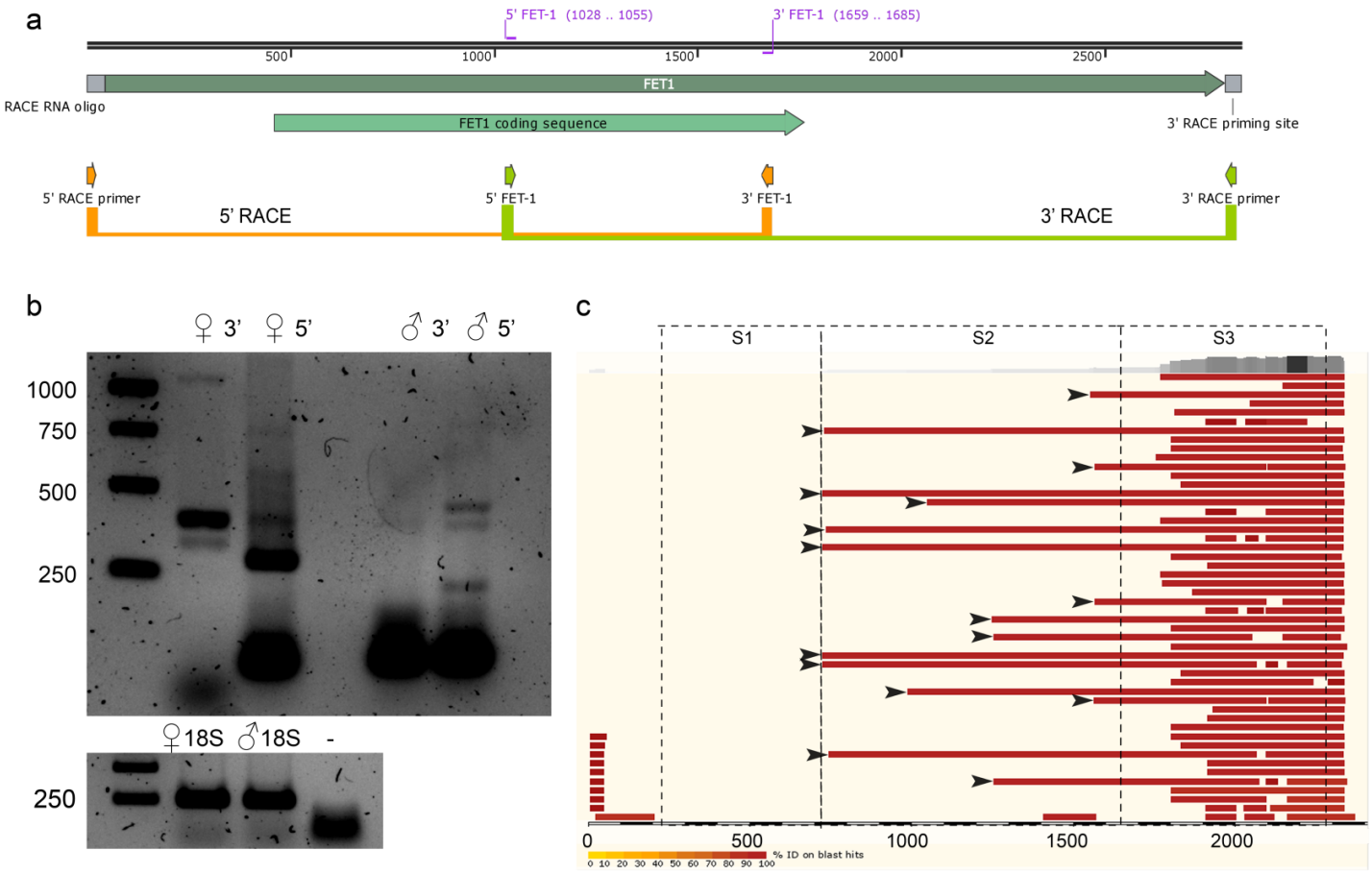


Figure 3. 6. Rapid Amplification of cDNA Ends (RACE) of *FET-1*. a) Diagram of RACE mechanism. The 5'cap of the mRNA is removed and replaced with a RNA oligo containing the 5' RACE primer binding sequence (orange arrow). Reverse transcription with the RACE Oligo dT adds a primer binding site for the 3' RACE primer (green arrow) The 5' RACE and 3' *FET-1* GSP primers; and 3' RACE and 5' *FET-1* primers are used for 5' and 3' RACE respectively, amplifying overlapping transcripts of the target mRNA. b) *FET-1* 3'and 5' RACE performed on E4.5 male and female RNA. ♂ - male, ♀ - female. 18S, RNA integrity control. -, no template control. c) lincRNA populations identified from *FET-1* a local alignment search against the Ensembl non-coding RNA database. S1 represent the *FET-1* (AY113681.1) transcript alone. Two lincRNA populations were identified (S2 and S3) showing homology to *FET-1*. S2 shows homology to the 728-1640 region of *FET-1* and S3 to 1641 to 2293 *FET-1* region. The grey bar above the sequence alignment is representative of the sequence coverage of *FET-1*. The black arrows indicate the lincRNA transcripts identified by RACE cloning and sequencing of individual RACE products.

Identification of FET-1-like lincRNA population

To identify the individual RACE products, the male and female RACE product was cleaned up using a PCR purification kit, ligated into the pGEM®-T Easy vector and cloned. This method was chosen instead of gel extraction of each RACE transcript since gel extraction yielded a suboptimal concentration for cloning purposes. Due to this it is possible that some of the less abundant RACE products did not ligate into a vector. Cloning was attempted multiple times, however some difficulties were experienced, likely due to the fact that *FET-1* has been identified as a retroviral transcript [75] and therefore not easily tolerated by the bacteria since it is recognised as a viral sequence and degraded or excluded from the cell. None the less, several male and female clones were isolated and sequenced. A local alignment search against the Ensembl94 non-coding RNA database identified the transcripts as long intergenic non-coding RNAs (lincRNAs) (Figure 3.6c). These transcripts, however, only aligned partially to *FET-1*, similarly to the truncated *FET-1* integrations observed during the local alignment search against Galgal5.0 in Ensembl94.

Most of my RACE products belonged to this population of lincRNAs aligning to *FET-1* (Figure 3.6, black arrows, Table 3.1). The *FET-1*-like lincRNAs could be grouped in two sets based on the length and homology to the full *FET-1* transcript (Figure 3.6c). Based on that, I annotated three different segments within the *FET-1* full length transcripts. Segment 1 (S1) would represent the 5'-most region of *FET-1* with no homology to any lincRNAs. S2 represents the middle region of *FET-1* that aligns to the longer lincRNAs, spanning across the *FET-1* CDS from 728-1640 bp.

S3 would represent the 3' region of *FET-1*, which aligns to all lincRNA from 1641-2293 bp.

Table 3. 1. *FET-1*-like lincRNA subpopulation identified from RACE**.**

Transcript	Genomic location	Size*	*FET-1* segment
ENSGALT00000075025.1	KQ759563.1:9800-9816	735	3
ENSGALT00000075472.1	KQ759563.1:7560-7576	739	3
ENSGALT00000061070.1	AADN04000843.1: 8,281-15,492	739	3
ENSGALT00000083673.1	AADN04008281.1: 1,315-2,409	1061	2
ENSGALT00000088774.1	AADN04013401.1: 1,423-2,528	1072	2
ENSGALT00000084186.1	AADN04003181.1: 6,597-7,702	1076	2
ENSGALT00000081113.2	12:55277-55293	1302	2
ENSGALT00000087372.1	Z:76639061-76639077	1344	2
ENSGALT00000075065.2	Z:76762329-76762345	1574	2
ENSGALT00000087498.1	AADN04001894.1: 11,089-12,702	1580	2
ENSGALT00000087546.1	AADN04001221.1: 12,577-14,204	1594	2
ENSGALT00000085805.1	AADN04000754.1: 648-2,290	1598	2
ENSGALT00000058686.2	W:2485168-2485184	1607	2
ENSGALT00000085969.1	Z:76762329-76762345	1609	2
ENSGALT00000084879.1	Z:74402582-74402598	1632	2

* To note is the fact that the transcript size in the table is slightly larger than the two segments sizes proposed in Figure 3.6c (S2 and S3). While the lincRNAs do align to *FET-1*, their alignment does not always start or end where at the first or last base pair of the lincRNA. The lincRNAs do, however, still largely show high similarity to *FET-1*.

In contradiction to what was shown previously, *FET-1* is clearly expressed in both males and females. Reed and Sinclair, 2002 used two whole mount ISH probes: a 151 bp probe and a 630 bp probe complementary to the 5' and 3' ends of the *FET-1* cDNA transcript respectively (Figure 3.7a). Additionally, the 5' probe sequence is also present in the 3' probe. While the *FET-1* clone isolated form a cDNA library of the chicken urogenital ridge[67] contains the 3' probe sequence, the region of 251 bp-630bp could not be found during an Ensembl94 and NCBI local alignment search against the genome assembly (Appendix 3, Figure 6A3.1). Additionally, no transcript within the Ensembl94 cDNA database or the lincRNA transcripts identified above (S2 and S3) contain this 3' probe sequence (Figure 3.6c).

A local alignment search against VecScreen, a database to screen for vector contamination, indicated no vector sequences within the *FET-1* CDS. The lack of proper description of the ISH probes used by Reed and Sinclair, 2002: "630 bp PCR fragment from the 3' end of the *FET-1* cDNA clone", makes their probe description open to the most obvious interpretation of its location; that the probe is located from the most 3' end, stretching 630 bp upstream into the *FET-1* cDNA. It still stands that the 251-630bp region of the 3' probe, as interpreted from Reed and Sinclair, 2002, does not appear in any of the nucleotide databases with the exception of aligning to the original *FET-1* cDNA sequence submitted to NCBI by Reed and Sinclair, 2002. The absence of this sequence in the genome, RNA database and the absence of vector contamination puts in question the use of the probe.

3.3. *FET-1*-related transcripts are expressed in gonads and prectoral muscle in sex-specific patterns

Because none of the previously published investigations reported on the existence of the *FET-1*-like lincRNAs or any male gonadal expression of *FET-1,* I decided to re-examine the spatiotemporal expression pattern of *FET-1* and related transcripts using *in situ* hybridisation. In an attempt to distinguish between the expression patterns of the three sets of transcripts, I used three probe sets (Figure 3.7): S1, S2 and S3. I reasoned that probe S1 would bind to only the full-length *FET-1* (transcribed from W chromosome). Probe S2 would bind both the full-length transcript and the longer S2 lincRNAs. Probe S3 would bind the full length transcript, transcripts within the S2 region as well as the shorter lincRNA transcripts (the S3 region). Therefore, any novel expression domains demonstrated by probe S2 but not probe S1 would be an indication of the expression of S2 lincRNAs. The same can be said for probe S3; any novel expression domains demonstrated by probe S3 is indicative of S3 lincRNA expression. Figure 3.7 shows the outline of the riboprobe preparation process.

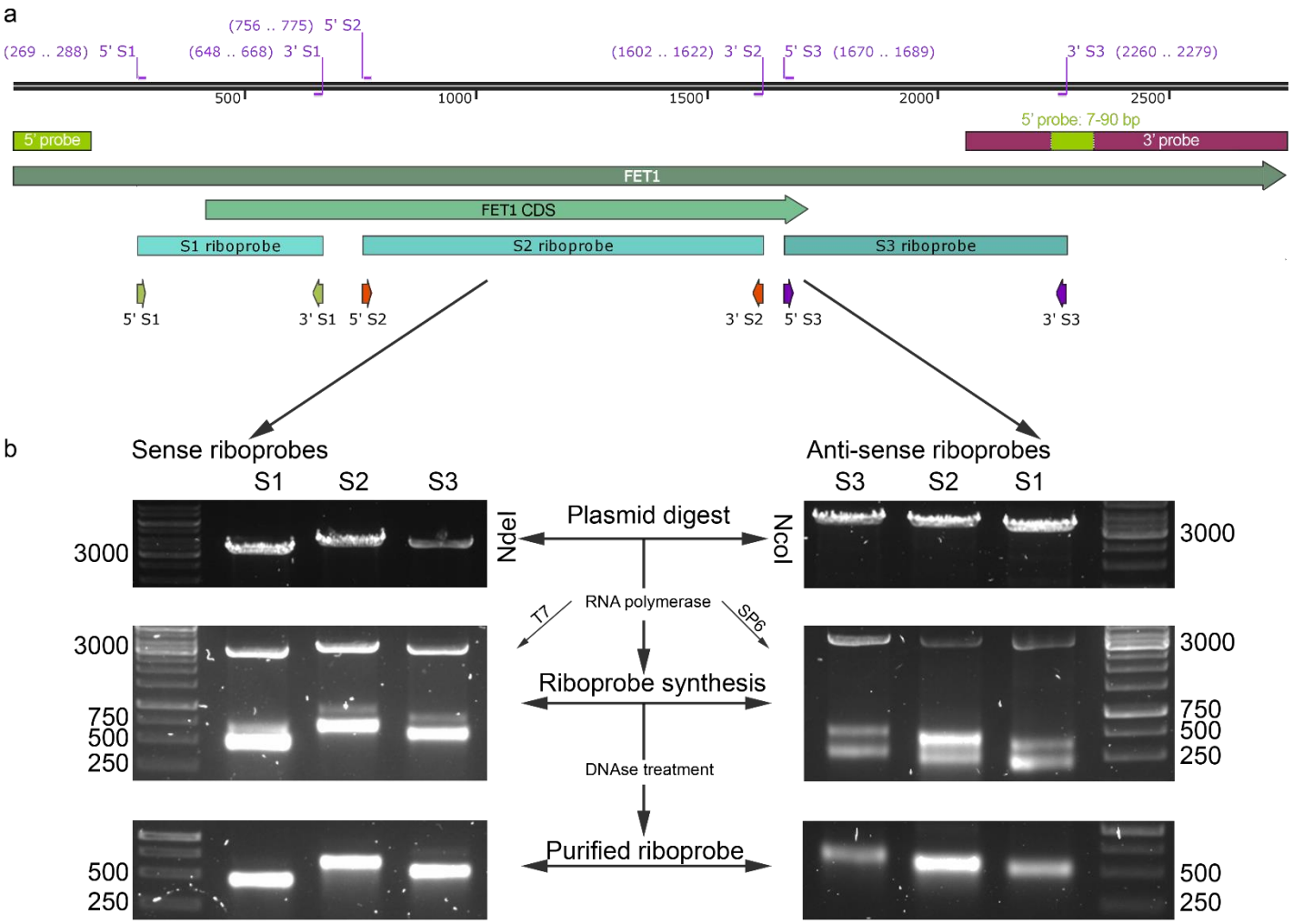


Figure 3. 7. Riboprobe design and synthesis. a) 5' probe and 3' probe: riboproes used by Reed and Sinclair. The 5' probe is repeated within the 3' probe. Riboprobe S1, S2 and S3: riboprobes and accompanying primers against the *FET-1* and the two lincRNA populations. b) Riboprobe synthesis. Top panels: plasmid linearization of pGEM®-T Easy plasmids containing PCR products of S1, S2 and S3; *Nde*I and *Nco*I was used for sense and anti-sense plasmids of each clone. Middle panels: riboprobe synthesis from linearized plasmids using T7 and SP6 RNA polymerase for the sense and anti-sense plasmids respectively. Bottom panel: purified and concentrated riboporbes after DNAse treatment to remove plasmid DNA from synthesized riboprobes.

It should be noted that the probes represented on the gel during synthesis are slightly smaller than expected, for which there are multiple possible contributing factors. In Figure 3.7, both the middle and lower panels show "ghost" bands (more apparent in the middle panel). This is most likely due to the probe secondary

structures not adequately resolved before being viewed on the gel. The probe was visualised on a standard 1% agarose gel, while the loading dye contained formamide to facilitate removal of secondary structures. It could be possible that the probes were not heated for long enough to ensure complete linearization of all the RNA probes. Since folded RNA migrates faster, the probes appears smaller than expected.

To be noted is the size of the S2 anti-sense probe, which is smaller than expected. This is due to the presence of an NcoI restriction site in the 5' end of the probe sequence. This shortened the probe by approximately 300 bp, but the probe was determined to still be adequate to bind S2 lincRNAs, since it was still specific to the S2 region of the FET-1 transcript.

Regardless of the smaller than expected probe size visualised during gel electrophoresis, sequencing of the recombinant vector containing the probe prior to probe synthesis confirmed synthesis of the correct probes sequences.

RNA in situ hybridization

ISH was performed on E3.5, E4.5 and E6.5 male and female embryos (Figure 3.8) using the sense and anti-sense riboprobes. No staining was observed for any of the probes in gonads of the male embryos (Figures 3.9, 3.10, 3.11 and Table 3.2), compared to females that showed gonadal staining in both the right and left gonads; more clearly for S1 and S3 at stages E4.5 and E6.5. S2 female gonadal staining was minimal for stage E4.5, while E6.5 showed no staining. However, there seems to be

no difference in staining between the left and right gonads of the females, as shown in by the RT-PCR (Figure 3.2b) as well as Reed and Sinclair, 2002. In the E4.5 female embryos stained with the S1 anti-sense probe, staining is observed in the outer edges of the kidneys, not seen for any other probe or embryo stage.

