8-29-2017

X Chromosome Dosage Compensation and Gene Expression in the Sheep

Kaleigh Flock
kaleigh.flock@uconn.edu

Recommended Citation

Flock, Kaleigh, "X Chromosome Dosage Compensation and Gene Expression in the Sheep" (2017). *Master's Theses*. 1144.
https://opencommons.uconn.edu/gs_theses/1144

X Chromosome Dosage Compensation and Gene Expression in the Sheep

Kaleigh Flock

B.S., University of Connecticut, 2014

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Masters of Science

at the

University of Connecticut

2017

APPROVAL PAGE

Masters of Science Thesis

X Chromosome Dosage Compensation and Gene Expression in the Sheep

Presented by

Kaleigh Flock, B.S.

Major Advisor____
Dr. Xiuchun (Cindy) Tian

Associate Advisor____
Dr. David Magee

Associate Advisor____
Dr. Sarah A. Reed

Associate Advisor____
Dr. John Malone

University of Connecticut

2017

Dedication

This thesis is dedicated to my major advisor Dr. Xiuchun (Cindy) Tian, my lab mates Mingyuan Zhang and Ellie Duan, and my mother and father. This thesis would not be possible without your hard work, unwavering support, and guidance. Dr. Tian, I am so thankful for the opportunity to pursue a Master's degree in your lab. The knowledge and technical skills that I have gained are invaluable and have opened many doors in my career as a scientist and future veterinarian. I strive to always be inquisitive and take full advantage of every learning opportunity. Mingyuan, it was a pleasure working with you and learning from you. The hours that we spent in the lab together performing experiments have culminated into great projects that will increase the knowledge base in the scientific community. Ellie, thank you for your hard work and dedication to the data analysis presented in this thesis. Data analysis by computer programming is very complex and I am thankful for your help and support. Mom and dad, thank you for giving me the opportunity and skills to succeed in my academic journey and in all my future endeavors. Having your support means the world to me.

ACKNOWLEDGEMENTS

My sincerest thanks and appreciation goes to my major advisor Dr. Xiuchun (Cindy) Tian. I can’t thank you enough for giving me the opportunity to join your lab, it has been such a wonderful and memorable experience. You are an inspiring person and scientist. Thank you for guiding me to become a skilled scientist, out-of-the-box thinker, independent worker, and for always supporting and believing in me. Lastly, thank you for pairing Mingyuan and I together for the sheep projects. It was the perfect match and I gained a lifelong friend.

I would like to thank Mingyuan Zhang for being the best experiment partner that anyone could ever ask for. Thank you for sharing your wealth of laboratory skills and knowledge with me and for encouraging me to grow as a scientist and person. It was a pleasure planning and performing experiments with you and effectively rising above the challenges of research together, learning and improving all the while.

I would like to thank my committee members Dr. David Magee, Dr. Sarah Reed, and Dr. John Malone for your support, advice, and insight. Thank you for your valuable feedback and comments that made this thesis possible.

I would like to thank all of the collaborators who made this project possible. Dr. Kristen Govoni, Dr. Sarah Reed, and Dr. Steven Zinn thank you for conceiving the experimental design, providing all of the samples used in the experiment, and for your help and support throughout the study. Thank you Dr. Ion Mandoiu and Sahar Al Seesi from computer science and engineering for your expert advice and creativity in engineering the optimal data analysis workflow to analyze our RNA seq data and solve our biological questions. Thank you, Dr. Jiang from Guangxi University for your support.

I would like to thank my lab mates: Mingyuan Zhang, Ellie Duan, Dr. Zongliang Jiang, Linkai Zhu, Dr. Nan Li, Dr. Limin Wang, Dr. Huan Yang, Dr. Junhe Yu, and previous lab mates Dr. Rashid Ali and Rose Wang. Thank you for your kindness in welcoming me into the lab, your advice and support, and your friendship. Ellie, thank you for all of your hard work and dedication to the data analysis of the two projects and for your help and support. Zongliang, thank you for training me, for your assistance with RNA seq questions, and your valuable advice in optimizing our experiments. Also, thank you Ling Wang and Delun Huang from Dr. Tang's lab, for your help and advice. It was a pleasure to work with such outstanding scientists and to learn about your projects.

I would like to thank Amanda Jones, Dr. Maria Hoffman, Sambhu Pillai, Joseline Raja, and Katelyn McFadden for their patience, help and support with the samples, experimental design, and answering any and all questions that we had.

I would like to thank Dr. Bo Reese and Lu Li at UCONN CGI for your expert training in library preparation, advice, and assistance with questions and sequencing. I would like to thank the Animal science department graduate students and faculty for your kindness and support.

I would like to thank my family and friends. Genevieve, thanks for encouraging me to pursue a master's degree, helping me to feel comfortable and connected in the department, and for your advice and support. Ashley, Jackie, and Shannon thank you for always being there for me. Mom and Dad thank you for always encouraging me to follow my dreams, believing in me, and supporting me through everything. I love you all very much and would not be able to accomplish my goals without your help. Thank you Luis Organista for your support and understanding. It means so much to me that we are able to help each other achieve our goals and follow our dreams.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 3

RESULTS

LIST OF FIGURES

INTRODUCTION & CHAPTER 1

CHAPTER 2

CHAPTER 3

RESULTS

LIST OF COMMON ABBREVIATIONS

ANOVA: analysis of variance
cDNA: complementary deoxyribonucleic acid
CON: control
CNVs: copy number variations
DNA: deoxyribonucleic acid
FPKM: fragments per kilobase of exon per million
GEO: gene expression Omnibus
GO: gene ontology
IUGR: intrauterine growth restriction
MAOA: monamine oxidase type A
mRNA: messenger ribonucleic acid
OVER: overfed
PAB: pseudoautosomal boundary
PAR(S): pseudoautosomal region(s)
RES: restricted
RIN: RNA integrity number
RNA: ribonucleic acid
RNA-seq: RNA sequencing
RPKM: reads per kilobase per million
RT-PCR: Real Time Polymerase Chain Reaction
RXE: relative X chromosome expression
SCNT: somatic cell nuclear transfer
SNP: single nucleotide polymorphism
TPM: transcripts per million
X:A: X to Autosome expression ratio
XCI: X chromosome inactivation
Xist: X-inactive specific transcript

INTRODUCTION

Proper dosage of gene and chromosome copy is essential in normal development. Gene dosage refers to the amount of gene product. Copy number refers to the number of copies of a gene present in a genome. Copy number variations (CNVs) are segments of DNA that are 1 kilobase or larger in which insertion or deletion events have occurred[1]. These DNA segments exist in a variable copy number compared to the reference genome. CNVs can influence genes or gene regions, such as those in livestock production traits. Dosage compensation is the balancing of expression between male and female sex chromosomes and between the sex chromosomes and the autosomes[2]. Dosage compensation mechanisms exist in eutherian mammals[3], marsupials[4], monotreme mammals[5], birds[6], and the non-mammalian invertebrates *C. elegans*[7] and *D. melanogaster*[8]. In eutherian mammals, dosage between male and female sex chromosomes and between the sex chromosomes and the autosomes must be balanced. The evolution of the sex-determining chromosomes X and Y led to a single functional X being present in males, while two X's exist in females. Without a mechanism to compensate, X-linked gene expression would be unequal between the sexes due to 'X aneuploidy' in males[9] (Figure 1).

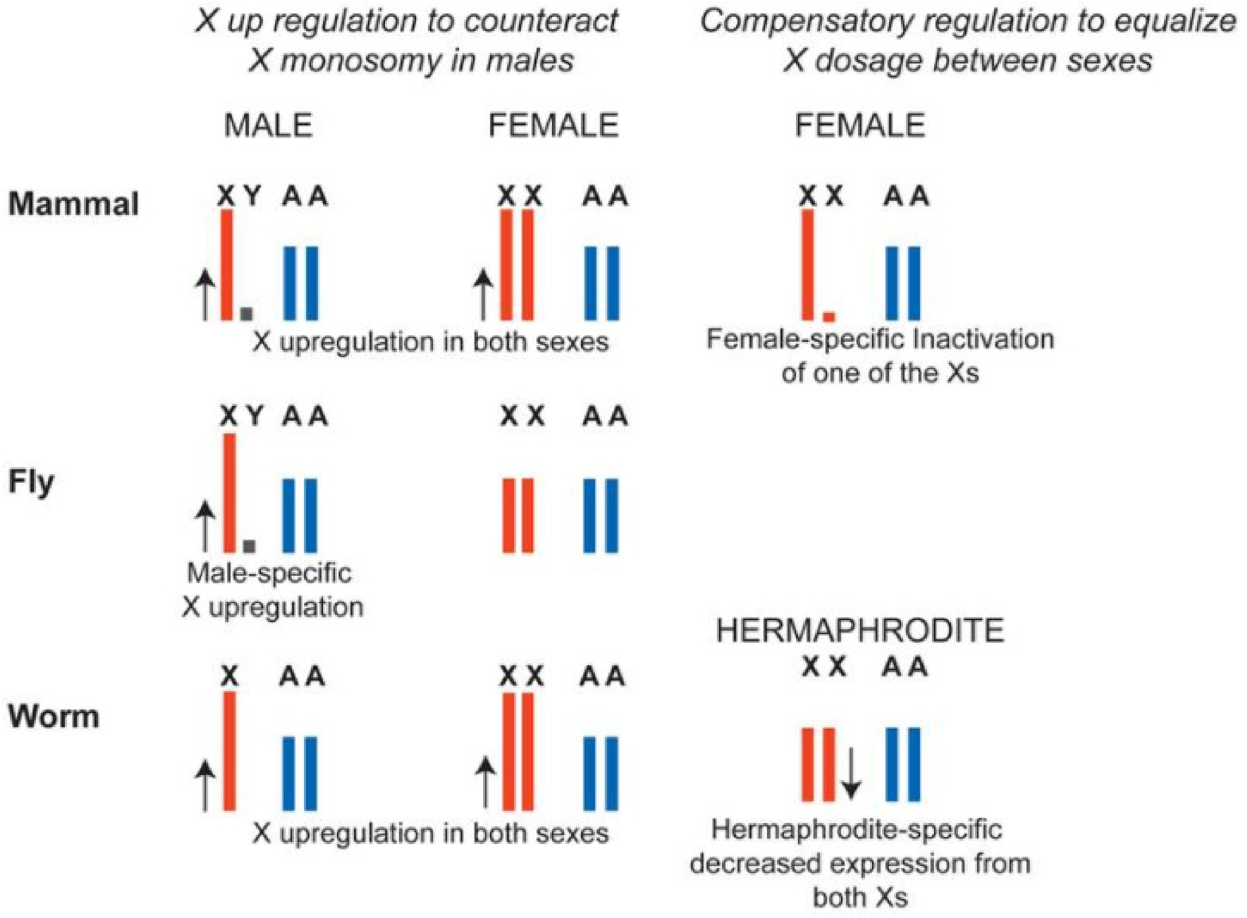


Figure 1. Ohno's hypothesis of dosage compensation in mammals, flies (*D. melanogaster*), and worms (*C. elegans*). In flies, X chromosome transcription is increased in males (XY) to compensate for X monosomy compared to the autosomes and to balance expression of the male X chromosome with the female X chromosomes. In worms, the X chromosome is upregulated in males (XO) to compensate for X monosomy. Upregulation of the X chromosome in female (XX) worms is counteracted by a female-specific dosage compensation mechanism, effectively balancing the sex chromosomes in males and females and with respect to the autosomes. Worms can be XX hermaphrodites and dosage compensation downregulates both X's by a factor of two to balance with the dosage of XO males. In mammals, the X chromosome is upregulated in both males (XY) and females (XX) followed by inactivation of one of the X chromosomes in females to balance the dosage of the sex chromosomes in males and females and with respect to the autosomes. From Ercan et al., 2015.