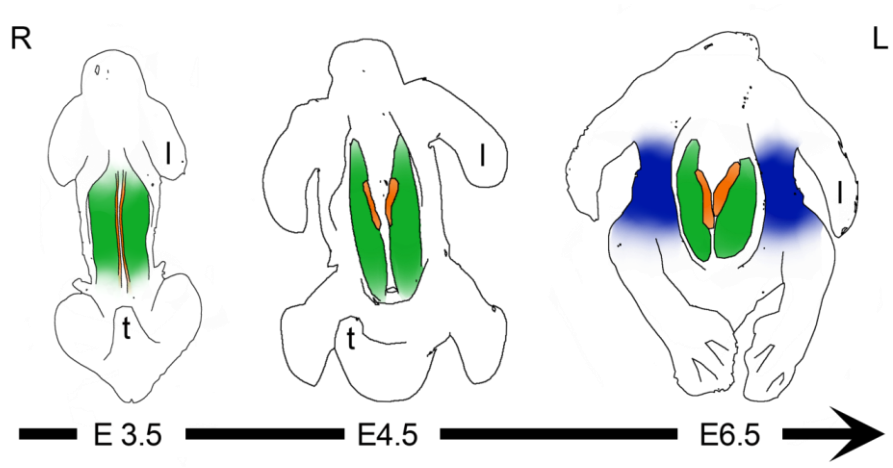
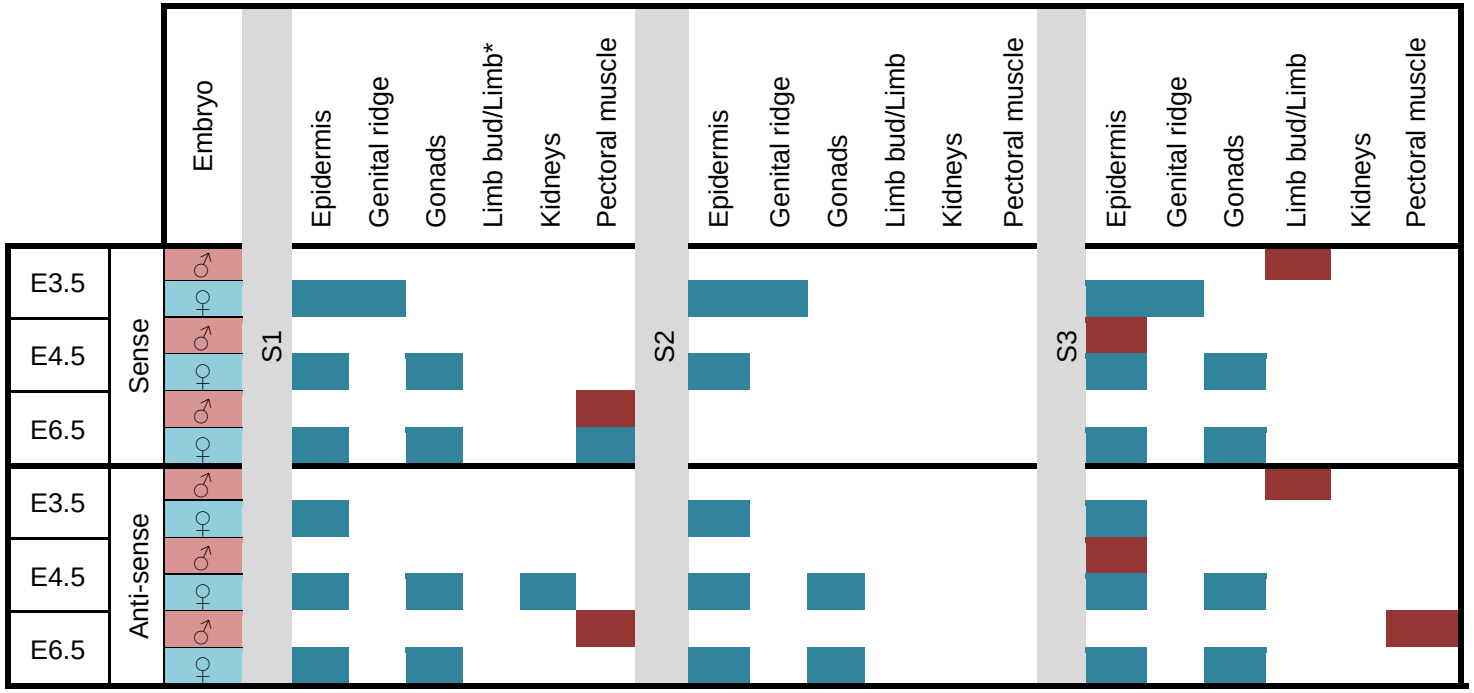


Figure 3. 8. Schematic of dissected chicken embryos at stages E3.4, E4.5 and E6.5. The kidneys are represented in green. The orange structures represents the genital ridge and gonads at E3.5 and E4.5/E6.4 respectively. The blue regions at E6.5 represent the region of developing pectoral muscle. t, tail. l, limb. R, right and L, left.

For both males and females, epidermal staining increases from probe S1 to S3, although this staining is higher in the female epidermis. Pectoral muscle is stained dimorphically in the S3 male embryos stained with the S3 anti-sense probe. Staining was observed for both the sense and anti-sense probes, suggesting transcription of lincRNA transcripts in both the forward and reverse orientations. It is also possible that the anti-sense staining is non-specific off target staining; however, the partial *FET-1* integrations identified in the genome do occur in both the sense and anti-sense directions.

Table 3. 2. Summary of ISH S1, S2 and S3 sense and anti-sense staining in male and female embryos. Blue and red represent female and male expression respectively.



*Limb buds are preset in E3.5 embryos, by E4.5 limbs are present

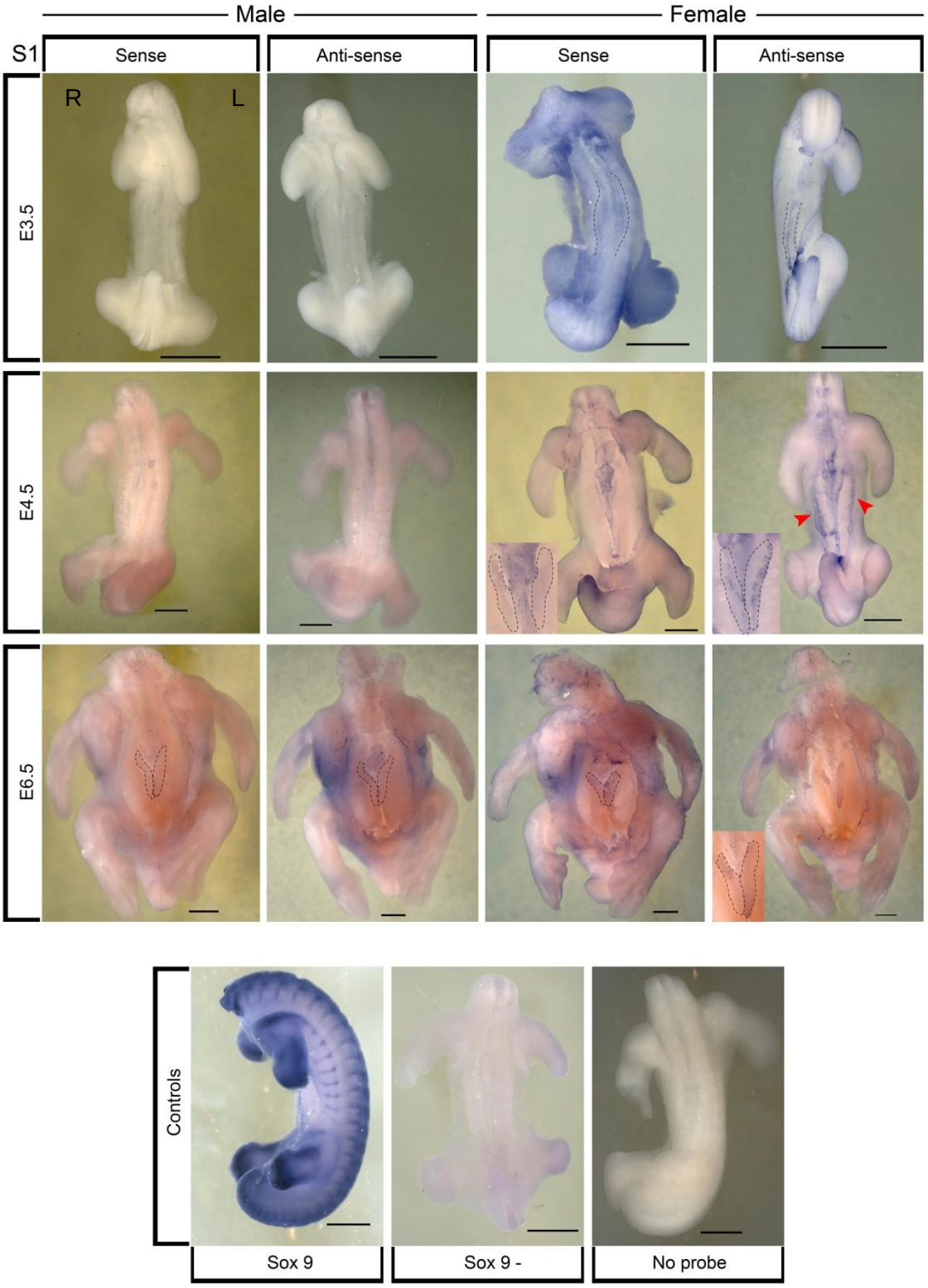


Figure 3. 9. S1 *in situ* hybridization. Female specific gonadal expression for E3.5, E4.5 and E6.5. Additionally, E4.5 and E6.5 females show epidermal expression. E6.5 males show staining specific in the pectoral muscle. The E4.5 anti-sense female shows staining in the outer edges of the kidneys (red arrows). Bottom panel: Sox9, positive control. Negative controls, Sox9- and no probe.

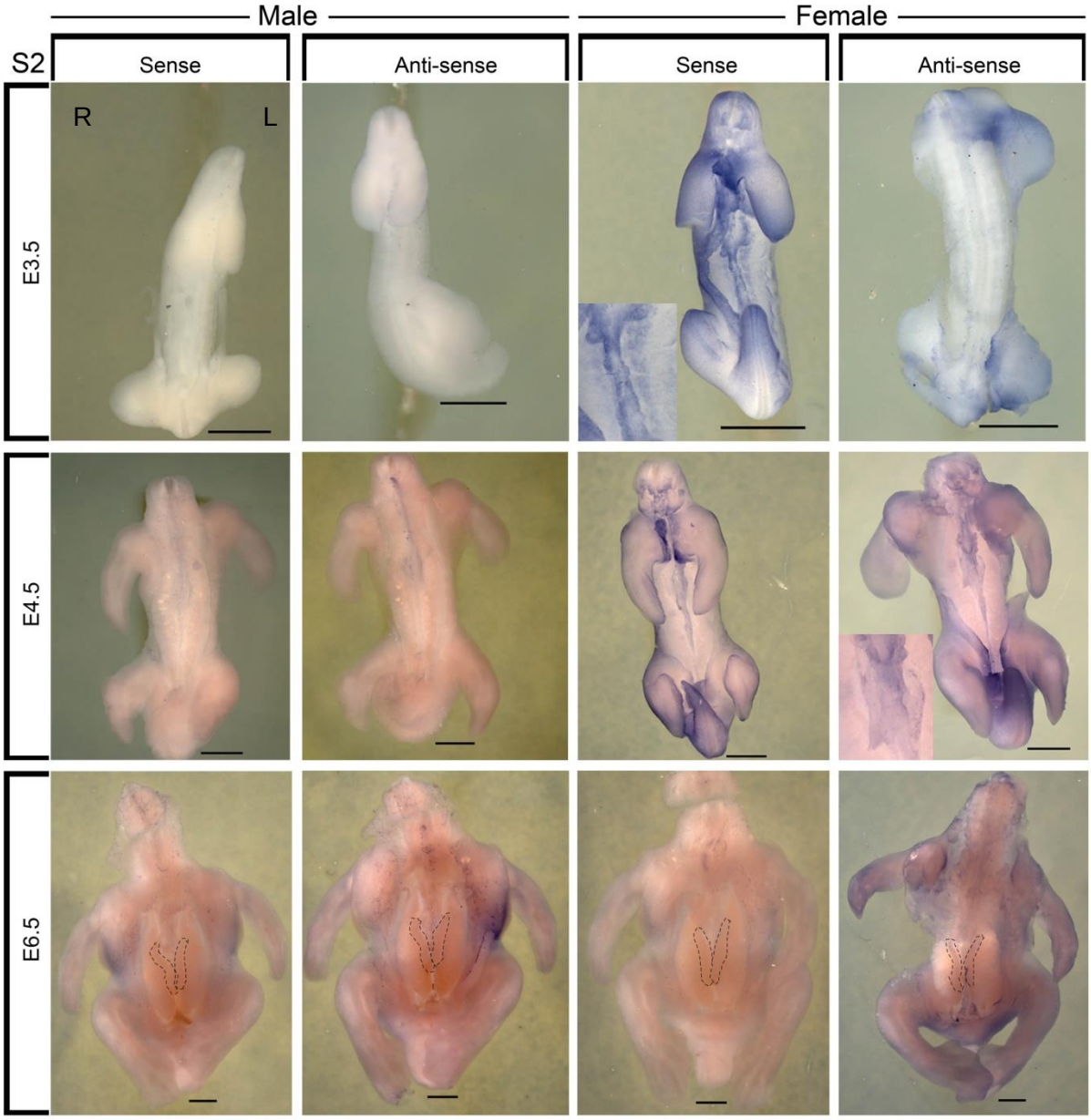


Figure 3. 10. S2 *in situ* hybridization. Female specific gonadal expression for E3.5, E4.5 and E6.5. Additionally, E4.5 and E6.5 females show epidermal expression.

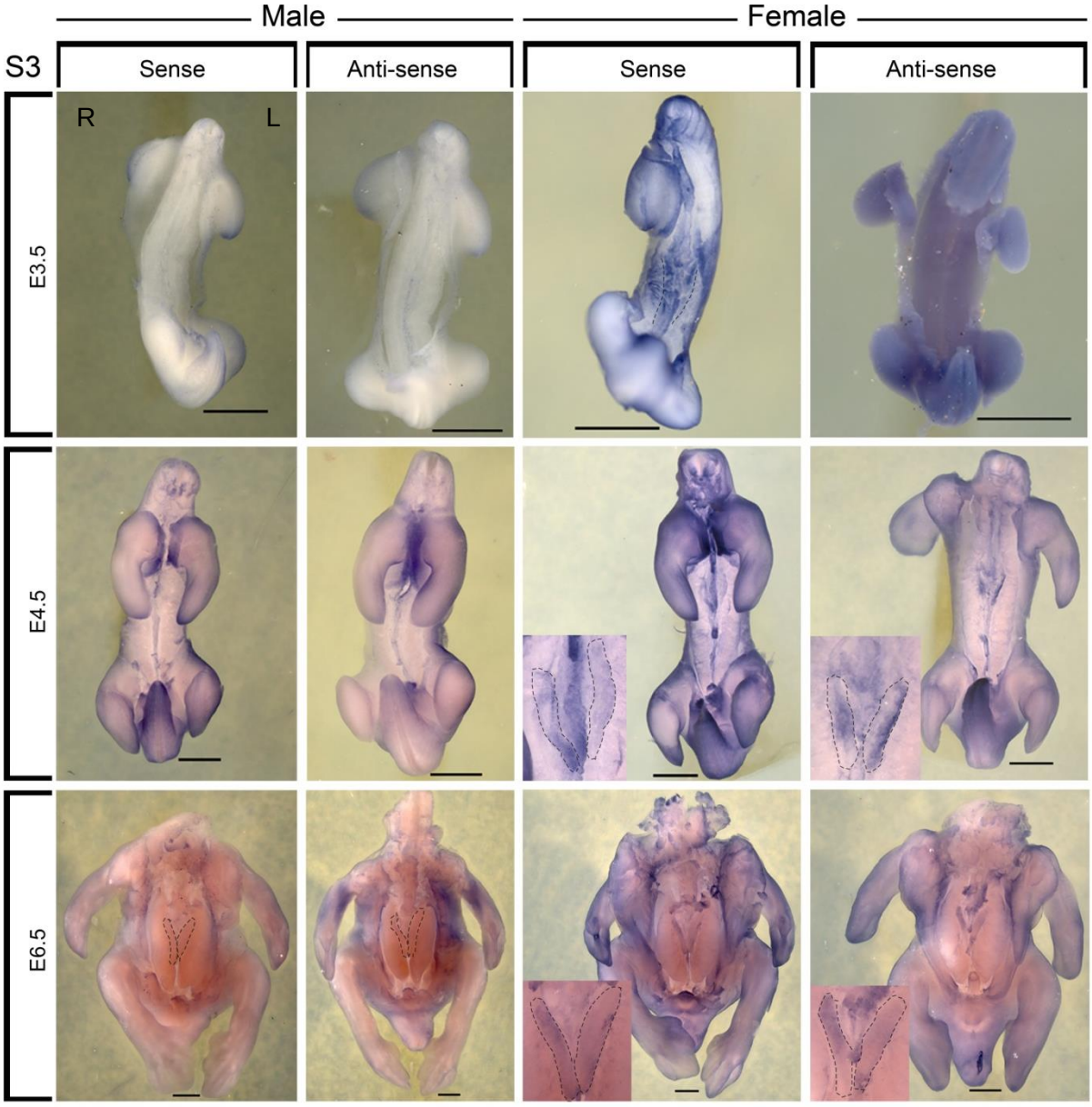


Figure 3. 11. S3 *in situ* hybridization. Female specific gonadal expression for E3.5, E4.5 and E6.5. E3.5 males show expression in the limb buds while E4.5 and E6.5 males show epidermal expression in addition to the females. E6.5 males show expression in the pectoral muscle

It would be expected that staining from the S1 to S3 probes would be additive since the *FET-1*-like lincRNA populations overlap *FET-1* and one another (Figure 3.6c). This however was not observed. This is likely due to insufficient staining, probe entry into the tissue due to insufficient permeabilization and/or potentially the removal of the probe due to excessive stringency washing, especially the male gonads where expression is very low. With this in mind, apparent dimorphic sex-specific expression patterns were observed in the 1) pectoral muscles in males, suggesting a role for *FET-1* related lincRNA during muscle development and 2) gonads, even though no staining was observed in the male gonads, the RT-PCR suggests a female-specific function for *FET-1* in the gonads. It should be noted that the RT-PCR amplified the same region the S2 probe was designed against; however different results were observed for the RT-PCR versus the ISH. This disparity between the two methods is most likely due to technical qualities of the methods and their abilities to detect the target transcript.