Aneuploidies are chromosomal deviations from the normal diploid dosage of an animal that reduce fitness and are usually lethal[10,11]. Human chromosome 21, the smallest chromosome containing the fewest genes, deviates from diploid due to non-disjunction in meiosis I where homologous chromosomes do not separate resulting in trisomy 21[12,13]. Trisomy 21 or Down's Syndrome has variable and complex clinical presentation including distinct facial dysmorphology, a brain that is smaller and hypocellular, and cognitive impairment[12]. Conversely, mammalian males have monosomy of the X chromosome when compared to mammalian females and are phenotypically normal. Aneuploidies of the X chromosome include Turner's syndrome (XO female), and Klinefelter's syndrome (XXY male). Characteristic features of Turner's syndrome include short stature and gonadal development failure, while Klinefelter males have taller stature, lower muscle tone, and decreased gonadal development[14]. X chromosome dosage compensation in mammals explains X chromosome regulation in the sex chromosome disorders Turner's syndrome and Klinefelter's syndrome and how males survive 'X aneuploidy'.

Ohno hypothesized that X-linked gene expression is doubled in both males and females, successfully balancing with the autosome expression in males[15]. Dosage compensation is necessary because in mammals, the X chromosome is large and gene rich and the Y chromosome is small and gene poor. The sex chromosome system in mammals denotes female mammals as XX and males as hemizygous with genotype XY. Due to dosage compensation, X-linked genes in males have a twofold upregulation. In females, the upregulation results in an overexpression from both X chromosomes and lead to the downregulation of an X chromosome to restore balance[16] (Figure 1). Dosage compensation is well studied in humans and mice, but little research has been done in other domestic species.

Another mechanism that helps achieve dosage compensation in mammalian females is X-chromosome inactivation (XCI) which randomly and globally inactivates one of the X-chromosomes[17]. X chromosome inactivation is unique to mammals[18], including monotremes[19], marsupials[20], and eutherians[21], but this thesis will focus mainly on dosage compensation and X chromosome inactivation in eutherian mammals.

Dosage compensation is species, tissue, and developmental stage specific, these differences warrant the need to study each species independently[22]. There is current debate in the scientific community over the confirmation of Ohno's hypothesis of dosage compensation in eutherian mammals. The debate is fueled by the difference in RNA sequencing and microarray experiments and differences in data analysis. Studies that reject Ohno's hypothesis of dosage compensation have included actively expressed, weakly expressed, and silent genes in their analysis. Xiong et al analyzed publicly available RNA-sequencing datasets and reported an X:AA ratio of ~0.5 and rejected Ohno's hypothesis[23]. Two additional studies have also rejected Ohno's hypothesis[24,25]. Later studies that have re-analyzed this dataset and other generated datasets, have reported an X:AA ratio around 1.0 and support Ohno's hypothesis[3,9,26,27]. These studies have only included actively expressed genes in their analysis of dosage compensation.

Sheep are a valuable model to study dosage compensation because little is known about dosage compensation in sheep and many production traits are linked to X-linked genes in sheep[28]. These X-linked genes are important in reproduction, linked to the immune system and disease, involved in biosynthetic pathways, and have human orthologs[28]. Genes on the X chromosome are evolutionarily conserved across mammalian species. X-linked genes that are X-specific have single expression due to X chromosome inactivation in females and XY males. Recombination between the X chromosome and autosome has been selected against because it would disrupt dosage compensation[29,30].

Maternal nutrition studies are valuable to sheep producers and sheep have been historically used as a model for human pregnancy[31,32]. Previous studies in *Ovis aries* have shown that maternal nutrition can produce gene expression changes in fetal tissues[33]. During the change in seasons, forage quality and quantity also changes and results in over and undernutrition of pregnant ewes. When forage quality and quantity are low in the fall and winter, intrauterine growth restriction (IUGR) occurs and poor growth of the developing fetus is a consequence. When ewes are fed ad-libitum feed or graze in high quality pasture, maternal overnutrition is common and affects the developing fetus[34]. Maternal nutrition can induce permanent changes in structure, physiology, and metabolism of offspring[35].

Through a collaboration with the labs of Dr. Govoni, Dr. Reed, and Dr. Zinn, we obtained ovine fetal day 135 tissues of brain, kidney, and lung and characterized dosage compensation in the sheep transcriptome by RNA-sequencing. Day 135 of gestation in sheep corresponds to late gestation, the maximal fetal growth period[33]. The effect of maternal overnutrition and undernutrition on dosage compensation was also investigated. Investigating the normal pattern of dosage compensation is important in understanding abnormalities that occur naturally and through biotechnology.

CHAPTER 1

LITERATURE REVIEW

1.1 DOSAGE COMPENSATION

Dosage compensation is present in non-mammalian and mammalian species and is achieved through different mechanisms such as increased X transcription in only the male or increased X transcription in males and females followed by a female-specific dosage compensation mechanism (Figure 1)[9]. In addition to species differences, dosage compensation is known to be both developmental stage specific and tissue specific[26,36]. Different sex chromosome systems exist in different mammalian and non-mammalian species. During the evolution of sex chromosomes X and Y in eutherian mammals and marsupials, divergence led to monosomy of the X chromosome in males (XY). During the evolution of sex chromosomes Z and W in birds and reptiles, divergence led to monosomy of the Z chromosome in females (ZW)[16]. The Y and W chromosomes became sex-limited through loss of gene activity through evolution. Both systems of sex determination resulted in an imbalance in gene dosage of X-linked or Z-linked genes in females and males respectively[37]. Susumu Ohno hypothesized that to compensate for monosomy of the sex chromosomes X or Z, upregulation of X or Z in the heterogametic sex would be necessary to return gene expression of the sex chromosomes to normal diploid level[15]. It was thought that sex chromosome evolution led to complete dosage compensation, as previously observed in *C. elegans*, *D. melanogaster*, and *M. musculus* until two independent studies in birds revealed incomplete dosage compensation of the Z chromosome[38,39]. Further investigation of Ohno's hypothesis has revealed incomplete dosage compensation in other species suggesting that a whole-chromosome regulation method is not employed in in all species[16]. In species with incomplete dosage compensation, there is direct compensation of a subset of genes known as dosage-sensitive genes, while loci that do not experience a dose effect are indirectly compensated[16].

1.1.2 Gene dosage

In diploid organisms, chromosome and gene dosage are closely regulated. Deviation from diploid on the gene level can have detrimental consequences and on the chromosome level is typically lethal. Aneuploidy is described as a deviation from the normal copy number of an individual chromosome[40]. There are two types of aneuploidy in multicellular organisms, conditional aneuploidy and somatic aneuploidy. Conditional aneuploidy has a presence in all cells and is marked by adverse effects, while somatic aneuploidy is more selective[40]. Turner's syndrome and Klinefelter's syndrome are sex chromosome disorders with aneuploidy of the X chromosome[41]. Mammalian XY males are monoallelic for most X-linked genes, making them a functional 'X aneuploidy'. Autosomal monosomies of chromosomes equal in size to X have lethal consequences[42]. XY males are able to avoid the deleterious effects of X monosomy through dosage compensation.

The delicate balance of gene dosage is explained by the evolution of the mammalian sex chromosomes. The mammalian XY pair evolved from a pair of autosomes as explained by H.J. Muller's proposed model (1914). Proto-X and proto-Y arose when one of the autosomes gained a sex determining locus. Proto-Y then continued to accumulate alleles that were advantageous to males, and X and Y recombination was lost[43] (Figure 2). The nonrecombining region gained mutations, deletions, and insertions of repetitive elements. Genes that did not have a sex-specific advantage became inactive and were lost from the Y chromosome[43]. As genes were gradually lost from the Y chromosome, making it haploinsufficient, the genes on the X chromosome had an increase in transcription[2].While present day X and Y differ in gene content, they share a region of sequence homology called the pseudoautosomal region (Figure 3). The pseudoautosomal region of the X and Y chromosome maintains 98-99% similarity in sequence and gene content, recombines frequently, has high GC content and high rate of mutation[44,45]. This region is conserved among mammalian species, but displays variation in gene content and size based on the species[46].