3.4. Full-length *FET-1* may be translated

Since *FET-1* has been shown to contain an open reading frame that encodes a predicted protein (acquisition number: AAM52407.1) it was decided to investigate the possibility of a FET-1 protein being transcripted in addition to the *FET-1*-like lincRNAs, as well as the possible function of this protein.

FET-1 *predicted protein*

The *FET-1* cDNA contains an open reading frame encoding a predicted 434 amino acid protein with 49.139 kDa molecular weight. It was shown to contain a conserved domain of the subfamily ENVV1-like HR1-HR2 that forms part of the Ebola HIV-1-like HR1-HR2 superfamily. This domain encodes for an envelope protein, containing a trans-membrane subunit found within various endogenous retroviruses [66] (Figure 3.11a). NCBI SmartBlast was used to identify possible orthologues and provide insight into the potential function of the FET-1 predicted protein. SmartBlast, a local alignment tool, compares a protein query against a landmark database consisting of 27 genomes across a wide taxonomic range. This tool aims to provide the five best matches to the query protein along with a conserved domain search through the Conserved Domain Database (NCBI). FET-1 was shown to have 35% and 32% identity to syncytin-1 precursors of human and mouse respectively (Figure 3.12b). It should be noted that the landmark database is not comprehensive in terms of vertebrate species represented; only containing human, house mouse and zebrafish genomes. The other 24 represented genomes are single cell eukaryotes, parasites, archaea and bacteria.

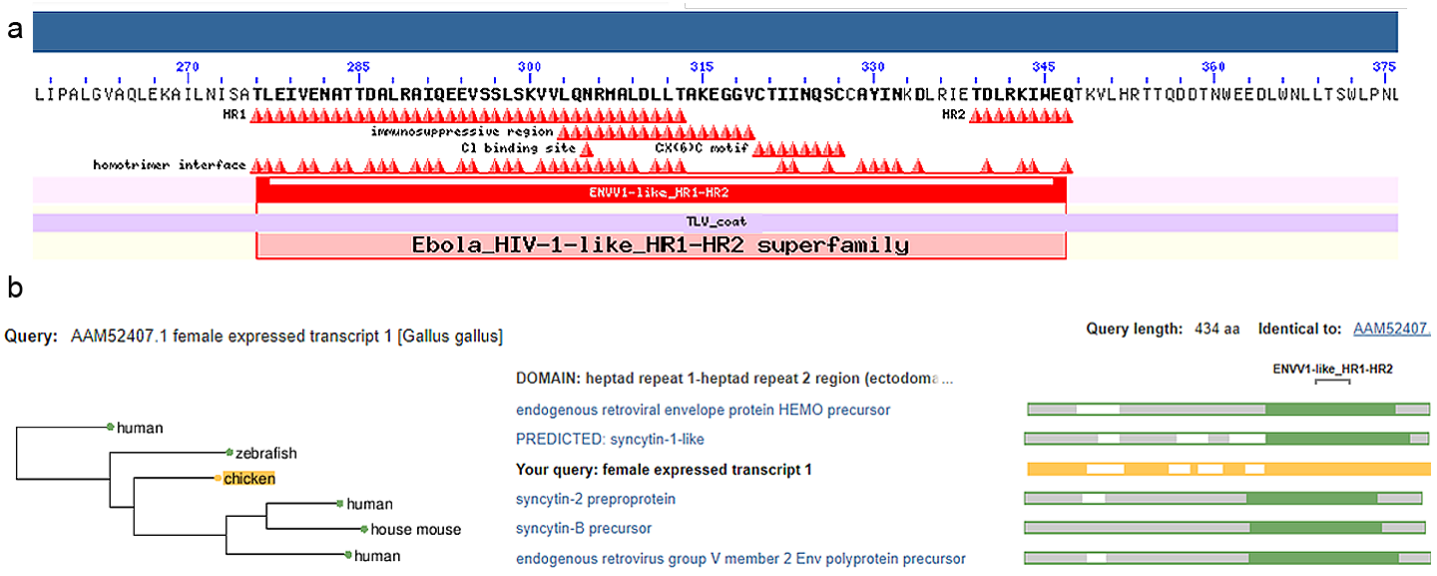


Figure 3. 12. FET-1 predicted protein, AMM52407.1. a) NCBI conserved domain database search of the *FET-1* predicted protein with hits against a retroviral envelope polyprotein as well as the ectodomain of the transmembrane of the human retrovirus ENNV1. b) SmartBlast of the *FET-1* predicted protein with similarity to the human, mouse and zebrafish syncytin precursors.

Immunohistochemistry of FET-1

To determine if the FET-1 predicted protein is expressed, a custom rabbit polyclonal anti-FET-1 antibody was synthesised by Genscript (Figure 3.13a). Indirect IHC was performed in parallel whole mount and sectioned E4.5 male and female embryo (Figure 3.13b and c). In addition to the anti-FET-1 antibody, a pre-immune serum control was included as a negative control for immunisation of the host. The pre-immune serum is extracted from the host species prior to immunisation. Since there is a risk that the host animal previously developed antibodies against the protein of interest or a similar protein, the pre-immune serum is extracted before immunization and used in place of the primary antibody. Staining using the pre-immune control suggests the presence of native antibodies in the host capable of binding to the target protein/ a similar protein to FET-1, making it difficult from distinguishing

between true signal due to the primary antibody and non-specific signal from the native host antibodies.

The tissue sections showed ubiquitous staining in the gonads and limbs in both the experimental and pre-immune samples, suggesting 1) that the anti-FET-1 antibody was not specific enough and 2) cross-reactivity of host native antibodies to chicken proteins. Since the antibody was produced in rabbit, BLASTp[76] was used to compare the predicted FET-1 protein to the rabbit protein database in both Ensembl and NCBI. Even though these databases contain the same information, differences in BALST parameters and discrepancies in the underlying genome/ proteome assembly due to different annotation methods produces different results when preforming local alignments against the same genome across databases.

Similar to the SmartBlast results, the FET-1 predicted protein showed 27% identity to the rabbit syncytin-Ory1 precursor (Appendix 3, Figure 6A3.1a) and 31% similarity to the uncharacterised protein LOC103345346 (isoforms X1 and X2; Appendix 3, Figure 6A3.1b). This suggest the presence of proteins similar to FET-1 in the rabbit that could be responsible for the initiation of antibody production prior to immunization with the FET-1 recombinant protein.

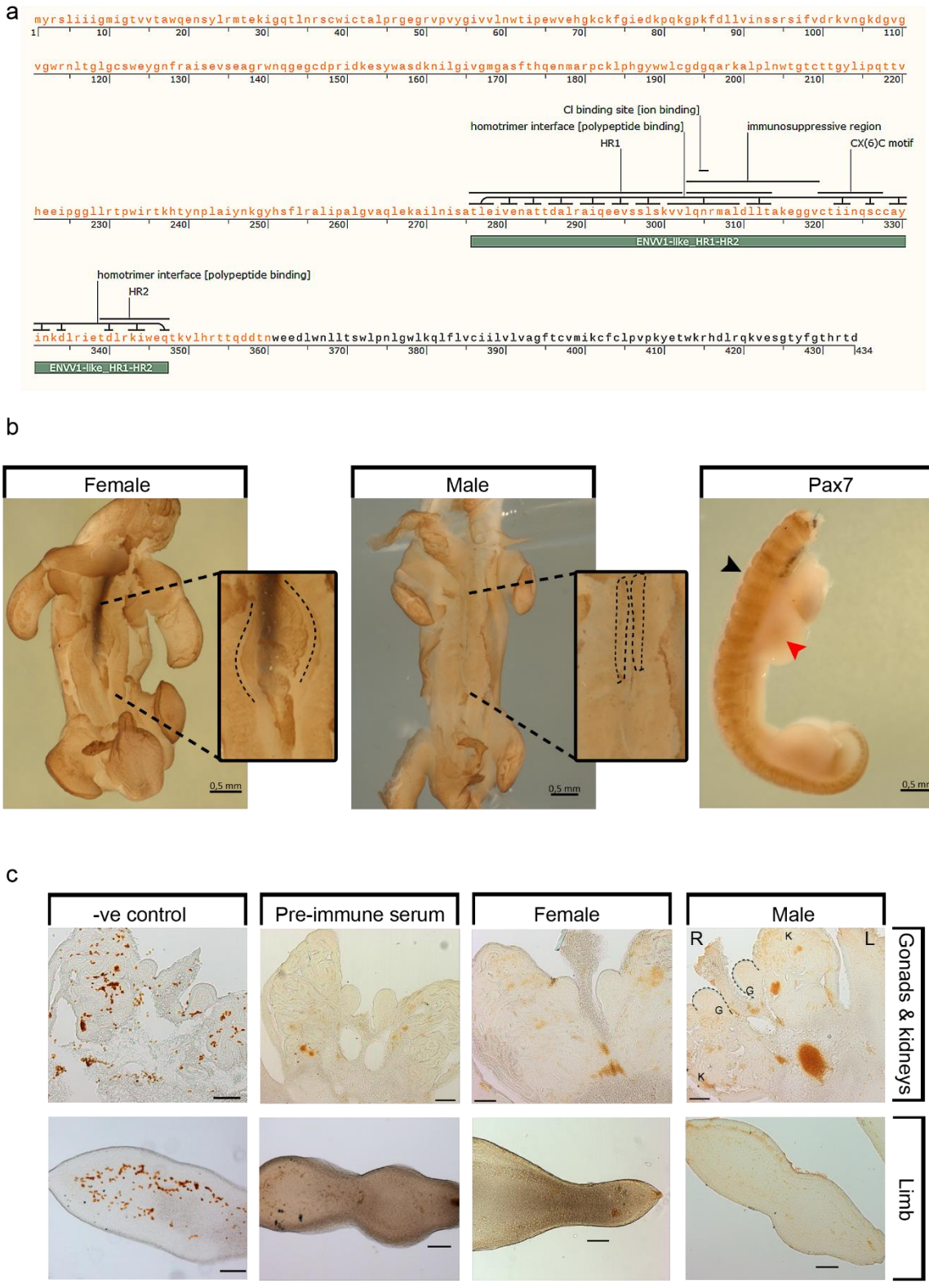


Figure 3. 13. Immunohistochemistry of the FET-1 predicted protein. a) FET-1 predicted protein sequence. The orange highlighted sequence, spanning across the conserved retroviral domain was used to produce a recombinant FET-1 protein for antibody production. b and c) IHC performed in parallel on whole mount embryos and tissue sections of the gonads/kidneys and limb. b) IHC on whole mount male and female embryos. Ubiquitous staining is present throughout both the female and male embryos. The enlarged images shown is the gonads outlined in black. Pax7 is the positive control, showing staining the Pax7 associated tissues; the premuscle mass in the wing/hind limb (red arrow) and the somite-derived muscle precursors (black arrow). c) IHC of male and female tissue sections. Kidney, gonad and limb sections showing ubiquitous staining in the male, female and pre-immune serum samples. Trapping of secondary antibody shows sporadic staining in the negative control, however the majority of the sections are not stained. Similar trapping is seen the other sections.

Dot blot and Western blot of FET-1

Given the ubiquitous staining of the IHC sections, the dot blot was used to investigate whether the same ubiquitous staining occurs across a wider range of tissues; serving as a less labour intensive method to confirm the expression observed during IHC. The dot blot was performed on a crude protein extract from the gonads, kidneys, limb and pectoral muscle from both male and female E4.5 embryos (Figure 3.14a). As with the IHC, staining was observed across all the tissues targeted, confirming ubiquitous staining seen during IHC.

The dot blot additionally allowed for several control to determine if the staining was background/non-specific. The “no primary antibody” negative control ensured that the staining is due to the binding of the FET-1 antibody and not non-specific background staining. Whether the staining was specific to the FET-1 predicted protein was not clear. The dot blot further served to test the function of the anti-FET-1 antibody (Genscript) using a positive control; the recombinant FET-1 protein used to produce the anti-FET-1 antibody. This demonstrated high affinity of the FET-1 antibody for the recombinant FET-1 protein, thus confirming the capability of the antibody of binding the predicted protein, if expressed. The use of a pre-immune serum control was used as a control for the animal host used for antibody production. This refers to the use of immune serum extracted from the host (New Zealand rabbit in the case of the FET-1 antibody) before immunization with the FET-1 recombinant protein; this control indicates if any component of the host immune system reacts with the target protein. Both the IHC and dot blot showed staining for the pre-immune serum control (Figure 3.13b and 3.14a), suggesting a component of

the rabbit pre-immune serum reacted with the protein/s of the tissue crude protein extract; the presence of antibodies in the host able to at least bind the FET-1 predicted protein (if expressed). This would indicate that the host already possessed antibodies capable of reacting to the FET-1 recombinant protein and suggests that the FET-1 antibody was not produced purely in response to the immunization of the FET-1 recombinant protein. As indicated previously in Figures 6A3.1a and b, there are a few proteins in the rabbit that show similarity to FET-1; these could include proteins that were recently acquired via retroviral integrations that still produce antigenic proteins similar to FET-1.

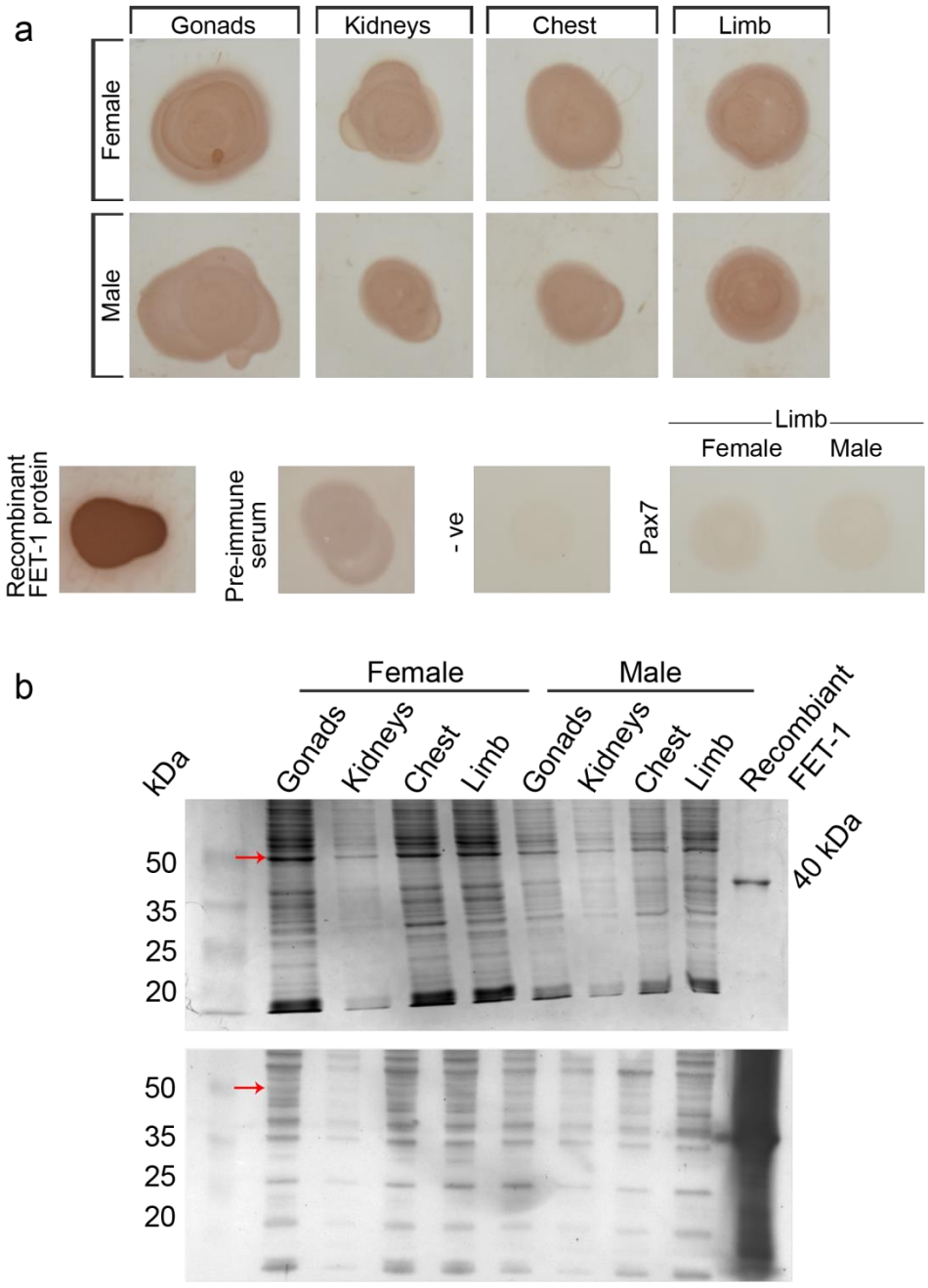


Figure 3. 14. Dot blot and Western blot of crude protein extract. a) Dot blot of gonad, kidney, prectoral muscle and limb crude protein extract, showing ubiquitous staining in all crude protein extracts. The recombinant FET-1 protein sample is a positive control for antibody function. –ve, represents the primary antibody (FET-1 antibody) negative control. Pax7 is the positive control used on crude protein extract from the limb. b) Coomassie blue stained SDS-PAGE and Western Blot of gonad, kidney, prectoral muscle and limb crude extract. Top panel: Coomassie blue stained SDS-PAGE gel giving an indication of the total protein content of the selected tissues. Bottom panel: Western blot of gonad, kidney, prectoral muscle and limb crude protein extracts showing binding of the *FET-1* antibody to a subset of the protein banding showed in the SDS PAGE. Both panels: The red arrows indicate the position where the *FET-1* predicted protein (49.139 kDa) is expected to be located. The last lane contains the *FET-1* recombinant protein at 40 kDa as a positive control.