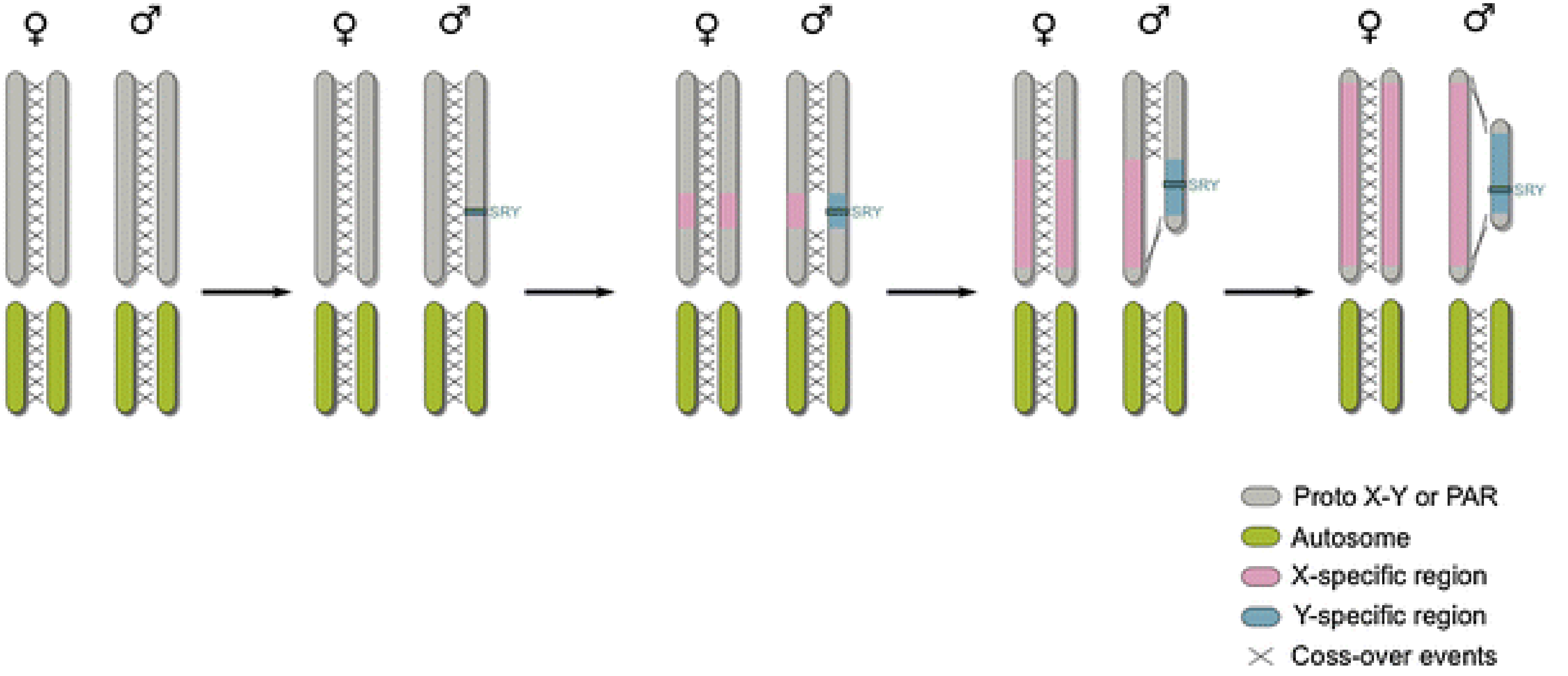


Figure 2. Sex chromosomes in mammals began as a pair of autosomes, proto-X and proto-Y, shown in gray with recombination shown with x's. An autosome pair is shown in green. The male-determining gene Sry initiated sex chromosome evolution in mammals. This blocked recombination between proto-X and proto-Y at the region of Sry and nearby genes, leading to the creation of X-specific (pink) and Y-specific (blue) regions. During evolution, recombination was further suppressed, the X-specific and Y-specific regions grew larger, and X and Y diverged. Pseudoautosomal regions (PARS) are the sections of the X and Y chromosome shown in gray that are still able to recombine. The Y chromosome is unable to recombine with the X chromosome in the male-specific region and has lost genes, becoming progressively smaller. From Pessia et al., 2014.

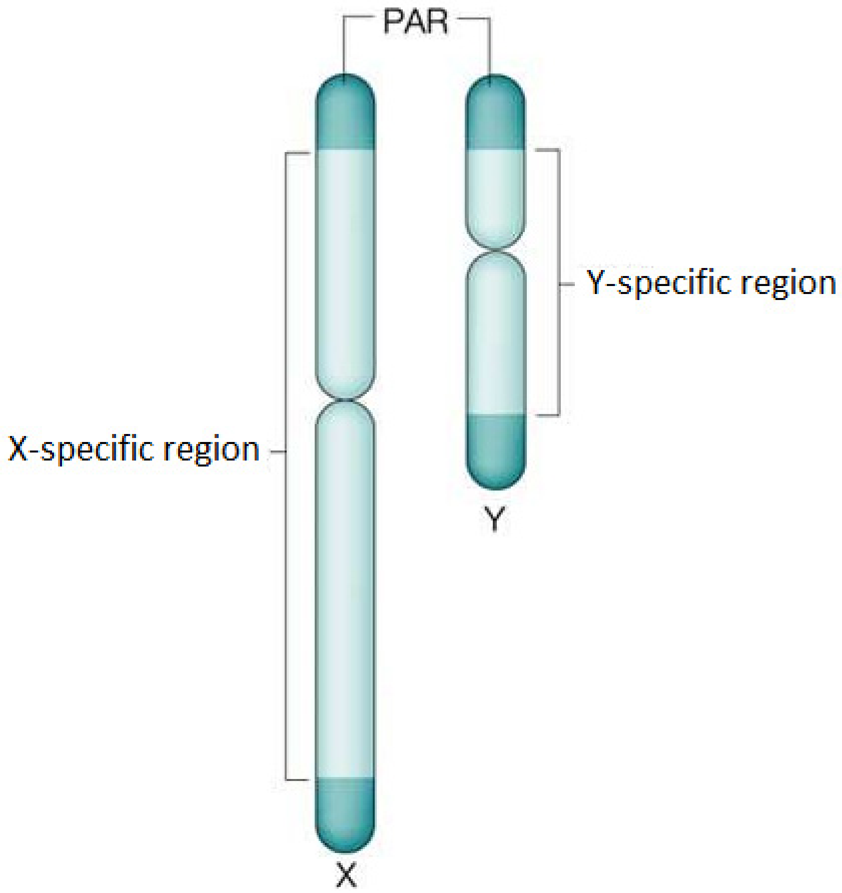


Figure 3. The human X and Y chromosomes have pseudoautosomal regions (PARs) which contain the same genes and undergo recombination. The X chromosome has an X-specific region and the Y chromosome has a Y-specific region which cannot recombine. X and Y chromosome evolution occurred in the same manner for mammalian species, but the size of the pseudoautosomal region(s) is species dependent. Adapted from Arnold et al., 2004.

1.1.3 Non-mammalian dosage compensation

The molecular mechanisms used to accomplish dosage compensation, although unique for the non-mammalian invertebrate species *D. melanogaster* (fly), *C. elegans* (worm), and mammalian *M. musculus* (mouse) have many similar elements including dosage compensation complex recruitment, a cis-spreading pattern, and transcription and chromatin structure regulation[47].*D. melanogaster* are diploid with one X chromosome per set of autosomes in males (X:AA) and two X chromosomes per autosome set in females (XX:AA)[36]. In fly somatic cells, male specific lethal complexes are deployed that double the transcription from the male X chromosome[7] (Figure 1).

In *C. elegans,* embryos that are XX are female and develop into hermaphrodites that are able to self-fertilize internally, but preferentially use XO male sperm when available for fertilization[7]. The X to Autosome (X:A) ratio determines sex and X-linked gene expression level. Males are XO with an X:A ratio of 0.5 while hermaphrodites are XX with an X:A ratio of 1.0[7]. Dosage compensation in *C. elegans* involves the decrease to one half the gene expression from both X chromosomes in hermaphrodites[48]. While Drosophila dosage compensation has been well studied, evidence for X repression in *C. elegans* is more recent. Dosage compensation is complete in the fly and the worm[49,8] . Lastly, in vertebrate birds, fish, and reptiles, dosage compensation appears to be partial and gene-specific[50].

1.1.4 Mammalian dosage compensation

The mechanism of dosage compensation equalizes the X chromosome dose in males and females and also minimizes the damaging effects of X-polysomy[15]. Ohno hypothesized that the mechanism of dosage compensation in mammals evolved by first doubling the expression of the X chromosome in both males and females. This solved the dosage imbalance problem of X-linked genes in males, and then inactivation of a single X chromosome in every cell of females by X chromosome inactivation (XCI) balanced gene dosage in both of the sexes[15]. Both X chromosome upregulation and X chromosome inactivation are necessary parts of the dosage compensation mechanism in mammals[3].

The status of dosage compensation in mammals has been analyzed with microarray and RNA-seq data by computing the mean expression of all X-linked genes to the mean expression of all autosomal genes. An X:A ratio of 1.0 indicates doubling of transcription of genes on the X chromosome, while an X:A ratio of 0.5 indicates that transcription is not doubled[26]. While global X upregulation has been seen in marsupials, global X upregulation is absent in placental mammals[4]. Partial to full upregulation of dosage sensitive X-linked genes is currently observed in eutherian mammals[4,3,51].

1.1.5 Developmental stage specificity

Dosage compensation in mammalian and non-mammalian species is essential to proper development and variation in the degree of dosage compensation can be studied throughout different developmental stages. Studies in *D. melanogaster* and *C. elegans* have discovered that improper dosage compensation is lethal, resulting in the death of *C. elegans* in embryogenesis or early larval stage[52,53,54] and male specific lethal mutations in Drosophila[55]. Also, mouse embryos die around day 10 with a lack of XCI[56].

In eutherian mammals, the developmental stage when XCI is achieved is unclear and may be largely varied among different species[57]. Okamoto et al. showed that mammalian species have diversity in the time that XCI is activated in early embryogenesis and in its regulation[21]. Random X chromosome inactivation transcriptionally inactivates one of the two X chromosomes in each cell at random[17]. In imprinted XCI, the paternally inherited X chromosome is preferentially silenced and the maternally inherited X remains active[58] (Figure 4).

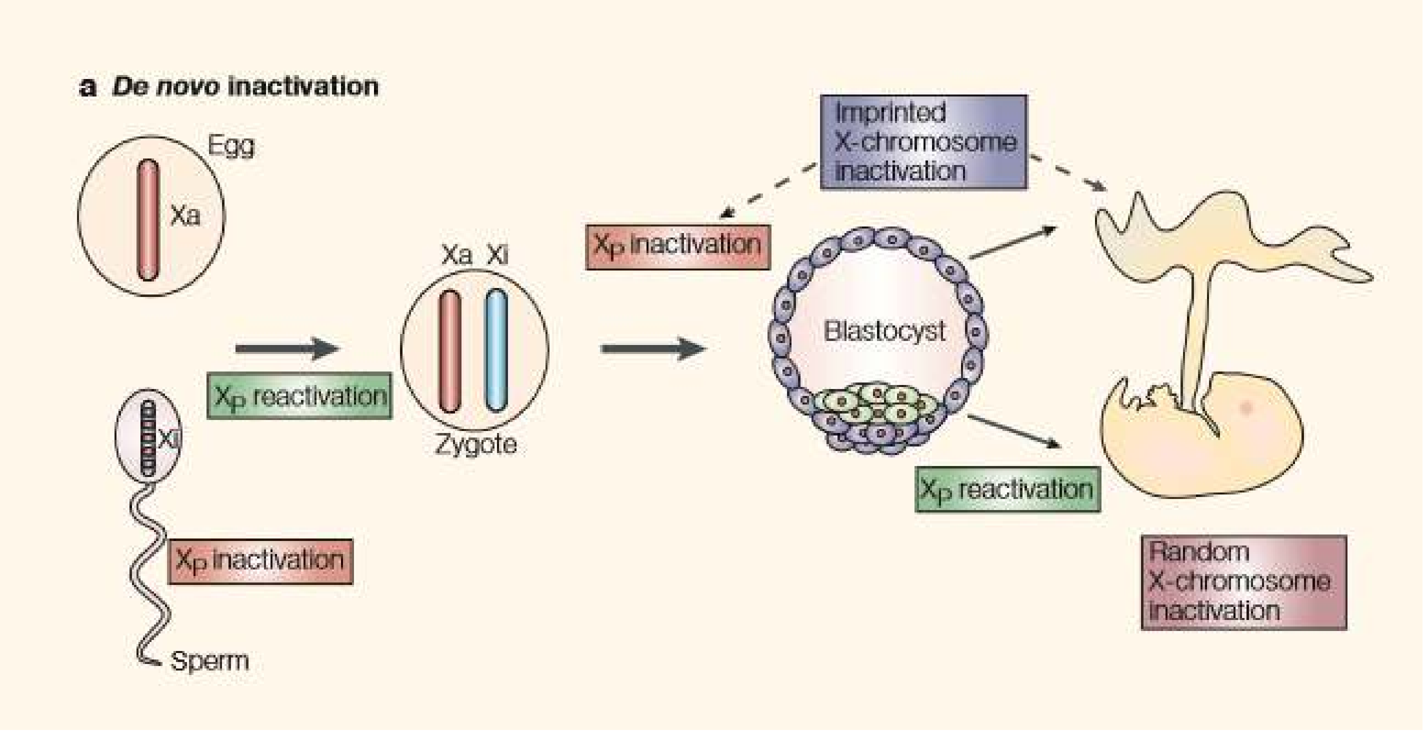


Figure 4. The de novo inactivation model of X chromosome inactivation in mammals is displayed. The egg has an active X chromosome and the sperm has an inactive X chromosome that is reactivated after meiosis. In the zygote, there are two active X chromosomes, and the paternal X (XP) chromosome is re-inactivated beginning at the 4-8 cell stage in mice. In the blastocyst, the trophectoderm (extra-embryonic cells in blue) maintains inactivation of XP and causes imprinted X-chromosome inactivation in the placenta. In the blastocyst, the inner cell mass (green cells) undergo reactivation of XP and then random X-chromosome inactivation. The fetus has random X-chromosome inactivation. From Huynh et al., 2005.

Random XCI initiation occurs at or after the blastocyst stage in humans, in cells of the morula and early blastocyst stage in rabbits, and in the late blastocyst stage in mice[59]. In mice, imprinted XCI initiation occurs first at the 2-4 cell stage. Another developmental stage specific event involves genes termed XCI escaping genes which are located outside the pseudoautosomal regions of the X chromosome and have bialellic expression[58]. In adult mouse tissues, genes escaping XCI are first inactivated in the mouse embryo and become reactivated during development[60]. The particular developmental stage in which genes escape X chromosome inactivation in humans is unknown[61]. The differences based on developmental stage highlight the advantage to studying early stages in embryonic development in various species to uncover species specific time points.

1.1.6 Tissue-specificity

Despite the difference in the mechanisms of dosage compensation in *Drosophila* and mammals, tissue-specific variation in dosage compensation has been observed in both species and may contribute to sex-biased gene expression[62]. In mammals, the expression of X-linked genes differs by tissue type. There is a high expression of X-linked genes in the brain, significantly higher compared to other somatic tissues[9]. The X chromosomes of mice and humans are enriched for genes related to brain function[26,63]. In addition, in the mouse brain, there is preferential expression of genes from the maternal X chromosome, indicating a bias in XCI[64]. Genes with tissue-restricted expression, such as those in the testis, ovary, and brain, accumulate on the sex chromosomes[9]. In pre-meiotic and post-meiotic stages, male specific X-linked genes are largely expressed in the testes[65,66]. X-linked genes that are conserved on chicken orthologs are characterized as the oldest X-linked genes and show high expression in mouse and human ovaries[67,68] Variation exists in the level of completion of XCI in different adult tissues[69]. X chromosome inactivation in eutherian mammals is random in somatic cells and the extra-embryonic structures can either follow random or imprinted inactivation.

1.1.7 Ovine dosage compensation

As there is no direct research on dosage compensation in the sheep, current research is aimed at the comparison of sheep X chromosome to that of the cow and human, pseudoautosomal regions, and X-linked genes in sheep with implications in artificial selection. In comparing the present day bovine X chromosome, ovine X chromosome, and human X chromosome, increased locus order differences were found between the ovine and bovine X chromosome than between the ovine and human X chromosome[70].

Studying pseudoautosomal regions can provide insight into sex chromosome evolution. Mammalian X and Y chromosomes share a region of sequence homology called the pseudoautosomal region (PAR) where recombination occurs during prophase of male meiosis[71]. This region contains 98-99% sequence similarity between the sex chromosomes, the same gene content, high GC content, and high recombination frequency[71]. The ruminant pseudoautosomal region is 5-9 Mb in size and shares the same genes as human pseudoautosomal region 1 (PAR1), with the exception of the gene PLCXD which is X-specific in ruminants[72]. The ruminant PAR begins with the gene GTPB6P and ends at the gene GPR143, marking the pseudoautosomal boundary (PAB)[72,73](Figure 5). This indicates that the boundary was established before Bovinae and Caprinae diverged ~18 million years ago[74]. Studying dosage compensation in the sheep can bridge the knowledge gap and uncover the mechanism of dosage compensation. The ovine X chromosome has been evaluated for artificial selection signatures that are useful in improving the desired phenotypic traits and guiding animal breeding[28]. X-linked genes in sheep are linked to reproductive function such as ovulation rate [75,76]. Because of sex-specific dosage compensation in mammals, X chromosome selection pressure is increased compared to the autosomes, highlighting more direct selection on the X chromosome[26,43].

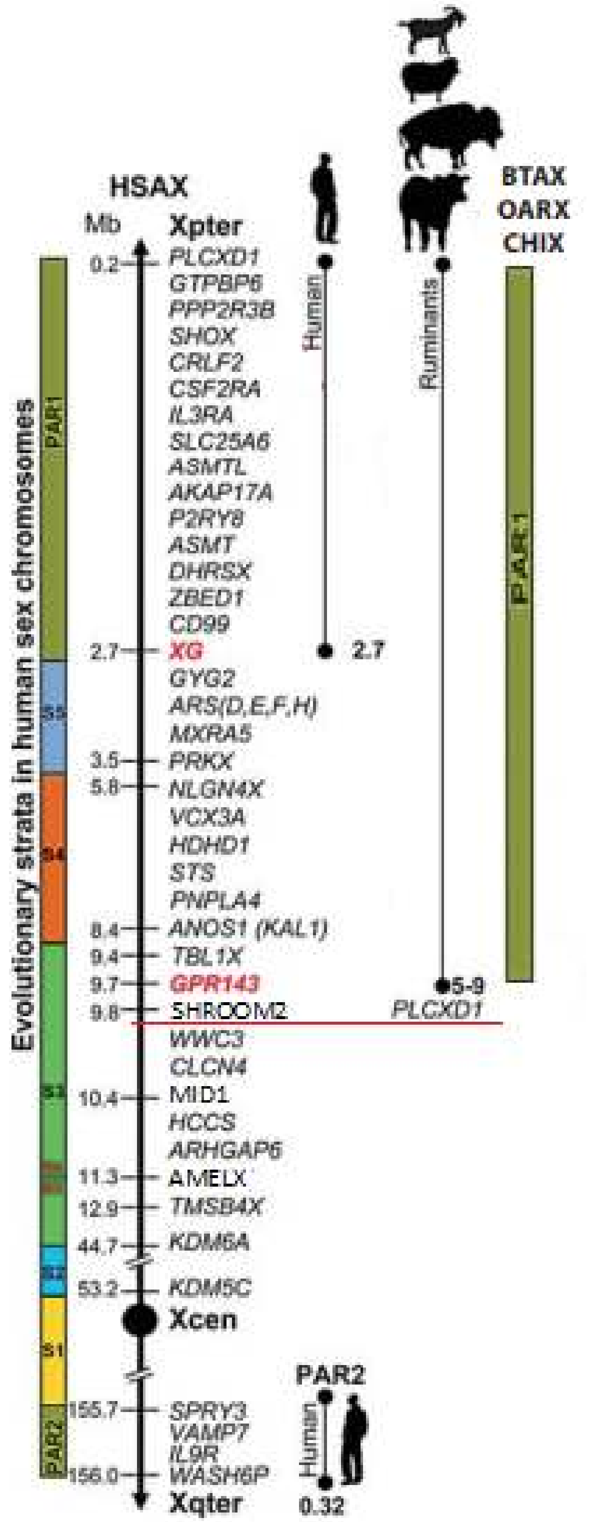


Figure 5. The Human X chromosome (HSAX) shows the organization of the human pseudoautosomal regions (PAR1 and PAR2) and X-specific regions evolutionary strata (S1-S5). Pseudoautosomal boundaries are shown in pink text. Human PAR1 is 2.7 Mb starting at the gene PLCXD1 and ending at the gene XG and human PAR2 is 0.32 Mb. The ruminant PAR starts at the gene GTPB6P and ends at the gene GPR143. The ruminant PAR is 5-9 Mb. The gene PLCXD1 is X-specific in ruminants and is not located in the pseudoautosomal region. Its location is marked by the red line. Adapted from Raudsepp et al 2015.

1.2 USING RNA SEQ TO CONFIRM MAMMALIAN DOSAGE COMPENSATION

A throughput method that assays the expression of a large number of X-linked and autosomal genes is required to study dosage compensation because expressed genes from the entire X chromosome and all autosomes are compared [77]. Throughput methods include microarray and RNA sequencing. Microarray assays a large number of genes with representative probes, but limitations exist in cataloguing and quantifying the differences in expression level and diversity of RNA molecules[78]. RNA sequencing is now the predominant method for studying transcriptomics and dosage compensation. The advancement of RNA sequencing (RNA-seq) technology introduced a high-throughput method where all gene transcripts in the transcriptome can be analyzed. RNA sequencing advantages include the ability to validate microarray results, study diverse species outside of the standard model organisms, and perform strand specific cDNA sequencing to study both sense and antisense transcripts[79]. Next generation sequencing technology continues to evolve, becoming specialized for a multitude of applications and more accessible to researchers.