Since both the IHC and dot blot showed ubiquitous staining across tissues, a Western blot was used to get an idea of the reactivity of the FET-1 antibody to the tissue crude protein extract and to determine if the predicted size of the FET-1 protein was stained. A Coomassie blue stained SDS-PAGE gel was made to visualize the natural banding pattern of the protein in the crude extract (Figure 3.14 b, top panel). Concurrently, a Western blot was performed using the anti-*FET-1* antibody to visualize the staining pattern (Figure 3.14b, bottom panel). As expected, the Commassie blue SDS PAGE showed a wide range of protein banding patterns, similar across all the tissues targeted.

The Western blot in turn showed staining across a subset of the protein banding pattern, including the size of the FET-1 predicted protein at 49.139 kDa (Figure 3.14b, red arrows). This could be explained by several factors. Firstly, that the FET-1 antibody is non-specific to the FET-1 predicted protein. Since FET-1 has similarity to a conserved domain (included in the region used to produce the antibody, Figure 3.13a) found in the trans-membrane subunit of various endogenous retroviruses, it is possible that this antibody reacted to other retroviral integrations expressed during embryonic development. Secondly, if the pre-immune serum staining is accurate, a percentage of the FET-1 antibody isolate consists of pre-immunization antibodies produced by the rabbit host against similar antigens to FET-1, able to bind to non-specific antigens in the crude protein extracts.

CHAPTER 4 - DISCUSSION AND CONCLUSION

Sex determination in avians is still poorly understood. While *DMRT1* plays a well-documented role to promote male gonadal development[28,31,77], the existence of a W-linked factor cannot be excluded. This is supported by ZZW triploid chickens that develop a left ovo-testis which later reverts back to a functional testis. This suggests the presence of a W-linked determinant that 1) influences sex determination and the downstream development of primary sex structures[61], and 2) is most likely recessive to that of the Z-linked factor, *DMRT1*. For a candidate gene to be considered a W-linked determinant it has to fulfil several requirements: 1) the gene has to be located on the W chromosome, with no Z gametologue, 2) gene expression has to occur before the onset of morphological differentiation and 3) the expression must be localised to the mesenphrenos/ gonads, likely also reflecting the development of the gonads; higher in the left ovary compared to the right ovary that regresses shortly after morphological differentiation.

This project aimed to confirm previous findings and to further evaluate *FET-1* as a possible female determinant. Reed and Sinclair, 2002 reported that *FET-1* is located on the short arm of the W chromosome, making it a promising candidate for the W-linked determinant. However, it is only in the last few years that the chicken genome has been annotated adequately, with the W chromosome still largely unannotated. As such, from the results above (Figure 3.4) multiple *FET-1* integrations were shown to be present on both the W and Z chromosomes when utilizing nucleotide basic local alignment search against the chicken genome assembly, Galgal5.0. Reed and Sinclair, 2002 only annotated one of the *FET-1* integrations to the W chromosome,

likely due to the limits of experimental techniques including the low resolution of DNA *in situ* hybridization. Additionally, access to transcriptome data such as the non-coding RNA database enabled identification of *FET-1*-like transcripts likely expressed from the autosomal genomic integrations.

The initial reports on *FET-1* demonstrated female-specific expression between E3.5 and E6.5[67]. The identification of multiple *FET-1* integrations and *FET-1*-like lincRNAs on both the Z and W chromosomes was well as their expression in both males and females is in contradiction to what as shown previously. Recently, Hirst et al, 2017 also re-examined *FET-1* expression by *in* situ hybridisation, demonstrating that it was also expressed in male gonads, albeit at much lower levels than in females. It would be expected for the W-linked determinant to be integrated into the canonical sex-determining pathways; i.e. the FOXL2/ 17βHSD/aromatase and Wnt4/Rspon1/ β-catenin pathways. Functional work performed by Hirst et al, 2017 showed *FET-1* expression to be unaffected after aromatase knockdown. This result is not unexpected, given that aromatase is expressed only from E6.5[79] compared to *FET-1* expression that is observed from E3.5. Co-localisation of *FET-1* and aromatase has not been demonstrated, however from separate studies it is shown that they co-localize in their temporal expression; aromatase from E6.5 into adulthood[65] and *FET-1* from E3.5[67] to at least E9.5[78]. Spatial localization places *FET-1* mRNA expression in the cortex (higher in the left than the right) and aromatase expression in the medulla, equal expression in both ovaries. The absence of a change in *FET-1* expression subsequent to aromatase knockdown suggests that *FET-1* is not a downstream component of the FOXL2/ 17βHSD/aromatase pathway. While *FET-1* is expressed well before aromatase, the dimorphic *FET-1* expression of the ovaries

does not reflect that of aromatase, which would be expected if *FET-1* acts upstream to trigger/regulate factors upstream of aromatase or even aromatase itself. The relationship between *FET-1* and the Wnt4/Rspon1/ β-catenin has not been investigated.

FET-1 only, with certainty, complies with the third requirement of its expression being localized to the developing gonads. The RT-PCR expression profile in the results above show the same results reported by Hirst et al, 2017; that *FET-1* expression is present in both males and females. None of the previous studies have investigated *FET-1* expression outside of the urogenital system (the gonads and kidneys). The S1 and S3 ISH results (Figure 3.9 and 3.11) have demonstrated *FET-1* expression not only in the gonads but in the pectoral muscle. It is shown that 1) higher *FET-1* expression is observed in female gonads compared to male gonads when using RT-PCR. There is, however, disparity between the RT-PCR and ISH results. RT-PCR, being the more sensitive technique; able to amplify the smallest amount of RNA present is more reliable toward the observation of higher expression in females. The absence of staining for S2 (the probe corresponds to the section amplified during RT-PCR) is concerning but likely due to intrinsic technical differences and the associated difficulties of the two techniques, such as reduced entry of the probes into the tissue due to improper permealisation and under staining. The staining patterns reported by Hirst et al, 2017 demonstrated dimporhic expression similar to the RT-PCR results, reported in Chapter 3, when using a probe targeting the same general region of the RT-PCR primers and the S2 probe. This further supports subpar execution of the technical methodology as the reason for no/reduced staining in the S2 embryos. 2) The transcripts represented by S1 (*FET-1* alone) and S3 (second

population of lincRNAs and *FET-1*) is expressed in male pectoral muscle, with the reverse complement expressed in female pectoral muscle. This suggests a possible sex-specific function in sexually dimorphic tissue, a quality of CASI genes.

Gynandromorphs have demonstrated that several secondary sex-dimorphic structures are developed not only under the influence of the sex steroid hormones produced from the developing gonads[63], but also the genetic identity of the somatic cells. These sexually dimorphic structures include the pectoral muscle, comb, leg spurs and wattle; all larger in males compared to females and are dependent on the cell autonomous sex identity of the individual somatic cells from which they develop[63]. Transgenic male chickens ectopically expressing aromatase were shown to develop female like wattle and comb growth, however retained male weight[80]. This suggests that different secondary structures show varying dependence on steroid hormones and somatic sex identity [63,81].This further suggests that differences in body mass due to muscle size are genetically controlled and are likely under the regulation of CASI genes.

Retroviruses have repeatedly infected vertebrates throughout evolution. The integration of proviruses into the germline has allowed for vertical transmission to the progeny. During this process retroviral genes *gag, pol* and *env* are “endogenised”, a process whereby accumulating mutations disrupt these genes. Occasionally one of these genes remains functional, contributing to the physiology of the host (Figure 4.1). Most notable is the *env* (envelope) endogenous retroviral gene (ERV) gene. *Env* encodes an envelope glycoprotein found to facilitate cell-cell fusion during

infection, thereby facilitating viral entry into the host cell. The endogenised env-derived gene often retains properties of the original gene that are beneficial to the host. For example, the env-derived genes confer resistance to infecting viruses of the same family by saturating viral cell receptors on the host cell membrane[82]. The endogenised env genes are also commonly referred to as syncytins, producing fusogenic proteins that mediate the formation of a syncytium; a multinucleated cell formed through fusion of neighbouring cells (Figure 4.1). A well known example of the fusogenic function of syncytins is placenta formation in vertebrates. In humans, HERV-W and HERV-FRD env proteins commonly known as syncytin 1 and 2 respectively are responsible for placenta syncytiotrophoblast formation[83,84]. All major clades of placental mammals have independently captured syncytin, these have been shown to have similar functions but are not orthologous[82].

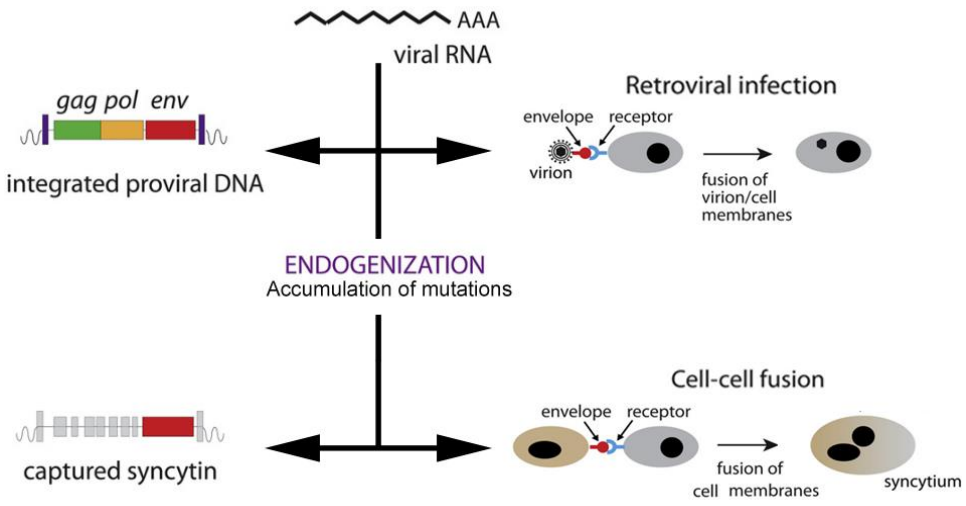


Figure 4. 1. Retroviral endoginisation. Integration of the proviral DNA into the germ line allows for vertical transmission to progeny. Top: Proviral DNA actively making infectious virions using the env protein to facilitate entry into uninfected cells. Bottom: Due to the accumulation of mutations, most of the proviral genes are disabled. The remaining active *env* gene is expressed on the host cell membrane providing a new function beneficial to the host eg. Cell-cell fusion. Image adapted from Dupressoir, et al., 2012.

The FET-1 predicted protein showed similarity to the syncytin precursors of multiple species, including human, mouse, zebrafish and rabbit. It is likely that chickens independently captured a syncytin-related retroviral *env* gene, FET-1. However, FET-1 is unlikely to play a similar reproductive role to syncytins in mammals since chickens have no need for placenta formation. Additionally, since no cell-cell fusion occurs during gonad development, it is therefore unlikely for FET-1 to have a fusogenic function in the gonads. It is however likely that the *FET-1*-like lincRNA transcripts contribute to gonad development.

It has previously been established that two-thirds of genomic DNA is transcribed, with only <2% translated into protein[85]. The assumption that developmental complexity correlates with increased protein-coding genes has been shown to be incorrect, with complexity seeming to correlate with the non-coding transcriptome of the particular organism[86]. Non-coding RNAs (ncRNAs) are RNA molecules that do not code for protein and have recently been shown to function in a regulatory capacity at all stages of development. There are several diverse types of lncRNAs which can be categorized in several different ways. Based on location five groups can be identified: 1) natural antisense transcripts found on the anti-sense strand of transcript producing genes, mostly overlapping in the 5' (promoter) and 3' (terminator) sequences. 2) Pseudogenes, genes that have lost their coding potential due to loss-of-function mutations. A portion of the pseudogene is expressed (2-20%) to regulate gene expression often that of its ancestral gene. 3) Long intronic ncRNA found within the annotated introns of known genes, observed to be differentially expressed, respond to stimuli and misregulated in cancer. 4) Promoter-associated and enhancer RNAs produced from the region of the transcription start site (TSS)

and the enhancer sequences respectively. Their functions are currently unclear. 5) Stand-alone lncRNA, often referred to as long intergenic/intervening ncRNAs (lincRNAs). These are found in between protein-coding genes and are almost always processed as mRNAs; capped, polyadenylated and spliced[87].

LincRNAs are a form of lncRNAs, defined by being localized between protein-coding genes[88]. LncRNAs have been shown to contribute to the sex determination in variety of organisms including *Drosophila melanogaster*, yeast and mice. Given the diverse set of mechanisms lncRNAs use to fulfil their biological function[88] it can be challenging to determine their exact mechanism of action. Regardless, multiple lincRNAs have been shown to contribute to the expression of major components within the canonical sex-determining pathways such as SRY and *DMRT1*[89]. Specific lincRNAs have also been shown to play a role in dosage compensation. In mammals this involves *Xist* acting in *cis* to induce heterochromatin on the X chromosome[90]. The anti-sense lincRNA to *Xist, Tsix*, induces repressive epigenetic modifications to the *Xist* promoter[85]. A similar gene is observed in chickens that may induce localized dosage compensation through the silencing of selected chromosomal regions of importance to sex determination on the Z chromosome. This refers to the *MHM* lncRNA accumulating upstream of the *DMRT1* gene, acting to prevent transcription factor binding through steric hindrance and/or conformation changes of the chromatin[55,91]. *DMRT1* itself acts in conjunction with an lncRNA to regulate its own function. In mice *DMRT1* acts as a transcriptional activator and repressor, participating in *trans* splicing (the joining of two or more different mRNA transcripts) with the lncRNA *Dmr*. This results in a *DMRT1* protein with an altered

carboxy terminus, which when over expressed mimics *DMRT1* loss-of-function phenotype[89]. This clearly acts to negatively regulate *DMRT1* expression.

lncRNA is also implicated in sex hormone function. The lncRNA *SRA* (Steroid receptor RNA Activator) has been shown to modulate ERα[92], androgen receptor and progesterone receptor function by direct association with their receptors[89] and recruitment of transcriptional activators and repressors. *SRA* can also be alternatively spliced into *SRAP/SRA1* to enhance steroid hormone mediated gene expression. Even though *SRA* is classified as a lncRNA, the human *SRA* was shown to produce a protein conserved within other vertebrates including chicken[93]. The *SRA* lncRNA and protein function independently of one another, although the exact mechanisms are still unknown. In addition is sex determination, lncRNAs also play a role during other developmental processes, such as patterning via HOX gene regulation, brain and central nervous system development as well as various organs, including the heart, skeletal muscle, skin and adipose tissue[85]. A summary of lncRNA transcripts involved in sex determination and general development can be found in Appendix 2, Table A2.1 and Table A2.2, adapted from Taylor et al, 2015.

The high homology between the *FET-1*-like lincRNA populations themselves and *FET-1* makes it difficult to distinguish between their respective potential functions. RT-PCR expression data shown supports gonadal staining and a likely role during sex determination and/or gonad differentiation and is in agreement with previously published data[67,78] supporting gonadal expression. ISH expression data supports the possibility of pectoral muscle expression. Further work is required to elucidate their

function. A possible method would be to investigate the additive loss of function during knockdown of the individual *FET-1*-like populations and the full-length *FET-1*. Alternatively, lincRNA interactions with that of DNA and proteins can be investigated with techniques such as RNA and DNA immunoprecipitation and variations thereof [88].

Another retroviral integration, *OVEX1*, is expressed in a female-specific manner, reflecting the right-left asymmetric development of the gonads[94]. This gene has been shown to be degenerated, indicative of endogenisation. This gene was proposed to contribute to fertility and folliculogenesis after expression was observed in somatic cells associated with the developing germ cells and the follicles. ERVs have been shown to regulate “stemness” and pluripotency through various mechanisms such as serving as alternative promoters for pluripotency genes, acting as long range enhancers, generating cell-specific lncRNAs that can bind proteins to regulate the pluripotency state and rewiring the gene regulatory networks by controlling pluripotency genes[95,96]. *FET-1* may not necessarily contribute to primary sex determination. Hirst et al, 2017 localized *FET-1* to the cortex; *FET-1* may therefore be involved in developmental processes related to the cortex, such as cortex differentiation and/or germ cell differentiation from the PGCs, similar to *OVEX1*.

While FET-1 is unlikely to have fusogenic function in the gonads, it has recently been shown that syncytin contributes to myocyte fusion in the developing pectoral muscle[97]. The process of muscle development in the chicken is a multistep process, taking place over four stages. The first stage involves the differentiation of myogenic

precursors from the somites. During the second stage the precursors proliferate and differentiate into myoblasts; the myoblasts differentiate and give rise to the myotubes during the third stage; and finally, the maturation of the myotubes into myofibers at the fourth stage[98]. It is during the third stage that the myoblasts exit the cell cycle and cell-cell fusion occurs, mediated in part by syncytins to form multi-nuclear myofibers. Syntycin was shown to play a sexually dimorphic role in pectoral muscle development in mice[97]. *In vivo* syncytin-B knockout resulted in a 20% reduction in muscle mass only observed in males, subsequently showing a female-like phenotype. *Ex vivo* knockout of both syncytin A and B in mice and the cognate syncytins in humans, sheep and dogs showed a reduction in myoblasts fusion[97]. The similarity of FET-1 to syncytin precursors of mice, human (Figure 3.12) and rabbit (Figure 6A3.1a and b) as well as the male-specific expression suggests similar sexually dimorphic function to that observed in mice. It is therefore possible that *FET-1* has taken a CASI role during pectoral muscle development. It could be hypothesised that reduced FET-1 expression will produce a female phenotype due to a decrease in cell-cell fusion in the developing pectoral muscle.

LincRNAs have also been shown to contribute to muscle development. One of the first lincRNAs identified to contribute to skeletal muscle development is Linc-MD1. Identified in the mouse, this lincRNA was found to be expressed before syncytium formation in the myoblasts, functioning as a competitive endogenous RNA (ceRNA) to sequester miRNA and regulate myoblast differentiation[98]. LincRNAs have additionally also been identified in chicken[98,99] and pig[100] skeletal muscle; however the potential sexually dimorphic expression was not considered for either. The sexually dimorphic expression of *FET-1* in the pectoral muscle shown above

suggests sex-specific function. It should also be noted that the sense transcript is expressed in female pectoral muscle. Given the sexually dimprohic nature of chicken pectoral muscle, it is possible that the sense transcript could act as an anti-sense transcript to maintain a female phenotype. Anti-sense transcripts have been implicated in regulation of their sense counterparts thorough various mechanism such as regulating transcription initiation, mediating chromatin modification and affecting the stability of their sense counterpart just to name a few[101]. This is, however, speculation until these transcripts can be fully investigated via expression and functional studies.

In conclusion, *FET-1* and the *FET-1*-like lincRNA populations have been shown to be expressed in a sexually dimorphic manner. Although their specific function is still unknown, there is evidence for their presence and possible function from multiple other ERVs and lncRNA in both the gonads and muscle. It has been shown that a single lncRNA can utilize different mechanisms depending on the cell type[85]. *FET-1* may therefore have independent functions in both gonads, contributing to fertility similar to OVEX1; and the pectoral muscle, in a CASI capacity.

CHAPTER 5 – FUTURE WORK

FET-1 clearly has a sex-dimorphic role during avian sex determination. Future work will involve investigating the precise function of *FET-1* and the *FET-1*-like lincRNAs in the female gonad and male pectoral muscle; how/where *FET-1* influences the main sex determining regulatory networks and CASI. Due to the nature of these transcripts and the fact that the two *FET-1*-like lincRNA populations overlap *FET-1*, it can be challenging. One possible method is shRNA-based knockdown functional studies of the two groups independently. This would have to be performed in parallel with the shRNA based knockdown of *FET-1* to distinguish between *FET-1* function and *FET-1*-like lincRNA function. The *FET-1*-like lincRNA knockdown will expectantly have an additive effect upon the knockdown of *FET-1* alone and/or will produce a slightly different morphology than that of *FET-1* knockdown. Additionally, the function of these lincRNAs can be investigated on many different levels[88]: 1) DNA interaction where the DNA-RNA interaction can be investigated using chromatin RNA immunoprecipitation (ChRIP); 2) protein interaction where identification of the protein partners of lincRNAs can provide insight as to their function. RNA pull down assaya would allow one to label a RNA probe with high affinity tags e.g. biotin. These RNA probes would be used in a cell lysate to bind potential protein binding partners and purified to identify the protein. This could potentially be confirmed under *in vivo* conditions using antibodies against the proteins identified; and 3) structural features, ncRNA form specific secondary and tertiary structures specific to their functions can be investigated.

The use of anti-syncytin Abs from syncytin orthologues, such as rabbit or mouse can be helpful to provide insight into the expression profile of the predicted protein, if expressed. Since antibody specificity is more dependent on tertiary structure than primary sequence, a proven antibody that binds a syncytin-related epitope may be more insightful than the custom FET-1 antibody that binds a range of protein containing the common transmembrane sequence found in many retroviruses.

CHAPTER 6 – APPENDICES

APPENDIX 1

Table 6A1. 1. Summary of lncRNA involved in sex determination and reproduction. Table adapted from Taylor et al.,2015

Reproductive process	lncRNAs involved	Mechanisms
Gonadogenesis	Sxl promoter proximal RNAs	Activates Sxl (sex-lethal) and recruits chromatin modifiers like Polycomb and Trithorax.
Sex determination	SRY	Acts as a miRNA sponge that competitively binds miR-138 to potentially influence sex determination.
	Dmr	Regulates splicing of *DMRT1* (Doublesex and Mab-3 related transcription factor 1).
Sex hormone responses	enhancer RNAs from ESR-bound enhancers	Controls estrogen receptor-regulated enhancer activity.
	PCGEM1	Binds androgen receptor and lncRNA PRNCR1 (prostate cancer- associated non-coding RNA 1); alters chromatin topology affecting androgen responses.
	PRNCR1	Binds androgen receptor and methyltransferase DOT1L (DOT1-like histone H3K79 methyltransferase); alters AR modification state and chromatin topology affecting androgen responses.
	SRA	Binds to and modulates activity of androgen receptor, estrogen receptor, and progesterone receptor; binds additional transcriptional regulators.
Spermatogenesis	mhrl	Regulates Wnt signaling in spermatogonial cells.
	Tsx	Suppresses apoptosis of pachytene spermatocytes.
Oogenesis	oskar	Non-coding functions of oskar mRNA required for early establishment of oocyte polarity.
	XIsirt and VegT	Required for cytoskeletal organization and oocyte polarity in Xenopus.
Placentation	HELLP and SPRY4-IT1	Controls cell survival and migration of trophoblast cells.
	Airn	Regulates imprinted expression of Igf2r, Slc22a2, and Slc22a3, which influence placental growth.
	Kcnq1ot1	Regulates imprinted expression of eight linked genes that influ ence placental growth.

Table 6A1. 2. Summary of lncRNA involved during the development of various organ systems. Table adapted from Taylor et al.,2015.

Tissue	lncRNA	Function
Cardiovascular	Fendrr	Regulates cardiac development
	Braveheart	Regulates cardiovascular development
	tie1AS	Regulates TIE1 (tyrosine kinase with immunoglobulin-like and EGF-like domains 1) and vascular development
Hematopoietic	H19	Maintains hematopoietic stem cell quiescence
	lncRNA-aGT	Necessary for embryonic to adult alpha-globin switching
	7 species	Controls terminal erythroid differentiation
Musculoskeletal	*SRA*	Enhances myogenic differentiation and myogenic conversion of non-muscle cells
	lincMD1	Enhances myoblasts differentiation
	Dum	Regulates myoblast differentiation
	MUNC	Induces myoblast differentiation
Neural	Six3OS	Regulates SIX3 (SIX homeobox 3) and retinal development
	TUNAR	Regulates pluripotency and neural differentiation
	Evf2	Represses DLX5 (distal-less homeobox 5) and controlling GABA circuitry
	Pnky	Regulates neurogenesis from neural stem cells
Mammary	Pinc1	Regulates alveolar development
	Zfas1	Regulates alveolar development
Endoderm	DEANR1/LINC00261	Regulates endoderm differentiation
Adipose	HOTAIR	Regulates preadipocyte differentiation
	10 species	Regulates preadipocyte differentiation

APPENDIX 2

Figure 6A2. 1. Pairwise alignment of *FET-1* (AY113681.1) to the proposed chromosome 4 location (NC_006091.4). The blocked red sequences represent the alignment of the 5' and 3' ends of the *FET-1* cDNA to the chromosome 4 location. There is a similarity of 46.7% between the two sequences with chromosome 4 missing the sequence of the *FET-1* CDS

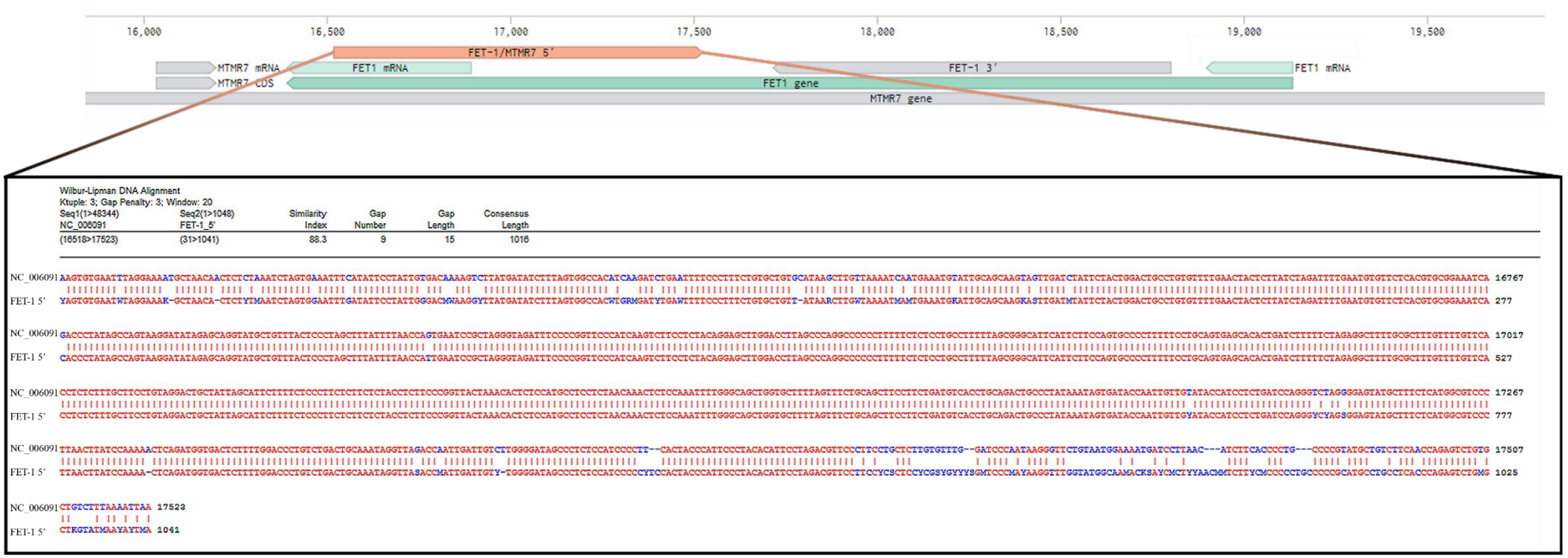


Figure 6A2.2 a. Confirmation of the chromosomal 4 *FET-1* sequence. Touchdown semi-nested PCR was performed using the gene overlapping the annotated *FET-1* on chromosome 4, to ensure amplification from chromosome 4. Sequencing of the semi-nested *FET-1* PCR product confirmed the absence of the *FET-1* CDS, as seen in Figure 6A2.1. The orange bar is representative of the 5' end of the *FET-1* semi-nested PCR product that used the same forward primer during both the touchdown *MTMR7* PCR and the semi-nested *FET-1* PCR. The lower part is the pairwise sequence alignment of the *FET*-1 chromosomal region on chromosome 4 (NC_006091) and the 5' end of the semi-nested *FET-1* PCR product (*FET-1* 5'). The pairwise alignment was performed using the Wilbur-Lipman DNA alignment, with the red nucleotide sequence indicating base matches, while the blue indicates mismatches.

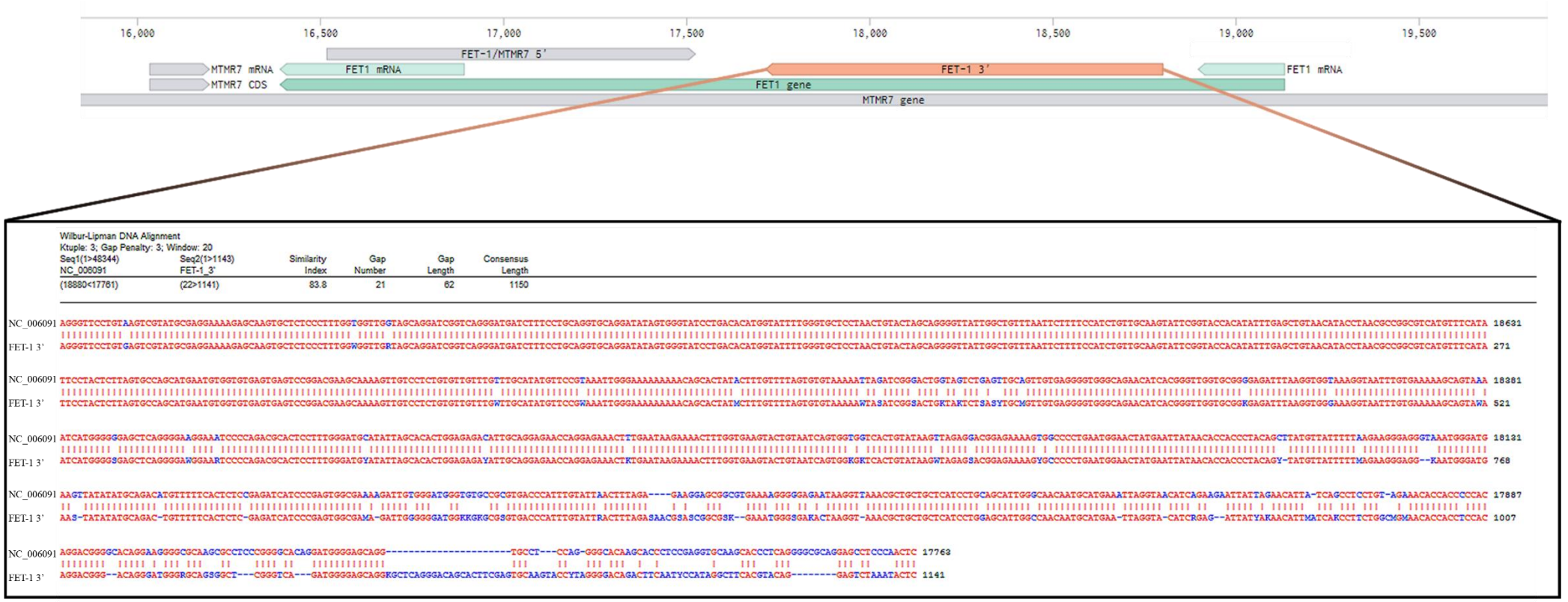


Figure 6A2.2 b. Confirmation of the chromosomal 4 *FET-1* sequence. Touchdown semi-nested PCR was performed using the gene overlapping the annotated *FET-1* on chromosome 4. Sequencing of the semi-nested *FET-1* PCR product confirmed the absence of the *FET-1* CDS, as seen in Figure 6A2.1. The orange bar is representative of the 3' end of the *FET-1* semi-nested PCR product that used a primer specific to the annotated *FTE-1* on chromosome 4. This primer was used on the amplification product of the *MTMR7* touchdown PCR, ensuring amplification from chromosome 4 and not any of the other possible *FET-1* integrations. The lower part is the pairwise sequence alignment of the *FET*-1 chromosomal region on chromosome 4 (NC_006091) and the 3' end of the semi-nested *FET-1* PCR product (*FET-1* 3'). The pairwise alignment was performed using the Wilbur-Lipman DNA alignment, with the red nucleotide sequence indicating base matches, while the blue indicates mismatches.