1.2.1 Microarray and RNA-seq studies

Since the early 2000's, microarray data has been analyzed to determine the extent of dosage compensation in eutherian mammals. Since 2010, RNA sequencing data has been analyzed and different conclusions have been reached on the status of dosage compensation in mammalian tissues due to the data analysis parameters and software used. Throughput data is needed to study dosage compensation because the mean expression of all the X-linked genes is compared with the mean expression of all the autosomal genes. There is current debate on the status of dosage compensation in mammalian tissues, with studies supporting Ohno's hypothesis of dosage compensation and studies rejecting Ohno's hypothesis of dosage compensation.

If X upregulation in males and females plus X inactivation in females are unsuccessful in dosage compensation, the X:AA mean expression ratio should be 0.5 in both sexes[3]. All the following studies reject Ohno's hypothesis of dosage compensation upon RNA-seq data analysis. Xiong et al re-analyzed RNA-seq datasets for human tissues[80-82], mouse tissues[83], and *C. elegans* developmental stages[84] and rejected Ohno's hypothesis due to calculated X:AA median expression ratios of 0.3 in mice, 0.5 in humans, and 1.0 in *C. elegans* embryos declining to 0.4 in *C. elegans* adults[23]. This group supports their conclusion that the X:AA ratio is ~ 0.5 with the importance of comparing actively expressed , weakly expressed, and silenced genes between X and autosomes[24]. Lin et al re-analyzed RNA-seq datasets of chicken, mouse, and human organs[85] by comparing 1:1 orthologs in human to chicken and mouse to chicken and concluded that after normalizing the AA:AA ratio to 1.0, the X:XX ratio is ~0.5[51]. These analyses found that comparing all the genes on the X chromosome and autosomes reveals a lack of dosage compensation.

Microarray data analysis has supported dosage compensation in mammals. If dosage is effectively compensated, the X:AA mean expression ratio should be 1.0 in both females and males[3]. All of the following studies support Ohno's hypothesis of dosage compensation upon RNA-seq data analysis. Ngyuen et al analyzed human and mouse microarray datasets and their own human tissue generated microarray data and concluded that dosage compensation is achieved in humans with an X:AA ratio of 0.94 and in mice with an X:AA ratio of 1.01 and these ratios did not show significant difference based on the tissue type[26].

The following studies re-analyzed the data in the publication by Xiong et al and declared that dosage is compensated in humans, mice, and *C. elegans*. Pessia et al re-analyzed the human tissues RNA-seq dataset and concluded that for dosage sensitive genes that code large protein complexes, the X:AA ratio is equal to 0.9[3]. Kharchenko et al re-analyzed the human and mouse RNA seq datasets and an additional mouse RNA-seq dataset by Gregg et al and found that the average X:AA ratio using reads per kilobase per million ($RPKM \geq 1$) is close to 1.0 in mouse and human tissues, specifically 0.93 ± 0.17 in all human tissues[86]. Deng et al analyzed Xiong et al's data and newly released RNA-seq datasets, removing skewed X chromosome gene content by using fragments per kilobase of exon per million greater than zero ($FPKM>0$) and found that *C. elegans* achieve dosage compensation with an X:AA ratio of 0.99 in adult animals lacking a germline.

In humans and mice, the X:AA ratios with the removal of genes that are weakly expressed and silenced reveal that the majority of genes that are expressed from the active X are upregulated and achieve comparable expression with the autosomes[9]. Lin et al proposed that dosage compensation can be effectively studied by measuring only genes that are actively expressed. Actively expressed genes in mouse embryonic stem cells, blastocysts and adult lymphocytes, are upregulated in both males and females[87]. Slightly more RNA-seq data analysis studies exist in favor of Ohno's hypothesis, while a final consensus has not been reached, it is clear that conclusions are heavily dependent on the analytical approach.

1.2.2 Effect of analytical approach on dosage compensation

RNA-seq data analysis is complex due to the vast number of software options available for processing data and the strong effect that the analytical approach can have on the final experimental results and conclusions. The scientific debate over the status of dosage compensation in mammals has resulted in extensive re-analysis of the same RNA-seq datasets, leading to different final conclusions. The factors that affect RNA-seq data analysis include trimming parameters, mapping parameters, reference genome annotation, library preparation, and depth of sequencing coverage.

Raw RNA-seq reads are trimmed using parameters for quality score and length. Since genes expressed at any level can be under selection for dosage compensation, trimming based on expression level can skew final X:AA ratio[2]. Because of the conflicting results of dosage compensation studies that calculate the X:AA ratio, RXE and X to proto-X (Z chromosome) expression analyses are becoming more popular. The trimming parameters used by Xiong et al were too stringent and likely introduced bias into their calculation of the X:AA ratio. Xiong et al utilized unique mapping and discarded reads spanning splice junctions, which results in lower Relative X expression (RXE) values. RXE is calculated for all tissues to standardize and compare dosage compensation. Relative X expression can be calculated by first doing a log2-transformation of FPKM, RPKM, or TPM values, making the data more normally distributed and reducing outliers. Next, the mean autosomal expression is subtracted from the mean X chromosome expression (i.e. relative X expression (RXE)= $\log_2(X)- \log_2(A)$)[2].

X to proto-X (Z chromosome) expression analyses compare the expression of X-linked genes in one species to their autosomal orthologs in another species[4]. Genes with 1:1 orthologs in chickens and humans have recently been studied by RNA-seq analysis of an amniote-wide dataset[4,51]. When one-to-one orthologs within *C. elegans* and *P. pacificus* are tested, there is lower expression of the orthologs that are X-linked, suggesting that X upregulation is absent[88]. An explanation for the difference in X upregulation status in the two analyses can be understood if X upregulation acts locally, instead of globally on tissues and dosage sensitive X-linked genes[88].

Mapping parameters include unique and non-unique mapping. Unique mapping aligns short reads to a single location in the genome and excludes reads that can map to multiple locations. Non-unique mapping allows for mapping of multi-mapping reads and paralogs, which are a result of gene duplication that occurs when a homolog is lost. Paralogous genes are important to study because gene duplication may be part of the mechanism of dosage compensation[2]. There are multiple genome annotation databases including RefSeq, Ensembl, Gencode, and the UCSC annotation database and the reference genome selected impacts mapping efficiency and gene expression estimation[89]. Jue et al found substantial variation in the estimation of RXE within the same human tissue type when using different genome annotations in the mapping program Cufflinks[2]. Library preparation methods vary in the initial selection. For example, Illumina-based library preparation selects for the poly-A tail of mature mRNA using oligodT magnetic beads and creates a 3' bias. In addition, the Illumina TruSeq stranded mRNA library preparation kit allows for strand specific sequencing. Lastly, depth of sequencing coverage is essential in detecting lowly expressed genes, which contribute to accurate RXE values[90].

CHAPTER 2

2.1 X CHROMOSOME INACTIVATION

X chromosome inactivation (XCI) involves the transcriptional silencing of one of the X chromosomes of either maternal or paternal origin in every diploid cell of female mammals[57]. It is also a mechanism of dosage compensation that is unique to female placental mammals. The inactive X chromosome was first observed in 1949 as condensed heterochromatin only present in the nuclei of female somatic cells, which is now called the 'Barr Body'[91] (Figure 1). The random form of X chromosome inactivation was first noted by Mary Lyon while studying coat color in mice, she observed that female mice that were heterozygous for an X-linked gene responsible for coat color displayed mosaic phenotypes of many different coat color patterns[17]. While XCI is random in mammalian somatic cells, it can either be random or imprinted in extra-embryonic tissues depending on the species. The imprinted form of X chromosome inactivation involves preferential silencing of the paternally inherited X chromosome, while the maternal X remains active[58]. The inactive X chromosome can be studied by fluorescence and immunostaining of several chromosome-wide markers[92]. The markers include X-inactive specific transcript (Xist) RNA coating, histone H3 lysine 27 methylation (H3K27me), and a chromatin modifying protein Eed[92]. In addition, the expression of Xist has been studied in mammalian species.

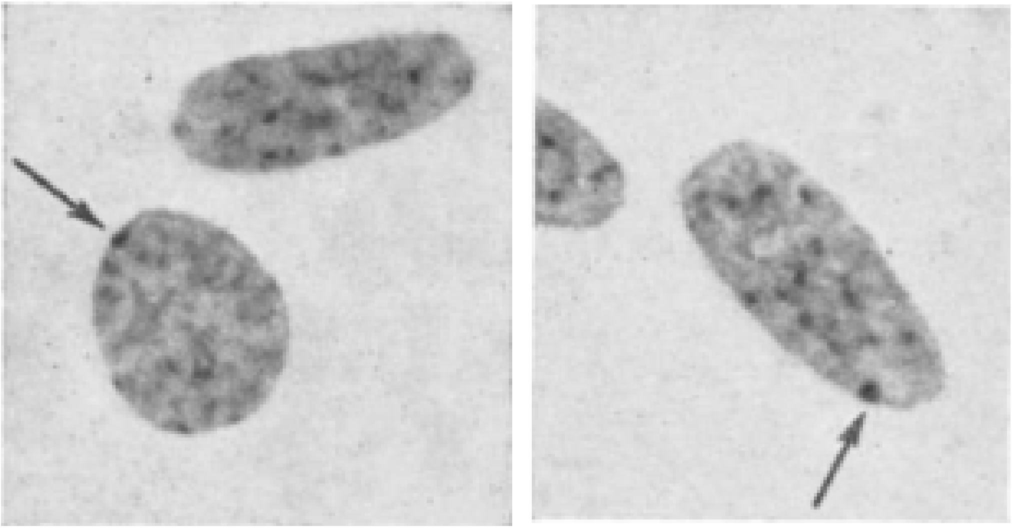


Figure 1. The black arrows point to the Barr body, the inactive X chromosome, within mouse nuclei. From Kenney et al., 1965.