APPENDIX 3

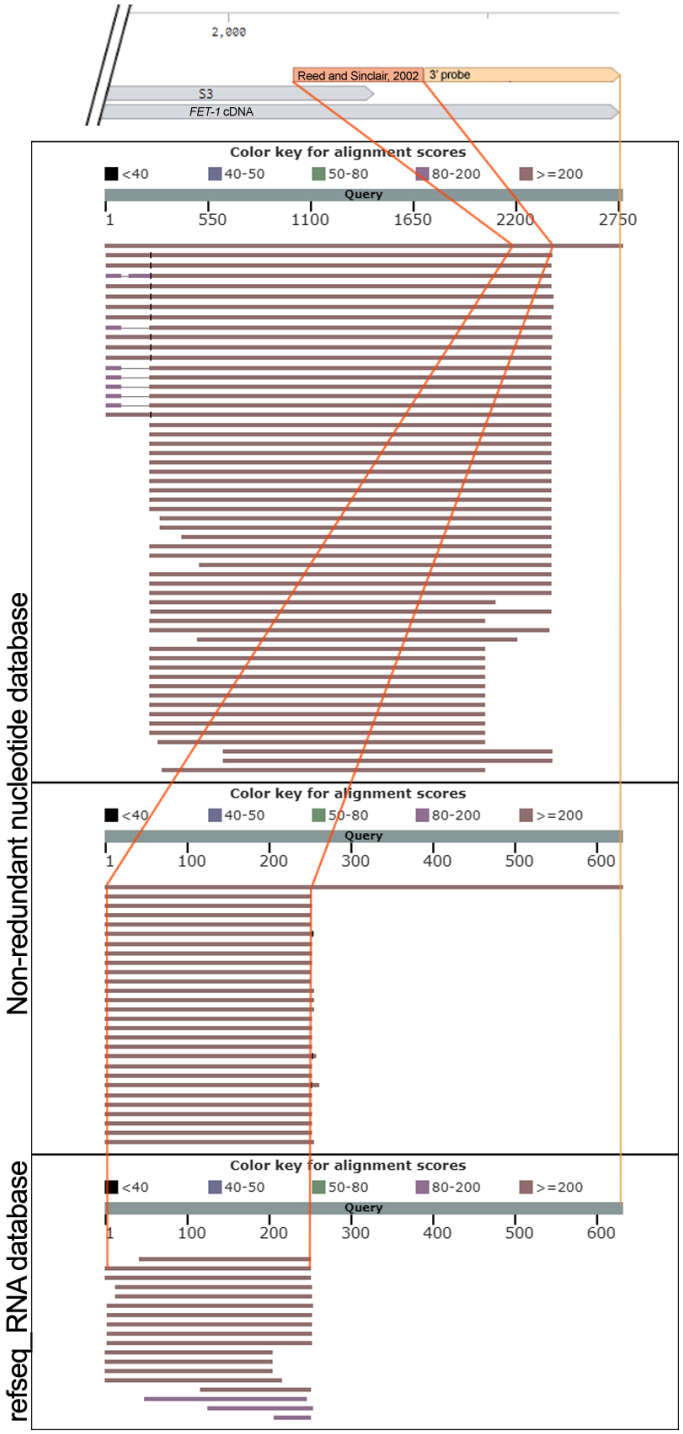


Figure 6A3. 1. Reed and Sinclair 3' probe analysis. The 3' probe Reed and Sinclair,2002 used for ISH was designed against the last 630 bp of the *FET-1* transcript (AY113681.1), represented by the top panel, with the S3 probe used in this project for perspective. The orange segment of the probe is represented in both the non-redundant nucleotide (nr/nt) database and the reference sequence (refseq) RNA database, while the yellow segment does not appear in any of the databases. The second and third panels represent the hits from a nucleotide BLAST analysis against the nr/nt database. In both these panels, the first hit is the *FET-1* transcript submitted to the database by Reed and Sinclair, 2002, aligning to the entire length of the 3' probe. Second panel: The full *FET-1* transcript showing the absence of the *FET-1* 3' end in the nr/nt database. Third panel: the full 3' probe against the nr/nt database (not all hits are shown). The fourth panel represents the BLAST analysis against the refseq RNA database. Since *FET-1* has been retracted from the database, it was not considered during this BLAST analysis. Similar to the second and third panels, hits are only present against the 5' end of the probe.

APPENDIX 4

Subject name	Gene hit	Subject start	Subject end	Subject ori	Genomic Location	Orientation	Query start	Query end	Length	Score	E-val	%ID
ENSOCUP00000026746	ENSOCUG00000029381	294	503	Forward	X:62571216-62571845	Forward	155	361	218	84.0	2e-17	27.52
ENSOCUP00000026867	ORY1	296	505	Forward	12:99950019-99950648	Forward	155	361	218	80.5	3e-16	26.61

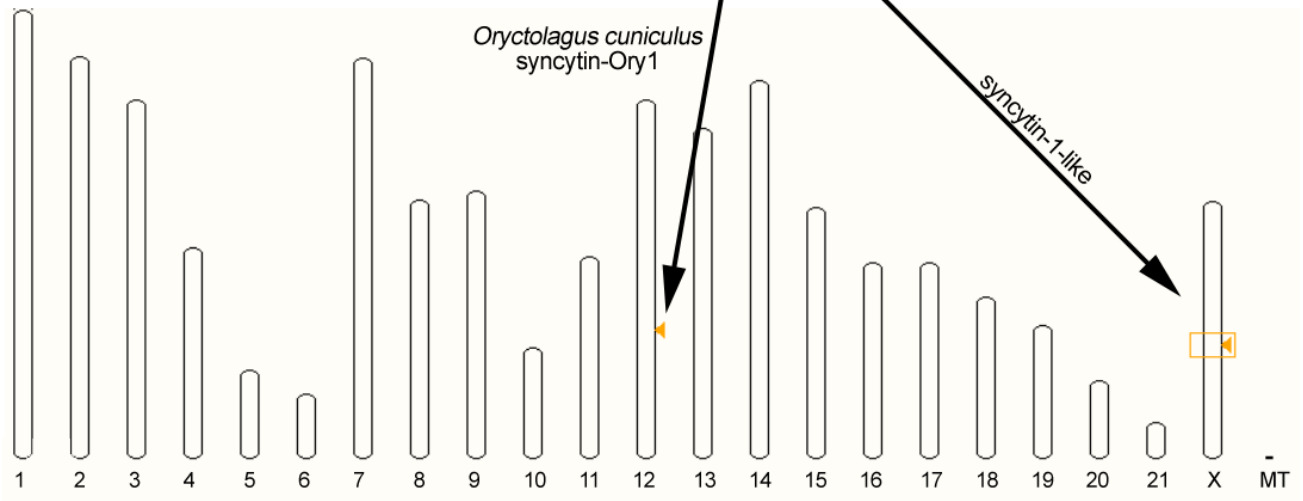
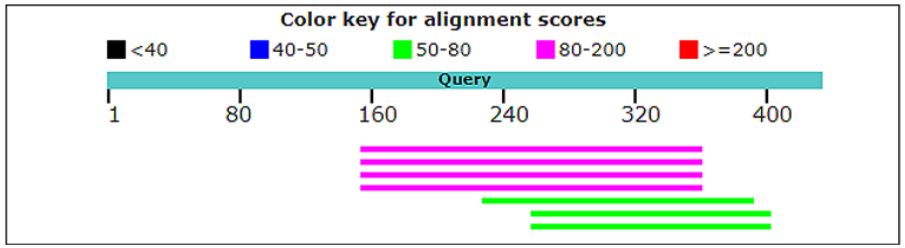


Figure 6A4.1 a. Local alignment of *FET-1* predicted protein against the Ensembl rabbit (*Oryctolagus cuniculus*) protein database. Top: Local alignment results showing hits of 27% and 26% identity against the rabbit syncytin-Ory1 on chromosome 12 and the syncytin-like-1 on the X chromosome. Bottom: Visual representation of alignment hit locations in the rabbit genome



	Description	Max score	Total score	Query cover	E value	Ident	Accession
■	PREDICTED: syncytin-2-like [Oryctolagus cuniculus]	84.0	84.0	47%	1e-17	27%	XP_017198287.1
■	PREDICTED: syncytin-2-like [Oryctolagus cuniculus]	83.6	83.6	47%	1e-17	27%	XP_017202187.1
■	PREDICTED: syncytin-1-like [Oryctolagus cuniculus]	84.0	84.0	47%	4e-17	28%	XP_017205165.1
■	syncytin-Ory1 precursor [Oryctolagus cuniculus]	82.8	82.8	47%	1e-16	27%	NP_001292522.1
■	PREDICTED: syncytin-2-like [Oryctolagus cuniculus]	78.6	78.6	38%	3e-16	32%	XP_017194466.1
■	PREDICTED: uncharacterized protein LOC103345346 isoform X1 [Oryctolagus cuniculus]	64.3	64.3	33%	7e-11	31%	XP_008247495.1
■	PREDICTED: uncharacterized protein LOC103345346 isoform X2 [Oryctolagus cuniculus]	63.9	63.9	33%	8e-11	31%	XP_017193518.1

Figure 6A4.1 b. Local alignment of *FET-1* predicted protein against the NCBI rabbit (*Oryctolagus cuniculus*) protein database. Top: graphic representation of local alignments against FET-1. Bottom: Local alignment hit names. Similar to the Ensembl search, syncytin shows between 27-32% similarity to FET-1. Additionally, two isogfoms of an uncharacterized protein shows 31% similarity to FET-1.

CHAPTER – 7 RECIPIES AND PROTOCOLS

A. CRUDE DNA EXTRACTION

1. 3x3x3mm tissue was collected from the from embryo
2. 100 µl lysis solution was added to tubes containing tissue

Crude lysis solution	
1M Tris-HCl (pH 8.0)	200 µl
0.5 M EDTA (pH8.0)	80 µl
1 M NaCl	4 ml
dH2O	15.72 ml
Total	20 ml

3. 1µl 20mg/ml Proteinase K was added
4. Vortexed briefly
5. Incubated in water bath/thermo cycler:
    a. 55°C for 15 min
    b. Vortex briefly
    c. 55°C 15 min
    d. 95°C for 5 min
6. Vortexed briefly
7. Centrifuged for 1 min at 10 000g
8. Used 1 µl of supernatant for PCR

B. PREPARATION OF DIG-LABELLED RIBOPROBES

Linearise DNA:

1. The plasmid containing the DNA of interest (10µg) was cut with an appropriate restriction enzyme according to the manufacture.

   The restriction enzyme should be cut distal of the promoter for transcription on the sense/ anti-sense strand.

2. The digested plasmid was cleaned with a PCR clean-up kit

3. Probe synthesis mix

Linearised plasmid (1 µg in DEPC ddH2O)	26.25 µl
5x Transcription Buffer	10 µl
0.1M DTT	5 µl
10x DIG RNA labelling mix	5 µl
RNase inhibitor	1.25 µl
RNA polymerase (T7/T3/SP6)	2.5 µl
Total	**50 µl**

4. Incubated for 2 hrs at 37°C in water bath/thermo cycler

After 1 hour another 1µl of the RNA polymerase was added

5. Checkpoint

Riboprobe synthesis mix	1 µl
DEPC ddH2O	3 µl
Loading Dye	2 µl
Total	**5 µl**

The solution was boiled for 3 min at 65°C to remove secondary structures, chilled on ice and electrophoresed on 1% agarose gel

Expected single RNA band + DNA band (RNA should be 5-10x more material than DNA)

6. Added 2 µl RNase-free DNase + 2µl DNase buffer
7. Incubated for 1-2hrs at 37°C in water bath/ thermo cycler
8. Cleaned riboprobe mix with RNA clean up/concentrator kit
9. Checkpoint

Riboprobe synthesis mix	1 µl
DEPC ddH2O	3 µl
Loading Dye	2 µl
Total	**5 µl**

The solution was boiled for 3 min at 65°C to remove secondary structures, chilled on ice and electrophoresed on 1% agarose gel

Expected single RNA band

Added 1ml hybridization mix to the clean riboprobe and stored at -20˚C

C. WHOLE MOUNT IN SITU HYBRIDIZATION

Fixing and dehydrating the embryos

1. Collected embryos and incubated in 4% PFA at 4°C overnight
2. Removed the 4% PFA solution
3. Washed in DEPC-PBT; shaked on nutator for 5-15mins.
   Discarded the solution and repeated this step three times
4. Dehydrated embryos – 15mins on nutator each
    a. 25% MeOH + PBT/DEPC
    b. 50% MeOH + PBT/DEPC
    c. 75% MeOH + PST/DEPC
    d. Added 100% MeOH. Agitated for 15mins.
    e. Discarded the solution and repeated this step three times
    f. Added 100% MeOH. Stored at -20^{0}C

Pre-treatment and Hybridisation- DAY 1

1. Rehydrated embryos – 5-10 mins each wash
    a. 75% MeOH + PBT/DEPC
    b. 50% MeOH + PBT/DEPC
    c. 25% MeOH + PBT/DEPC
2. Washed twice with PBT/ DEPC for 5mins
3. Washed again with PBT/DEPC for 10mins
4. Treated with 10µg/ml Proteinase K in PBT/DEPC with **NO SHAKING**
5. Incubation time depended on the stage of the embryo

No. of somites	Stage	Time (min)
4-6ss	8	8mins
7-9ss	9	9mins
10-12ss	10	10mins
13-15ss	11	11mins
20-22ss	14	14mins
3days	20	30mins

6. Washed in 2mg/ml glycine in PBT/DEPC for 10mins. **NO SHAKING**
7. Washed twice with PBT/DEPC for 5mins. **NO SHAKING**
8. Post-fixed with 4% PFA+0.2% glutaraldehyde for 20mins at room temperature
9. Rinsed 4X with PBT/DEPC for 5mins
10. Washed in PBT/DEPC:Hybridisation Mix (1:1) for 10min at room temperature
11. Washed in 100 % hybridization Mix for 10min at room temperature; added another hybridisation mix wash if necessary to get close to 100% hybridisation Mix as possible.

In Hybridisation oven

12. Changed to new hybridisation solution and pre-hybridised: incubated at 70˚C for at least 3 hours
13. Hybridisation: added new pre-warmed (70ºC) hybridisation solution + RNA probe (1-10µl/ml hybe). Incubated at 70ºC overnight (at least 16 hours). Positioned tubes horizontally if using an oven.

Post-hybridisation washes (most important for specificity)-DAY2

In Hybridisation oven

14. Used hybridisation mix with no tRNA added
15. Removed the probe in hybridisation solution and stored it at -20ºC
16. Washed 2X with pre-warmed hybridisation solution (50ml tube O/N in the oven) at 70ºC for 15 min
17. Washed 4X with warmed hybridisation solution at 70ºC for 30-45mins
18. Washed with Hybridisation mix:MABT (1:1) @ 70ºC for 30mins
19. Washed 4X with MABT for 30 minutes at room temperature
20. Changed the MABT to MABT+3%BSA for 15 mins at room temperature (blocking solution)
21. Replaced with fresh MABT + 3% BSA for 1 hour at room temperature
22. Changed to Anti-DIG-AP Ab (1:2000) in MABT+ 3% BSA for 5 min at room temperature
23. Replaced with fresh Anti-DIG-AP AB in MABT + 3% BSA and leave overnight at 4ºC on the shaker

Post-Ab washes and Histochemistry- DAY3

1. Rinsed 2X with MABT for 5mins each wash at room temperature
2. Washed 2X with MABT for 30mins each at room temperature
3. Washed 6X with MABT for 1 hour each wash at room temperature
4. Incubated at 4^{o}C overnight

Shaker- DAY4

1. Washed 2X with 100mM Tris-Cl (pH 9.5) for 15 min each at room temperature
2. Washed 2-4X with NTMT 15 min at room temperature
3. Changed the NTMT solution to NBT/BCIP in NTMT (filtered). Ensured that the vials were covers with foil. Incubated at 4^{o}C overnight
   NB: for some embryos, the developing time was short
   After the desired colour had developed, proceed with the washing steps.
4. Washed twice with 100mM Tris-HCl pH 9.5 for 5mins each at room temperature (keep dark – cover with foil)
5. Wasedh 3X with PBT for 5mins each at room temperature covered in foil
6. Changed to 4% PFA, leave at 4^{o}C overnight OR 2 hours room temperature covered in foil
7. Washed 3x with PBT for 10mins at room temperature covered in foil
8. Changed solution to MeOH in PBT- 5mins each
   a. 25% MeOH + PBT/DEPC
   b. 50% MeOH + PBT/DEPC
   c. 75% MeOH + PBT/DEPC
   d. Washed twice with 100% MeOH for 5mins each
9. Kept at 4^{o}C (hermetically closed)
10. Stored at -20^{o}C OR rehydrated gradually to PBT for photographing
11. Rehydrated embryos before taking pictures
12. Changed solution to MeOH + PBT – 5mins each
    a. 75% MeOH + PBT/DEPC
    b. 50% MeOH + PBT/DEPC

c. 25% MeOH + PBT/DEPC
d. Washed 3X with PBT for 5 mins each

13. Processed to 50% glycerol for photographing
    a. 10% glycerol for 10 min
    b. 25% glycerol for 10 min
    c. 50% glycerol for 10 min

IN SITU HYBRIDIZATION REAGENTS

10X PBS (1L)

$Na_2HPO_4.7H_2O$ 11.5 g

NaCl 80g

KH_2PO_4 2g

KCl 2g

Make up to 1L ddH_2O

pH 7.2

4% PFA 10ml

1.1 ml of 37% formaldehyde

8.9 ml of PBS + 2 mM EDTA

Hybridization mix (500ml) DNAase/ RNAse free

Reagent	Volume	Final concentration
Formamide	250ml	50%
20X SSC/DEPC (pH 5 citric acid)	32.5ml	1.3X
0.5 EDTA (pH 8)	5ml	5mM
tRNA (20mg/ml in H2O/DEPC	5ml	200µg/ml
Tween-20	1ml	0.2%
10%CHAPS in H2O/DEPC	25ml	0.5%
Heparin	50mg	100µg/ml

Fill up to 500ml with DEPC ddH_2O

5X MAB(T) 1l 1XMAB + 0.1% Tween-20

Maleic Acid 58g

ddH_2O 500ml

pH to 7.5 with Tris-base 100-150g

NaCl 43.5g

20% Tween-20 5ml or 100% Tween-20 1ml

Fill up to 1l with ddH_2O OR 50µL Tween-20 + 50ml 1X MAB

10% Boehringer blocking reagent in MAB (no Tween-20) 200ml

BBR 20g

1X MAB 180ml

Heat to 70 to dissolve

Autoclave or microwave to boil

Aliquot in 1ml

Freeze at -20^oC

PBT/DEPC (make fresh)

PBS/DEPC with 0.1%Tween-20

20X SSC (200ml) RNAse/ DNAse free

NaCL 35g

Nacitrate 17.6g

pH with 1M citric acid to 5.0

Fill up to 200ml with DEPC ddH_2O

NTMT(50ml)

Reagent	V1(50ml)	V2(1L)	Final concentration
5M NaCl	1ml	20ml	100mM
1M Tris-HCl pH 9.5	5ml	100ml	100mM
1M MgCl2	2.5ml	50ml	50mM
10% Tween-20	0.5ml	10ml	0.1%

Fill up to final volume with ddH_2O

Glycine

0.1g in 50ml DEPC-H_2O

Make fresh immediately prior to use

D. IMMUNOHISTOCHEMISTRY ON SECTIONS

TISSUE SECTIONS - Embryo preparation

Dissected embryos and fixed in 4% PFA for 2-3 hours at room temperature or overnight at 4˚C.

Washed 3X in PBS for 5 min

Cryoprotect:

5% sucrose for 6 hours

15% sucrose overnight at 4˚C

Incubated in embedding medium for a minimum of 6 hours or overnight at 4˚C

Flash freezed in liquid nitrogen

Sectioned embryos in 10-15 µm sections

Placed sections on microscope slide and let dry overnight

* Sections could have been stored at -20˚C until required for IHC

Removal of embedding medium

Washed 2X 5min in PBS at room temperature

WHOLE MOUNT

Dissected embryos and fixed in 4% PFA for 2-3 hours at room temperature or overnight at 4˚C.

Washed 3X in PBS for 10 min

* Embryos could have been stored in PBS at 4˚C until needed, ensured that there was not fungal or bacterial growth before use.

IMMUNOHISTOCHEMISTRY – DAB

Given the bigger size of the whole embryo, all washing steps was be doubled in time, as well as the incubation for inhibition of endogenous peroxidise activity. However, this was optimised as needed. Primary and secondary antibody incubation times needed not be adjusted when incubated at 4˚C overnight.

Washed 3X for 5min in TBSTxD

Incubated in 0.3% H_2O_2 in MeOH for 15 min

Washed in TBSTxD for 2X for 5min

Blocked in 3% BSA for 1-2 hrs at room temperature or overnight at 4˚C

Incubated in primary Ab overnight at 4˚C

Washed in TBSTxD 6-8X for 15min

Incubated in secondary Ab for 2-3 hrs at room temperature or 4˚C overnight

Washed 5-6X for 15min in TBSTxD

Incubated in 0.5mg/ml DAB for 20 min

Dissolved 1 tablet DAB (final concentration - 0.5mg/ml) in 20 ml TBST

Incubateed slides/embryos in 1.5 µl 50% H_2O_2 + 1ml DAB for colour development, **in the dark.** Checked for colour development every 10-20 min.

Rinsed in ddH_2O to stop colour reaction

For tissue sections, the coverslip was mounted with 1-2 drops of mounting medium while slide was still wet

Imaged the whole embryos/ tissue sections as appropriate

IMMUNOHISTOCHEMISTRY REAGENTS

TBS

0.8g NaCl

0.2g KCl

3g Tris base

800 ml dH_2O

Dissolve all reagents in water. Adjust pH to 7.4 and add dH_2O to final volume of 1L and autoclave

TBSTxD

TBS + 1% Trition X and 1% DMSO (add after autoclaving)

Blocking solution

3% BSA in TBSTxD

E. SDS PAGE AND WESTERN BLOT

SDS PAGE GEL PREPERATION

The SDS PAGE gel in a single electrophoresis run was divided into a stacking gel and separating gel.

The stacking gel (acrylamide 5%) was poured on top of the separating gel (after solidification) and a gel comb is inserted in the stacking gel. The acrylamide percentage in SDS PAGE gel depended on the size of the target protein in the sample.

Acrylamide %	M.W. Range
7%	50 kDa - 500 kDa
10%	20 kDa - 300 kDa
12%	**10 kDa - 200 kDa**
15%	3 kDa - 100 kDa

Volumes of stacking gel and separating gel differed according to the thickness of gel casting.

Thickness of the gel (mm)	Vol. of stacking gel (ml)	Vol. of separating gel (ml)
0.75	2	4
1.0	3	6
1.5	4	8

5 ml stacking gel

Component	Volume
H2O	2.975 ml
0.5 M Tris-HCl, pH 6.8	1.25 ml
10% (w/v) SDS	0.05 ml
Acrylamide/Bis-acrylamide(30%/0.8% w/v)	0.67 ml
10% (w/v) ammonium persulfate (AP)	50 µl
TEMED	5 µl

10 ml separating gel

Component	Volume				
Acylamide percentage	6%	8%	10%	**12%**	15%
H2O	5.2 ml	4.6 ml	3.8 ml	**3.2 ml**	2.2 ml
Acrylamide/Bis-acrylamide (30%/0.8% w/v)	2 ml	2.6 ml	3.4 ml	**4 ml**	5 ml
1.5M Tris (pH=8.8)	2.6 ml	2.6 ml	2.6 ml	**2.6 ml**	2.6 ml
10% (w/v)SDS	0.1 ml	0.1 ml	0.1 ml	**0.1 ml**	0.1 ml
10% (w/v) Ammonium persulfate	100 µl	100 µl	100 µl	**100 µl**	100 µl
TEMED	10 µl	10 µl	10 µl	**10 µl**	10 µl

Note: AP and TEMED was added right before each use.

1x Running Buffer, volume dependant on electrophoresis system:

Component	Concentration
Tris-HCl	25 mM
Glycine	200 mM
SDS	0.01 % (w/v)

SDS PAGE PROTOCOL

Make the separating gel

The casting frames were set up (clamped two glass plates in the casting frames) on the casting stands.

Prepared the gel solution (as described above) in a separate small beaker.

Swirled the solution gently but thoroughly.

Pipetted the appropriate amount of separating gel solution (listed above) into the gap between the glass plates.

Filled the gap with isopropanol overflow to make the top of the separating gel be horizontal

Waited for 20-30min to let it polymerize.

Make the stacking gel

Discarded the isoproponal to see the separating gel left.

Pipetted in the stacking gel until it overflowed.

Inserted the well-forming comb without trapping air under the teeth.

Waited for 20-30min to let it polymerize.

Made sure the stacking gel was completely polymerized and took out the comb.
Took the glass plates out of the casting frame and set them in the cell buffer dam.
Poured the running buffer (electrophoresis buffer) into the inner chamber and kept

pouring after overflow until the buffer surface reached the required level in the outer chamber.

Prepare the samples

Heated them in boiling water for 5-10 min.

Loaded the prepared samples into wells and make sure not to overflow. Made sure to load the protein marker into the first lane. Then covered the top and connect the anodes.

Set an appropriate volt and ran the electrophoresis.

Total running time: stopped SDS-PAGE running when the down most sign of the protein marker almost reached the foot line of the glass plate. Generally, about 1 hour for a 120V voltage and a 12% separating gel. For a separating gel possessing higher percentage of acylamide, the time was be longer.

WESTERN BLOT

For each gel, one piece of 0.45 nm nitrocellulose membrane was cut and two pieces of filter paper to the dimensions of the gel.

Equilibrated the membranes in Transfer buffer for 15-30 min.

Transfer buffer

Reagents	Volume
Glycine	28.8 g
Tris base	6.04 g
methanol	200 ml
ddH_2O	1.6 L

Dissolve Tris base and glycine together in 1.6 L of ddH_2O.

Add methanol and mix.

Add ddH_2O to a final volume of 2 L.

Protein Blotting by the Capillary Method, adapted from Walker, J.M., 2009

Placed two sheets of 3MM filter paper on a glass plate, and thoroughly soaked them in blotting buffer. The ends were dipped in reservoirs containing blotting buffer to ensure this filter paper pad remained wet overnight.

Placed the gel to be blotted on top of this filter paper bed. Made sure the bed was fully wetted and that no bubbles were trapped between the gel and filter paper. Thoroughly wet the top of the gel with blotting buffer.

A piece of nitrocellulose paper was cut to the size of the gel to be blotted. Wetted the paper by dipping it in blotting buffer. Then lay the nitrocellulose sheet on the gel surface. Took great care to ensure no air bubbles are trapped.

Added dry 3 MM filter paper, cut to the size of the, on top of the nitrocellulose membrane and then placed on top of this a pad of your absorbent material.

Finally, placed a heavy weight on top of the absorbent material, e.g., a sheet of thick glass that supported a 2–3 L flask filled with water (Figure 7E.1)

Allowed blotting to take place overnight (preferably for 24 h).

If the absorbent material is was after an overnight run, it was due to the transfer buffer preferentially moving around the gel. This would have been due to the top layers of filter paper or absorbent material coming in contact with the bottom wet layers of filter paper or buffer tank.

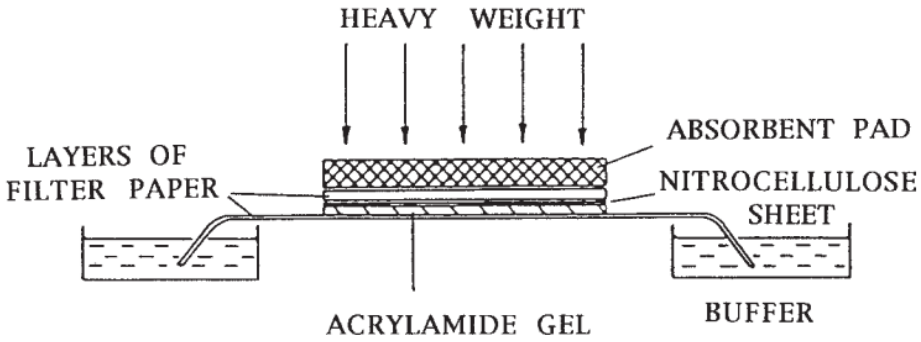


Figure 7E. 1. Ca ɔaper saturated in transfer buffer. The filter paper dipped in buffer reservoirs on both sides to ensure constant saturation. The nitrocellulose membrane is place face down on the DSD PAGE gel. Two to three layer of 3 MM filter paper is placed in the nitrocellulose membrane with absorbent material over the filter paper. A heavy weight is placed on the set up to ensure proper contract between the SDS PAGE gel and nitrocellulose membrane for protein transfer overnight. Image from Walker, J.M., 2009.

CHAPTER 8 – REFERENCES

1. Bull, J. J. Sex determining mechanisms: An evolutionary perspective. *Experientia* **41**, 1285–1296 (1985).

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

3. Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. & Loverll-Badge, R. Male development of chromosomally female mice transgenic for SRY. *Nature* **351**, 117–121 (1991).

4. Smith, C. A. & Sinclair, A. H. Sex determination: Insights from the chicken. *BioEssays* **26**, 120–132 (2004).

5. Nakamura, M. Sex determination in amphibians. *Semin. Cell Dev. Biol.* **20**, 271–282 (2009).

6. Sarre, S. D., Georges, A. & Quinn, A. The ends of a continuum: Genetic and temperature-dependent sex determination in reptiles. *BioEssays* **26**, 639–645 (2004).

7. Devlin, R. H. & Nagahama, Y. Sex determination and sex differentiation in fish: An overview of genetic, physiological, and environmental influences. *Aquaculture* **208**, 191–364 (2002).

8. Vicoso, B., Emerson, J. J., Zektser, Y., Mahajan, S. & Bachtrog, D. Comparative Sex Chromosome Genomics in Snakes: Differentiation, Evolutionary Strata, and Lack of Global Dosage Compensation. *PLoS Biol.* **11**, e1001643 (2013).

9. Gamble, T., Castoe, T. A., Nielsen, S. V., Banks, J. L., Card, D. C., Schield, D. R., Schuett, G. W. & Booth, W. The Discovery of XY Sex Chromosomes in a Boa and Python. *Curr. Biol.* **27**, 2148–2153.e4 (2017).

10. Graves, J. A. M. Avian sex , sex chromosomes , and dosage compensation in the age of genomics. *Chromosom. Res.* **22**, 45–57 (2014).

11. She, Z. Y. & Yang, W. X. Molecular mechanisms involved in mammalian primary sex determination. *J. Mol. Endocrinol.* **53**, R21–R37 (2014).

12. Nakamura, M. The mechanism of sex determination in vertebrates-are sex steroids the key-factor? *J. Exp. Zool. Part A Ecol. Genet. Physiol.* **313 A**, 381–398 (2010).

13. Huang, S., Ye, L. & Chen, H. Sex determination and maintenance : the role of DMRT1 and FOXL2. *Asian J. Androl.* **19**, 619–624 (2017).

14. Eggers, S., Ohnesorg, T. & Sinclair, A. Genetic regulation of mammalian gonad development. *Nat. Publ. Gr.* **10**, 673–683 (2014).

15. Matson, C. K., Murphy, M. W., Sarver, A. L., Griswold, M. D., Bardwell, V. J. & Zarkower, D. DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature* **476**, 101–104 (2011).

16. Lindeman, R. E., Gearhart, M. D., Bardwell, V. J., Lindeman, R. E., Gearhart, M. D., Minkina, A., Krentz, A. D. & Bardwell, V. J. Sexual Cell-Fate Reprogramming in the Ovary. *Curr. Biol.* **25**, 764–771 (2015).

17. Murphy, M. W., Lee, J. K., Rojo, S., Gearhart, M. D., Kurahashi, K., Banerjee, S., Loeuille, G. A., Bashamboo, A., McElreavey, K., Zarkower, D., Aihara, H. & Bardwell, V. J. An ancient protein-DNA interaction underlying metazoan sex determination. *Nat. Struct. Mol. Biol.* **22**, 442–451 (2015).

18. Maatouk, D. M., Dinapoli, L., Alvers, A., Parker, K. L., Taketo, M. M. & Capel, B. Stabilization of β-catenin in XY gonads causes male-to-female sex-reversal. *Hum. Mol. Genet.* **17**, 2949–2955 (2008).

19. Cesar dos Santos, A., Carvalho Viana, D., Benevides de Oliveira, G., Miguel Lobo, L. & Chaves Assis-Neto, A. Intrauterine Sexual Differentiation: Biosyntesis and action of Sexual Steroid Hormones. *Brazilian Arch. Biol. Technol.* **58**, 395–405 (2015).

20. Mank, J. E. & Ellegren, H. Parallel divergence and degradation of the avian W sex chromosome. *TRENDS Ecol. Evol.* **22**, 389–391 (2007).

21. Schmid, M., Smith, J., Burt, D. W., Aken, B. L., Antin, P. B., Archibald, A. L., Ashwell, C., Blackshear, P. J., Boschiero, C., Brown, C. T., Burgess, S. C., Cheng, H. H., Chow, W., Coble, D. J., Cooksey, A., Crooijmans, R. P. M. A., Damas, J., Davis, R. V. N., Koning, D.-J. de, *et al.* Third Report on Chicken Genes and Chromosomes. *Cytogenet. Genome Res.* 78–179 (2015). doi:10.1159/000430927

22. Smith, C. A. Rowley review. Sex determination in birds: A review. *Emu* **110**, 364–377 (2010).

23. Stiglec, R., Ezaz, T. & Graves, J. A. M. A new look at the evolution of avian sex chromosomes. *Cytogenet. Genome Res.* **117**, 103–109 (2007).

24. Cheng, H. Evolution of avian sex chromosomes from an ancestral pair of autosomes. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 8147–8152 (1998).

25. Ellegren, H. Evolution of the avian sex chromosomes and their role in sex determination. *Trends Ecol. Evol.* **15**, 188–192 (2000).

26. Nanda, I., Zend-Ajusch, E., Shan, Z., Grützner, F., Schartl, M., Burt, D. W., Koehler, M., Fowler, V. M., Goodwin, G., Schneider, W. J., Mizuno, S., Dechant, G., Haaf, T. & Schmid, M. Conserved synteny between the chicken Z sex chromosome and human chromosome 9 includes the male regulatory gene DMRT1: a comparative (re)view on avian sex determination. *Cytogenet. Cell Genet.* **89**, 67–78 (2000).

27. Raymond, C. S., Murphy, M. W., Sullivan, M. G. O., Bardwell, V. J. & Zarkower, D. Dmrt1 , a gene related to worm and fly sexual regulators , is required for mammalian testis differentiation. *GENES Dev.* **14**, 2587–2595 (2000).

28. Smith, C. a, Roeszler, K. N., Ohnesorg, T., Cummins, D. M., Farlie, P. G., Doran, T. J. & Sinclair, A. H. The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature* **461**, 267–271 (2009).