2.1.1 Random, imprinted, and skewed XCI

Two distinct patterns of X chromosome inactivation occur in eutherian mammals, random XCI and imprinted XCI (Figure 1). Somatic cells of female eutherian mammals undergo random X chromosome inactivation in which one of the two X chromosomes in each cell is transcriptionally inactivated at random[17]. The inactivated X chromosome is condensed heterochromatin termed a Barr body and can be of either maternal or paternal origin[91]. A counting mechanism within the zygote first distinguishes the number of X chromosomes present in the cell and then the initiation mechanism selects an X chromosome to be inactivated. The multi-step process ensures a single, functional X chromosome in each adult diploid cell[93]. X chromosome inactivation is also apparent in human females with more than two copies of the X chromosome, as all X chromosomes except one are inactivated. Ohno described this as the "n-1" rule where an individual with n X chromosomes will have n-1 inactivated[77]. Random XCI is present in the embryonic and extra-embryonic tissues of the horse, mule, rabbit, and human[21,94,95].

Imprinted XCI is a nonrandom inactivation whereby the paternally inherited X chromosome is preferentially silenced and the maternally inherited X remains active[58]. Initiation of imprinted XCI begins in murine preimplantation embryos and is maintained in extra-embryonic tissues, most notably the placenta[96]. Placental tissues of cow[97,98] and rat[99] exhibit imprinted XCI. Random and imprinted XCI can be present in different cells from different embryonic layers within the same animal[58,100,101,102].

Skewed XCI is a disruption in normal X chromosome inactivation patterns. Inactivation predominantly occurs to either the paternally inherited X or the maternally inherited X[103]. Studies characterize skewed or non-random XCI as 75% of cells have the same inactive X[58].

2.1.2 Human and mouse XCI

XCI is random in human somatic cells. X chromosome inactivation has been observed in human preimplantation embryos starting at the eight cell stage[104]. It has been unclear if XCI is random or imprinted in the human placenta[101]. A recent study of allele specific expression of SNPs on the X chromosome of human placenta samples revealed variable patterns of X chromosome inactivation organized into patches with the paternal or maternal X inactivated respectively[94].

XCI is random in mouse somatic cells. XCI follows an imprinted pattern in the mouse and rat placenta[105,99]. The trophectoderm and primitive endoderm display imprinted XCI and later become the placenta and yolk sac[106]. In the mouse, the paternal X chromosome is preferentially silenced and imprinted XCI is established at the four cell stage in embryonic preimplantation development[107]. In the inner cell mass of the blastocyst, the paternal X is reactivated and is followed by random X chromosome inactivation of the maternal or paternal X chromosome[108]. The inner cell mass gives rise to the embryo proper and later adult somatic tissues, which maintain random XCI[109].

2.1.3 Bovine XCI

XCI occurs in cattle[97,110]. The gene Monoamine oxidase type A (MAOA) has been used to study XCI in cattle because it is an X-linked housekeeping gene in humans[111] and mice[112] that is subjected to XCI. Xue et al first reported that Xist and MAOA were expressed in bovine kidney, brain, liver, heart, and spleen[97]. In bovine in vitro embryos, the MAOA gene displayed only maternal expression in the morula stage indicating the establishment of imprinted XCI. The establishment of imprinted XCI occurs later in the bovine than in the human and the mouse. Subsequently, the paternal X was reactivated in the inner cell mass of the blastocyst[113].

XCI is random in bovine somatic cells. Random X chromosome inactivation is re-established between day 7 and day 14, corresponding to the blastocyst and early elongation stages[114]. In addition, two X-linked genes ubiquitin conjugating enzyme E2A (UBE2A) and spermidine/sperine N-acetyltransferase 1 (SAT1) were shown to partially escape XCI between these embryonic development stages[114]. In re-analysis of bovine blastocyst microarray expression data, higher expression of X-linked genes was observed in female blastocysts than male blastocysts indicating incomplete X chromosome inactivation and dosage compensation[115].

In cattle, different data are reported for the status of XCI in the placenta. In extra-embryonic tissue, the MAOA maternal allele is solely expressed[97]. Conversely, a recent study of the bovine intercotyledonary chorioallantois described random expression of the alleles of X-linked genes[116]. Day 15 bovine extraembryonic membranes showed biallelic Xist expression, while the trophectoderm cell line CT1 showed monoallelic Xist expression[116].

2.1.4 Ovine XCI

X chromosome inactivation is known to exist in sheep, but little is known about its onset and regulation. An inactive X chromosome was first observed in sheep fetal ovaries by radioactive staining and autoradiography as a heterochromatic chromosome located peripherally in comparison to the other chromosomes[117]. Studies have investigated the transcriptional activity of X-linked genes on the active X in ovine female embryos. It has been found that X-linked and Y-linked genes are transcriptionally active in ovine preimplantation embryos starting at the 2-cell stage[118]. This is consistent with RT-PCR reports in mouse and human embryos[119-122] .

A study of Xist in sheep by Zhao et al found that Xist does not have tissue specific expression in female sheep, consistent with its expression from the inactive X in somatic cells. In brain, kidney, liver, spleen, lung, small intestine, ovary, muscle, and heart of two-day old lambs, Xist mRNA had little difference in expression[123].

2.1.5 Genes escaping XCI

Genes that escape X chromosome inactivation show expression from both the active and inactive X chromosomes[124]. An escaped gene has ≥10% expression from the allele that was inactivated[125]. Variation exists in the amount of expression from the inactive allele for individual genes and for the same gene in different tissues[61]. 10-15% of genes on the human X chromosome escape inactivation. The escaping genes are located in both the still recombining pseudoautosomal regions (PARS) and in the X-specific region in X-linked genes with an active homolog on the Y chromosome[126]. In somatic cells of mice, almost all X-linked genes remain inactivated and 3% escape inactivation[127,128]. Escape from XCI is also common in other placental mammals such as cows[129]. There is a pattern in the distribution of escaping genes, possibly due to relative distance from Xist as most genes reveal clustering in the distal portion of the short arm of the X chromosome[130]. Studying genes that escape XCI helps to compare the difference in epigenetic marks and to better understand the mechanism of inactivation on the rest of the X chromosome.

CHAPTER 3

3.1 EFFECT OF MATERNAL NUTRITION ON FETAL EPIGENETICS AND DEVELOPMENT

While the fetal genome plays a large role in the growth and development of the fetus, increasing evidence supports a strong influence of the intrauterine environment on fetal development[131]. The phenomenon of fetal programming maintains that the intrauterine environment can induce changes in expression of the fetal genome and can permanently alter offspring physiology, structure, metabolic function, and growth postnatally[132].

Maternal nutrition can affect the intrauterine environment and epigenetically alter the fetal genome. Possible epigenetic modifications include DNA methylation and histone acetylation. Maternal nutrition studies in mice and rats have revealed that maternal diet can alter fetal gene expression through epigenetic modification, inducing physiological changes in the developing fetus[133,134]. Limited data exists pertaining to the effect of maternal nutrition on epigenetic and gene expression changes in sheep. In a recent study, pregnant ewe nutrition was found to change the expression of specific genes under strong epigenetic regulation known as imprinted genes in the sheep[135].

3.1.1 Poor Maternal Nutrition

Understanding the effect of poor maternal nutrition on sheep fetal development and the fetal genome is relevant to sheep production and metabolic disease[136]. Over and undernutrition of pregnant ewes are both representative of poor maternal nutrition during gestation. Because the sheep industry uses forage based systems[137], pregnant ewes are subjected to both under and overnutrition based on the quality and quantity changes in forage in different seasons. Pregnant ewes have been used extensively as models for human pregnancy. Maternal-fetal interactions such as metabolic function and nutrient transport can be studied due to the ability to sample from fetal and maternal vasculature in sheep without the use of anesthesia[31]. In addition, researchers are also able to study nutritional programming, such as that induced in fetal growth restriction. In different stages of pregnancy, maternal nutrition can significantly alter offspring physiology, structure, and metabolism[138]. Maternal nutrition contributes to both fetal and placental growth[139]. Nutrient deficiency as a result of poor maternal nutrition during gestation has been shown to severely impair normal fetal and placental growth[140]

Sheep research by the Govoni, Reed, and Zinn labs has focused on the effect of poor maternal nutrition on both the pregnant ewe and offspring by utilizing a control-fed (100% NRC), restricted-fed (60% NRC), and over-fed (140% NRC) treatment design[141-146]. Ewes compensate for poor maternal nutrition by either reducing or increasing their own body weight. Pillai et al Restricted ewes and overfed ewes had decreased and increased body weight and body condition score respectively when compared to the control ewes at day 135 and birth ($P \leq 0.05$)[146]. Fetal body weight did not differ between maternal nutrition groups at day 45, day 90, or day 135[146]. Reed et al found that ewe body weight at the end of gestation was reduced by 18.9% (18.3 ± 3.6 kg; $P < 0.01$) in the restricted group and ewe body weight increased by 6.6% (119.7± 3.6 kg; $P < 0.10$) when compared to the control group (112.9± 3.6 kg)[141].

Maternal nutrition has been found to affect metabolic processes in sheep blood and fetal tissues. Hoffmann et al reported that overnutrition in ewes was found to affect metabolism in 3 month old sheep by increasing circulating triglycerides indicating possible future metabolic disease[147]. The poor maternal nutrition studies by the aforementioned labs have also shown that poor maternal nutrition effects fat, muscle, and bone development prenatally and postnatally[145]. In addition, maternal nutrition can affect critical organ development such as the trend for increased heart size in sheep born to overfed mothers[147]. Pillai et al reported that no differences were found between maternal nutrition groups in fetal kidney and liver weight at day 90, day 135, and birth[146].

In a separate fetal sheep transcriptome study, muscle and adipose tissue were evaluated for the effect of maternal diet during mid-to-late gestation. Different maternal diet resulted in gene expression and energy metabolism changes in both tissues[33]. In twins and offspring from undernourished ewes, epigenetic modifications occur in sheep fetal hypothalamic pathways that regulate energy balance, altering these pathways increases the offspring's chance of obesity and/or metabolic disease later in life[136]. The findings of maternal diet altering fetal epigenetics is of particular interest because XCI is an epigenetically regulated process.

3.2 SUMMARY

Two mechanisms have evolved in mammals to balance the expression of X-linked genes between the sexes and to balance the expression ratio of the X chromosome to the autosomes. In Ohno's hypothesis, X-linked gene expression is doubled in both males and females, successfully balancing with the autosome expression in males. While it is well characterized in *Drosophila*, this pattern has only recently been observed in mammalian species. In mammalian females, X-chromosome inactivation (XCI) randomly and globally inactivates one of the X-chromosomes. Dosage compensation is known to be species, tissue, and developmental stage specific. This suggests the importance of studying dosage compensation and X chromosome inactivation in different species where research is limited such as the sheep. Sheep are a good model to investigate dosage compensation normally and under the effect of poor maternal nutrition. Poor maternal nutrition, both under and overnutrition are common based on the changes in quality and quantity of forage with the change of seasons. In humans, obesity and type 2 diabetes are late onset diseases that occur in response to earlier nutritional conditions that effect epigenetic marks such as histone tail modification and DNA methylation[148]. It has also been shown that supplementing or restricting folate, choline, or methionine in the maternal diet can affect DNA methylation pattern establishment in offspring[149,150]. Maternal nutrition can influence epigenetic modifications of the fetal genome and may result in changes in expression of X-linked genes.

3.3 OBJECTIVES

Dosage compensation and X chromosome inactivation have been studied thoroughly in mice and humans, but research in domestic species, livestock in particular, has lagged behind. Proper dosage compensation and XCI are needed for viable offspring and disruptions are associated with disease. With the advancement of RNA sequencing technology, more transcriptomic studies can be performed in domestic species to uncover the complex epigenetic mechanisms. Very few RNA sequencing experiments have been conducted in the sheep. Here we present the first RNA-seq experiment evaluating sex chromosome dosage compensation. In combination with the data of this study, two additional RNA-seq datasets (PRJEB6169) and (PRJNA254105) were added to achieve a more global view of dosage compensation in the sheep[151,152]. In addition, while it is known that maternal diet can influence epigenetic changes in the developing fetus, the effect of maternal diet on dosage compensation is unknown. By investigating the effect of maternal under and overnutrition on X-linked genes, we hope to uncover more insight into the compensatory nature of dosage compensation under an environmental stressor.

The first objective of this study was to characterize global dosage compensation in the sheep using data from this study and additional RNA-seq datasets (PRJEB6169) and (PRJNA254105). In the three combined datasets, we were able to analyze sheep dosage compensation in fetal brain, kidney, lung, day 14 embryos (PRJNA254105), adult and juvenile heart, liver, muscle, and rumen in both males and females (PRJEB6169). Specific brain tissue and female and male specific tissues (PRJEB6169) were also evaluated. We hypothesized that dosage compensation in the sheep would be incomplete, similar to that of the cow[115]. The second objective was to determine the effects of maternal control, restricted, and overfed diets on the expression of X-linked genes in fetal tissues at 135 days of gestation. We hypothesized that maternal diet may influence the X-linked genes that are expressed and their expression level in ovine fetuses.

3.4 MATERIALS AND METHODS

3.4.1 Animals

All animal protocols[141-147] were reviewed and approved by the University of Connecticut Institutional Animal Care and Use Committee. Animal breeding, feeding and care, necropsy, and sample collection were performed by the labs of Dr. Govoni, Dr. Reed, and Dr. Zinn. The research animals, fifteeen western white faced ewes and four line bred Dorset rams were purchased and shipped from Midwest farms and Indiana respectively. The relatedness of the rams is included in a pedigree chart (Figure 1) and the relatedness of the ewes is unknown. Estrous synchronization[153] of ewes was accomplished with progesterone controlled intravaginal drug release devices (Pfizer Animal Health; New York, NY, USA) and Lutaylase (Pfizer Animal Health). Ewes were bred live cover to one of the Dorset rams as previously described[141-147]. Pregnancy was confirmed by ultrasound on day 20 of gestation if a ewe was not re-marked by a ram, day 0 represents the initial marking of the ewe by the ram. On day 30 of gestation, pregnant ewes were individually housed and randomly assigned to control (100% NRC requirement), restricted (60%), or overfed (140%) diets calculated by the National Research Council requirement for total digestible nutrients for a ewe pregnant with twins[141,154]. Ewes were weighed weekly to track body weight gain and to adjust diets throughout the pregnancy. The ewes remained on their respective diets until day 135 of gestation, when they were euthanized and necropsied to collect fetal tissues.

3.4.2 Fetal brain, kidney, and lung sample collection and selection

Samples were collected and provided by the labs of Dr. Govoni, Dr. Reed, and Dr. Zinn. Within each treatment group, brain, kidney, and lung were collected from each fetus. The control group consisted of four ewes and eight fetuses (8 brain, 8 kidney, and 8 lung samples). The restricted group consisted of seven ewes and thirteen fetuses (13 brain, 13 kidney, and 13 lung samples). The overfed group consisted of six ewes and ten fetuses (10 brain, 10 kidney, and 10 lung samples) (Table 1). Samples were further selected for this study genetically representing each ram present in each nutritional treatment and selecting singletons over twins and triplets to increase the genetic diversity.

Brain, lung, and kidney were selected from seven control fetuses (3 females and 4 males), four restricted fetuses (1 female and 3 males), four overfed fetuses (3 females and 1 male) (Table 2). We included an increased number of fetuses in the control group to increase the power of studying normal dosage compensation in the sheep. Tissues were flash frozen in liquid nitrogen and were stored at -80°C until RNA extraction was performed.

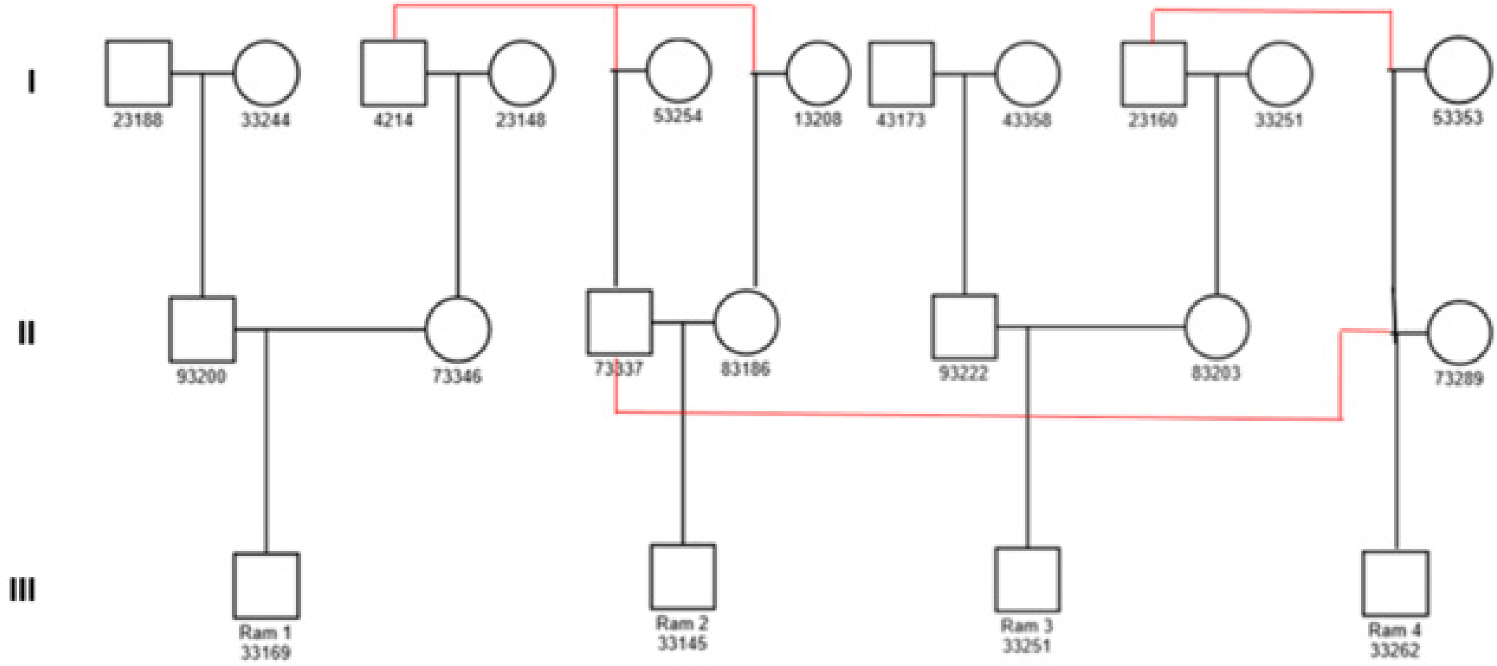
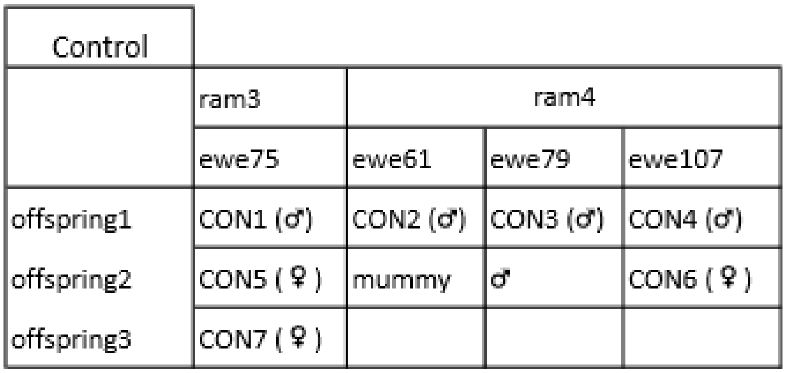
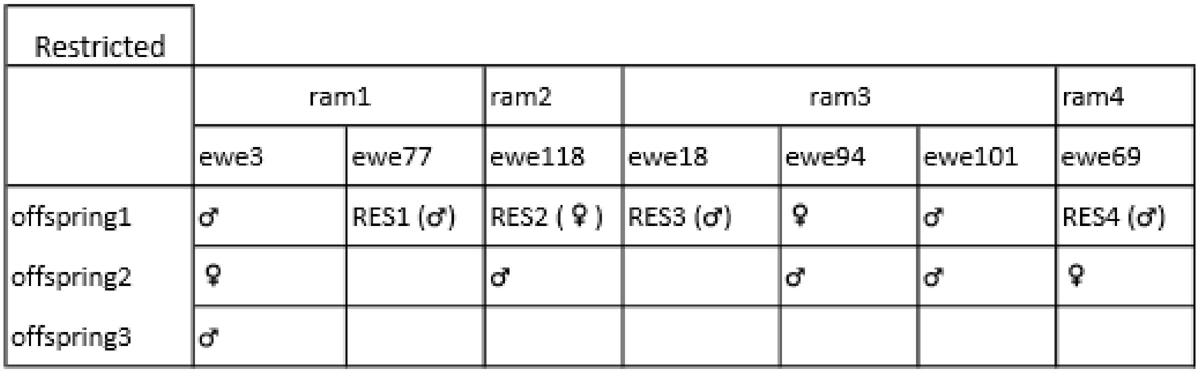


Figure 1. Ram Pedigree chart. The Pedigree shows the relatedness of the rams used in the study. Ram 2 & Ram 4 are half siblings with the same father 73337 and different mothers, Ram 1 & 2 have the same grandfather, Ram 3 & Ram 4 have the same maternal grandfather, Ram 1's mother is half sibling to Ram 2's father and mother, and Ram 3's mother is half sibling to Ram 4's mother.