29. Raymond, C. S., Kettlewell, J. R., Hirsch, B., Bardwell, V. J. & Zarkower, D. Expression of Dmrt1 in the Genital Ridge of Mouse and Chicken Embryos Suggests a Role in Vertebrate Sexual Development. *Dev. Biol. 215,* **215**, 208–220 (1999).

30. Shan, Z., Nanda, I., Wang, Y., Schmid, M., Vortkamp, A. & Haaf, T. Sex-specific expression of an evolutionarily conserved male regulatory gene , DMRT1 , in birds. *Cytogenet. Cell Genet.* **89**, 252–257 (2000).

31. Lambeth, L. S., Raymond, C. S., Roeszler, K. N., Kuroiwa, A., Nakata, T., Zarkower, D. & Smith, C. A. Over-expression of DMRT1 induces the male pathway in embryonic chicken gonads. *Dev. Biol.* **389**, 160–172 (2014).

32. Smith, C. A., Roeszler, K. N., Hudson, Q. J. & Sinclair, A. H. Avian sex determination : what , when and where ? *Cytogenet. Genome Res.* **117**, 165–173 (2007).

33. Oreal, E., Pieau, C., Mattei, M., Josso, N., Carre, L. E. & Picard, J. Early Expression of AMH in Chicken Embryonic Gonads Precedes Testicular SOX9 Expression. *Dev. Dyn.* **212**, 522–532 (1998).

34. Takada, S., Wada, T., Kaneda, R., Lim, Y., Yamashita, Y. & Mano, H. Evidence for activation of Amh gene expression by steroidogenic factor 1. *Mech. Dev.* **123**, 472–480 (2006).

35. Nakata, T., Ishiguro, M., Aduma, N., Izumi, H. & Kuroiwa, A. Chicken hemogen homolog is involved in the chicken-specific sex-determining mechanism. *Proc. Natl. Acad. Sci.* **110**, 3417–3422 (2013).

36. Govoroun, M. S., Pannetier, M., Pailhoux, E., Cocquet, J., Brillard, J. P., Couty, I., Batellier, F. & Cotinot, C. Isolation of chicken homolog of the FOXL2 gene and comparison of its expression patterns with those of aromatase during ovarian development. *Dev. Dyn.* **231**, 859–870 (2004).

37. Smith, C. A., Smith, M. J. & Sinclair, A. H. Expression of Chicken Steroidogenic Factor-1 during Gonadal Sex Differentiation. *Gen. Comp. Endocrinol.* **113**, 187–196 (1999).

38. Guioli, S. & Lovell-Badge, R. PITX2 controls asymmetric gonadal development in both sexes of the chick and can rescue the degeneration of the right ovary. *Development* **134**, 4199–4208 (2007).

39. Ishimaru, Y., Komatsu, T., Kasahara, M., Katoh-Fukui, Y., Ogawa, H., Toyama, Y., Maekawa, M., Toshimori, K., Chandraratna, R. A. S., Morohashi, K. -i. & Yoshioka, H. Mechanism of asymmetric ovarian development in chick embryos. *Development* **135**, 677–685 (2008).

40. Ayers, K. L., Sinclair, A. H. & Smith, C. A. The molecular genetics of ovarian differentiation in the avian model. *Sex. Dev.* **7**, 80–94 (2012).

41. Nakabayashi, O., Kikuchi, H., Kikuchi, T. & Mizuno, S. Differential expression of genes for aromatase and estrogen receptor during the gonadal development in chicken embryos. *J. Mol. Endocrinol.* **20**, 193–202 (1998).

42. Hudson, Q. J., Smith, C. A. & Sinclair, A. H. Aromatase Inhibition Reduces Expression of FOXL2 in the Embryonic Chicken Ovary. *Dev. Dyn.* **233**, 1052–1055 (2005).

43. Andrews, J. E., Smith, C. A. & Sinclair, A. H. Sites of Estrogen Receptor and Aromatase Expression in the Chicken Embryo. *Gen. Comp. Endocrinol.* **108**, 182–190 (1997).

44. Smith, C. A., Shoemaker, C. M., Roeszler, K. N., Queen, J., Crews, D. & Sinclair, A. H. Cloning and expression of R-Spondin1 in different vertebrates suggests a conserved role in ovarian development. *BMC Dev. Biol.* **8**, 1–16 (2008).

45. Lee, H. C., Lim, S. & Han, J. Y. Wnt/β-catenin signaling pathway activation is required for proliferation of chicken primordial germ cells in vitro. *Sci. Rep.* **6**, 34510 (2016).

46. Chassot, A., Ranc, F., Gregoire, E. P., Roepers-, H. L., Taketo, M. M., Camerino, G., Rooij, D. G. De, Schedl, A., Chaboissier, M., Umana, P. & Generale, B. Activation of b-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. *Hum. Mol. Genet.* **17**, 1264–1277 (2008).

47. Tomizuka, K., Horikoshi, K., Kitada, R., Sugawara, Y., Iba, Y., Kojima, A., Yoshitome, A., Yamawaki, K., Amagai, M., Inoue, A., Oshima, T. & Makoto, K. R-spondin1 plays an essential role in ovarian development through positively regulating Wnt-4 signaling. *Hum. Mol. Genet.* **17**, 1278–1291 (2008).

48. Eggers, S. & Sinclair, A. Mammalian sex determination-insights from humans and mice. *Chromosom. Res.* **20**, 215–238 (2012).

49. Payer, B. & Lee, J. T. X Chromosome Dosage Compensation: How Mammals Keep the Balance. *Annu. Rev. Genet.* **42**, 733–772 (2008).

50. Kuroda, Y., Arai, N., Arita, M., Teranishi, M., Hori, T. & Harata, M. Absence of Z-chromosome inactivation for five genes in male chickens. *Chromospme Res.* **9**, 457–468 (2001).

51. McQueen, H. A., McBride, D., Miele, G., Bird, A. P. & Clinton, M. Dosage compensation in birds. *Curr. Biol.* **11**, 253–257 (2001).

52. Mank, J. E. & Ellegren, H. All dosage compensation is local: Gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. *Heredity (Edinb).* **102**, 312–320 (2009).

53. Itoh, Y., Melamed, E., Yang, X., Kampf, K., Wang, S., Yehya, N., Nas, A. Van, Replogle, K., Band, M. R., Clayton, D. F., Schadt, E. E., Lusis, A. J. & Arnold, A. P. Dosage compensation is less effective in birds than in mammals. *J. Biol.* **6**, 2 (2007).

54. Smith, C. A., McClive, P. J., Western, P. S., Reed, K. J. & Sinclair, A. H. Conservation of a sex-determining gene. *Nature* **402**, 601–602 (1999).

55. Teranishi, M., Shimada, Y., Hori, T., Nakabayashi, O., Kikuchi, T., Macloed, T., Pym, R., Sheldon, B., Solovei, I., Macgregor, H. & Mizuno, S. Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. *Chromosom. Res.* **9**, 147–165 (2001).

56. Roeszler, K. N., Itman, C., Sinclair, A. H. & Smith, C. A. The long non-coding RNA , MHM , plays a role in chicken embryonic development , including gonadogenesis. *Dev. Biol.* **366**, 317–326 (2012).

57. Ayers, K. L., Davidson, N. M., Demiyah, D., Roeszler, K. N., Grützner, F., Sinclair, A. H., Oshlack, A. & Smith, C. a. RNA sequencing reveals sexually dimorphic gene expression before gonadal differentiation in chicken and allows comprehensive annotation of the W-chromosome. *Genome Biol.* **14**, R26 (2013).

58. O'Neill, M., Binder, M., Smith, C., Andrews, J., Reed, K., Smith, M., Millar, C., Lambert, D. & Sinclair, A. ASW : a gene with conserved avian W-linkage and female specific expression in chick embryonic gonad. *Dev. Genes Evol.* 243–249 (2000).

59. Hori, T., Asakawa, S., Itoh, Y., Shimizu, N. & Mizuno, S. Wpkci , Encoding an Altered Form of PKCI , Is Conserved Widely on the Avian W Chromosome and Expressed in Early Female Embryos : Implication of Its Role in Female Sex Determination. *Mol. Biol. Cell* **11**, 3645–3660 (2000).

60. Smith, C. A., Roeszler, K. N. & Sinclair, A. H. Genetic evidence against a role for W-linked histidine triad nucleotide binding protein (HINTW) in avian sex determination. *Int. J. Dev. Biol.* **53**, 59–67 (2009).

61. Lin, M., Thome, M. H., Martin, I. C. A., Sheldon, B. L. & Jones, R. C. Development of the gonads in the triploid (ZZW and ZZZ) fowl, gallus domesticus, and comparison with normal diploid males (ZZ) and females (ZW). *Reprod. Fertil. Dev.* **7**, 1185–1197 (1995).

62. Graves, J. A. M. Sex and death in birds: A model of dosage compensation that predicts lethality of sex chromosome aneuploids. *Cytogenet. Genome Res.* **101**, 278–282 (2003).

63. Zhao, D., McBride, D., Nandi, S., McQueen, H. a, McGrew, M. J., Hocking, P. M., Lewis, P. D., Sang, H. M. & Clinton, M. Somatic sex identity is cell autonomous in the chicken. *Nature* **464**, 237–242 (2010).

64. Clinton, M., Zhao, D., Nandi, S. & McBride, D. Evidence for avian cell autonomous sex identity (CASI) and implications for the sex-determination process? *Chromosom. Res.* **20**, 177–190 (2012).

65. Vaillant, S., Dorizzi, M., Pieau, C. & Richard-Mercier, N. Sex Reversal and Aromatase in Chicken. *J. Exp. Zool.* **290**, 727–740 (2001).

66. Marchler-bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., *et al.* CDD : NCBI ' s conserved domain database. *Nucleic Acids Res.* **43**, D222–D226 (2015).

67. Reed, K. J. & Sinclair, A. H. FET-1: A novel W-linked, female specific gene up-regulated in the embryonic chicken ovary. *Mech. Dev.* **119**, S87–S90 (2002).

68. Liebert, M. A., Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. A Greedy Algorithm for Aligning DNA Sequences. **7**, 203–214 (2000).

69. Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chicken embryo. *J. Morphol.* **88**, 49–92 (1993).

70. Li, H., Xu, H., Zhao, C., Sulaiman, Y. & Wu, C. A PCR amplification method without DNA extraction. *Electrophoresis* **32**, 394–397 (2011).

71. Clinton, M., Haines, L., Belloir, B. & Mcbride, D. Sexing chick embryos : A rapid and simple protocol. *Br. Poult. Sci.* **42**, 134–138 (2001).

72. Kawakami, A., Kimura-Kawakami, M., Nomura, T. & Fujisawa, H. Distributions of PAX6 and PAX7 proteins suggest their involvement in both early and late phases of chick brain development. *Mech. Dev.* **66**, 119–130 (1997).

73. Walker, J. M. *Protein blotting by the capillary method*. *The Protein Protocols Handbook* (Humana Press Inc., 2009). doi:10.1385/1-59259-169-8:335

74. Korbie, D. J. & Mattick, J. S. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat. Protoc.* **3**, 1452–1456 (2008).

75. Bolisetty, M., Blomberg, J., Benachenhou, F., Sperber, G. & Beemon, K. Unexpected diversity and expression of avian endogenous retroviruses. *MBio* **3**, 1–10 (2012).

76. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).

77. Smith, C. A., Melissa, K. & Sincair, A. H. DMRT1 Is Upregulated in the Gonads During Female-to-Male Sex Reversal in ZW Chicken Embryos. *Biol. Reprod.* **68**, 560–570 (2002).

78. Hirst, C. E., Major, A. T., Ayers, K. L., Brown, R. J., Mariette, M., Sackton, T. B. & Smith, C. A. Sex Reversal and Comparative Data Undermine the W Chromosome and Support Z-linked DMRT1 as the Regulator of Gonadal Sex Differentiation in Birds. *Endocrinology* **158**, 2970–2987 (2017).

79. Smith, C. A., Andrews, J. E. & Sinclair, A. H. Gonadal Sex Differentiation in Chicken Embryos : Expression of Estrogen Receptor and Aromatase Genes. *J. Steroid Biochem. Mol. Biol.* **60**, 295–302 (1997).

80. Lambeth, L. S., Morris, K. R., Wise, T. G., Cummins, D. M., Neil, T. E. O., Cao, Y., Sinclair, A. H., Doran, T. J. & Smith, C. A. Transgenic Chickens Overexpressing Aromatase Have High Estrogen Levels but Maintain a Predominantly Male Phenotype. *Endocrinology,* **157**, 83–90 (2016).

81. Owens, I. P. F. & Short, R. V. Hormonal basis of sexual dimorphism in birds: implications for new theories of sexual selection. *Trends Ecol. Evol.* **10**, 44–47 (1995).

82. Dupressoir, A., Lavialle, C. & Heidmann, T. From ancestral infectious retroviruses to bona fide cellular genes: Role of the captured syncytins in placentation. *Placenta* **33**, 663–671 (2012).

83. Sha, M., Lee, X., Li, X. ping, Veldman, G. M., Finnerty, H., Racie, L., LaVallie, E., Tang, X. Y., Edouard, P., Howes, S., Keith, J. C. & McCoy, J. M. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* **403**, 785–789 (2000).

84. Dupressoir, A., Vernochet, C., Harper, F., Guegan, J., Dessen, P., Pierron, G. & Heidmann, T. A pair of co-opted retroviral envelope syncytin genes is required for formation of the two-layered murine placental syncytiotrophoblast. *Proc. Natl. Acad. Sci.* **108**, E1164–E1173 (2011).

85. Fatica, A. & Bozzoni, I. Long non-coding RNAs : new players in cell differentiation and development. *Nat. Rev. Genet.* **15**, 7 (2013).

86. Taft, R. J., Pheasant, M. & Mattick, J. S. The relationship between non-protein-coding DNA and eukaryotic complexity. *BioEssays* **29**, 288–299 (2007).

87. Ulitsky, I. & Bartel, D. P. lincRNAs: Genomics, Evolution, and Mechanisms. *Cell* **154**, 26–46 (2013).

88. Kung, J. T. Y., Colognori, D. & Lee, J. T. Long Noncoding RNAs: Past, Present, and Future. *Genetics* **193**, 651–669 (2013).

89. Taylor, D. H., Chu, E. T.-J., Spektor, R. & Soloway, P. D. Long non-coding RNA regulation of reproduction and development. *Mol. Reprod. Dev.* **82**, 932–956 (2015).

90. Augui, S., Nora, E. P. & Heard, E. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet.* **12**, 429–442 (2011).

91. Rastetter, R. H., Smith, C. A. & Wilhelm, D. The role of non-coding RNAs in male sex determination and differentiation. *Reproduction* **150**, R93–R107 (2015).

92. Deblois, G. & Giguère, V. Ligand-independent coactivation of ERalpha AF-1 by steroid receptor RNA activator (SRA) via MAPK activation. *J. Steroid Biochem. Mol. Biol.* **85**, 123–131 (2003).

93. Chooniedass-Kothari, S., Emberley, E., Hamedani, M. K., Troup, S., Wang, X., Czosnek, A., Hube, F., Mutawe, M., Watson, P. H. & Leygue, E. The steroid receptor RNA activator is the first functional RNA encoding a protein. *FEBS Lett. 566* **566**, 43–47 (2004).

94. Carré-Eusèbe, D., Coudouel, N. & Magre, S. OVEX1, a novel chicken endogenous retrovirus with sex-specific and left-right asymmetrical expression in gonads. *Retrovirology* **6**, (2009).

95. Gerdes, P., Richardson, S. R., Mager, D. L. & Faulkner, G. J. Transposable elements in the mammalian embryo: Pioneers surviving through stealth and service. *Genome Biol.* **17**, (2016).

96. Meyer, T. J., Rosenkrantz, J. L., Carbone, L. & Chavez, S. L. Endogenous Retroviruses: With Us and against Us. *Front. Chem.* **5**, 1–8 (2017).

97. Redelsperger, F., Raddi, N., Bacquin, A., Vernochet, C., Mariot, V., Gache, V., Blanchard-Gutton, N., Charrin, S., Tiret, L., Dumonceaux, J., Dupressoir, A. & Heidmann, T. Genetic Evidence That Captured Retroviral Envelope syncytins Contribute to Myoblast Fusion and Muscle Sexual Dimorphism in Mice. *PLoS Genet.* **12**, 1–18 (2016).

98. Li, Z., Ouyang, H., Zheng, M., Cai, B., Han, P., Abdalla, B. A., Nie, Q. & Zhang, X. Integrated analysis of long non-coding RNAs (LncRNAs) and mRNA expression profiles reveals the potential role of lncRNAs in skeletal muscle development of the chicken. *Front. Physiol.* **7**, 1–14 (2017).

99. Li, T., Wang, S., Wu, R., Zhou, X., Zhu, D. & Zhang, Y. Identification of long non-protein coding RNAs in chicken skeletal muscle using next generation sequencing. *Genomics* **99**, 292–298 (2012).

100. Zou, C., Li, J., Luo, W., Li, L., Hu, A., Fu, Y., Hou, Y. & Li, C. Transcriptome analysis reveals long intergenic non-coding RNAs involved in skeletal muscle growth and development in pig. *Sci. Rep.* **7**, 1–11 (2017).

101. Pelechano, V. & Steinmetz, L. M. Gene regulation by antisense transcription. *Nat. Rev. Genet.* **14**, 880–893 (2013).