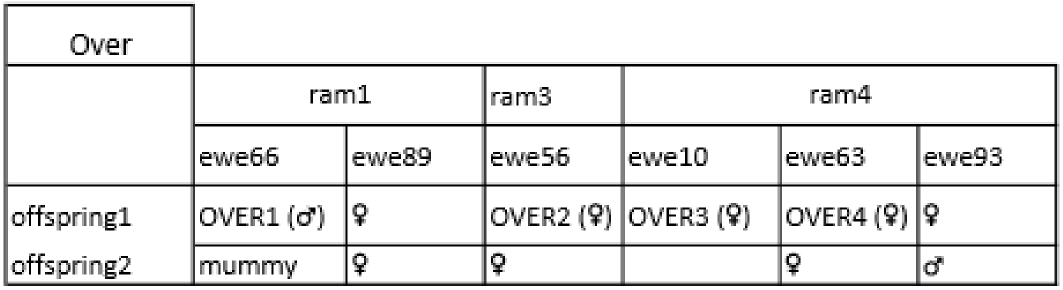
Table 1 Sample Selection



Control				
	ram3	ram4		
	ewe75	ewe61	ewe79	ewe107
offspring1	CON1 (♂)	CON2 (♂)	CON3 (♂)	CON4 (♂)
offspring2	CON5 (♀)	mummy	♂	CON6 (♀)
offspring3	CON7 (♀)			



Restricted							
	ram1		ram2	ram3			ram4
	ewe3	ewe77	ewe118	ewe18	ewe94	ewe101	ewe69
offspring1	♂	RES1 (♂)	RES2 (♀)	RES3 (♂)	♀	♂	RES4 (♂)
offspring2	♀		♂		♂	♂	♀
offspring3	♂						



Over						
	ram1		ram3	ram4		
	ewe66	ewe89	ewe56	ewe10	ewe63	ewe93
offspring1	OVER1 (♂)	♀	OVER2 (♀)	OVER3 (♀)	OVER4 (♀)	♀
offspring2	mummy	♀	♀		♀	♂

Table 2 Total Sample Collection

Treatment	EWE ID	Ram ID	Number of offspring	Kidney	Brain	Lung
Control	61*	4	2*	F1	F1	F1
	75	3	3	F1, F2, F3	F1, F2, F3	F1, F2, F3
	79	4	2	F1, F2	F1, F2	F1, F2
	107	4	2	F1, F2	F1, F2	F1, F2
Over	10	4	1	F1	F1	F1
	56	3	2	F1, F2	F1, F2	F1, F2
	63	4	2	F1, F2	F1, F2	F1, F2
	66*	1	2*	F1	F1	F1
	89	1	2	F1, F2	F1, F2	F1, F2
	93	4	2	F1, F2	F1, F2	F1, F2
Restricted	3	1	3	F1, F2, F3	F1, F2, F3	F1, F2, F3
	18	3	1	F1	F1	F1
	69	4	2	F1, F2	F1, F2	F1, F2
	77	1	1	F1	F1	F1
	94	3	2	F1, F2	F1, F2	F1, F2
	101	3	2	F1, F2	F1, F2	F1, F2
	118	2	2	F1, F2	F1, F2	F1, F2

* one fetus sampled, 2nd mummy. F=fetus

3.4.3 RNA isolation and quality control

RNA was extracted from fetal brain, kidney, and lung using TRIzol (Invitrogen, Grand Island, NY) according to the manufacturer's instructions. The homogenization procedure was optimized at a frequency of 30 Hz for 3-4 minutes using the TissueLyser II homogenizer (Qiagen Sample and Assays Technology, USA). RNA samples were purified using the Qiagen RNeasy kit (Qiagen, Valencia, CA, USA) and a Dnase treatment (Qiagen, Valencia, CA, USA) was applied. RNA quality was examined by spectrophotometry and with the Agilent Total RNA Nano kit (Agilent Technologies, Santa Clara, CA). All samples had RNA Integrity Numbers (RINs) greater than or equal to 7 (RIN $\geq$ 7) (Table 3). The concentration of total RNA used in library preparation was calculated using a Qubit 2.0 Fluorometer (Thermo Fisher scientific; Waltham, MA, USA) (Table 3).

Table 3 RNA Quality and Quantity

Tissue	Treatment	Sample ID	RIN	260/280	Qubit RNA concentration (ng/μl)
Brain	Control	1	8.3	2.20	1176
		2	8.7	2.15	230
		3	7.0	2.23	85.8
		4	8.1	2.22	472
		5	8.8	2.14	1170
		6	7.0	2.14	576
		7	8.0	2.14	1158
	Restricted	1	8.3	2.19	724
		2	7.5	2.14	252
		3	8.4	2.13	103
		4	8.1	2.16	858
	Overfed	1	8.2	2.18	720
		2	7.4	2.19	220
		3	8.5	2.19	468
		4	9.0	2.18	1266
Lung	Control	1	9.5	2.18	796
		2	9.4	2.14	396
		3	9.4	2.18	510
		4	9.0	2.17	464
		5	10	2.14	573
		6	9.7	2.14	1860
		7	9.9	2.14	1920
	Restricted	1	8.2	2.16	1448
		2	8.9	2.14	422
		3	8.8	2.13	384
		4	9.4	2.12	342
	Overfed	1	9.0	2.17	988
		2	9.1	2.18	540
		3	9.3	2.19	488
		4	9.1	2.18	684

Tissue	Treatment	Sample ID	RIN	260/280	Qubit RNA concentration (ng/µl)
Kidney	Control	1	8.9	2.12	450
		2	7.3	2.15	1264
		3	9.3	2.08	736
		4	7.8	2.11	1160
		5	9.1	2.14	1470
		6	7.3	2.14	1092
		7	8.3	2.14	2226
	Restricted	1	8.8	2.04	2412
		2	9.1	2.10	1996
		3	7.5	2.18	1280
		4	9.3	2.16	330
	Overfed	1	9.2	2.10	494
		2	8.4	2.10	524
		3	8.9	2.15	648
		4	7.8	2.20	1224

3.4.4 Library preparation, quality control, and quantification

An average input of 2 µg of total RNA/sample was used to prepare 45 cDNA libraries for sequencing following the manufacturer's instructions by the Illumina TruSeq mRNA library prep kit. This protocol involves purification of mRNA by poly-T oligo-attached magnetic beads, fragmentation, priming with random hexamers, incorporation of dUTTP during second strand synthesis for strand specific sequencing, and adapter ligation. Library quantification was performed using the KAPA qPCR kit and the Agilent DNA 1000 kit (Agilent Technologies, Santa Clara, CA) was used with the Agilent Bioanalyzer to obtain average library lengths and to ensure the absence of adapter dimerization (Table 4).

Table 4 Library Preparation

Tissue	Treatment	Sample ID	Total RNA input for library prep (ng)	nanodrop cDNA library concentration (ng/ul)	Bioanalyzer average library lengths (bp)	KAPA qRT-PCR Average Stock Library Concentration (nM)
Brain	Control	1	2352	84.1	247	348
		2	1380	66.2	266	254
		3	1030	66.8	240	418.4
		4	2832	68.8	255	699.3
		5	2340	86.7	257	194.1
		6	2304	78.5	261	280.3
		7	2316	78.8	246	361.7
	Restricted	1	3982	90.1	268	551.7
		2	1512	71.6	247	489.5
		3	1236	86.2	251	577
		4	2145	84.6	248	433.5
	Overfed	1	2160	82.9	247	330.5
		2	1320	71.7	252	565
		3	2808	90.7	245	657.8
		4	2279	84.0	250	294.2
Lung	Control	1	2388	79.2	240	472.3
		2	2376	86.2	243	400
		3	3060	89.1	277	445
		4	2088	70.4	274	302
		5	2292	86.3	252	299
		6	2232	82.3	275	228
		7	2304	78.9	275	328
	Restricted	1	2172	56.1	268	285
		2	2532	70.4	245	298
		3	2304	85.7	257	573.6
		4	2052	87.1	247	661
	Overfed	1	2174	80.1	264	310.8
		2	3240	85.3	243	263.4
		3	2928	79.6	255	563.3
		4	2052	78.0	244	247.2

Tissue	Treatment	Sample ID	Total RNA input for library prep (ng)	nanodrop cDNA library concentration (ng/ul)	Bioanalyzer average library lengths (bp)	KAPA qRT-PCR Average Stock Library Concentration (nM)
Kidney	Control	1	2700	100.2	233	2229.3
		2	2275	76.4	262	403
		3	2208	89.5	242	372.3
		4	2320	83.4	245	273
		5	2205	87.0	285	313.3
		6	2190	79.9	267	393.9
		7	2226	82.2	268	356
	Restricted	1	2412	49.0	220	363.1
		2	2395	61.1	242	426.2
		3	2304	86.6	246	324
		4	1980	89.8	251	2065.7
	Overfed	1	2470	92.7	244	1454.8
		2	2096	50.9	267	344
		3	3888	82.2	259	1276.8
		4	2203	82.9	261	385

3.4.5 RNA sequencing

Barcoding adapters were added in the processing of the 45 cDNA libraries, they were pooled at a concentration of 4 nM, and libraries were sequenced on the Illumina NextSeq 500 Platform at the Center for Genome Innovation, University of Connecticut. The RNA-seq libraries were sequenced with 2 ×75 bp paired-end reads on the NextSeq500 in three sequencing runs and one percent PhiX DNA was spiked in as an internal control each time. A total of 1160, 576, and 413 million raw sequencing reads passing filtering were obtained for sequencing runs 1, 2, and 3 respectively. Overall, we obtained 2149 million raw sequencing reads passing filtering from three sequencing runs of 45 fetal tissue samples.

Table 5 RNA Sequencing Runs and Adapters

RNA-Seq Run number	Tissue	Treatment	Sample ID	Adapter number	Adapter sequence	Mean % Reads Identified (Post Filtering)	% Mapped Reads	Number of Mapped Reads
Run 1	Brain	Control	1	11	GGCTAC(A)	3.8	0.89	17,261,538
			2	7	CAGATC(A)	1.5	0.90	6,629,510
			3	6	GCCAAT(A)	3.1	0.90	15,055,094
			4	13	AGTCAA(C)	2.0	0.91	9,932,518
			5	22	CGTACG(T)	8.2	0.90	40,040,450
			6	23	GAGTGG(A)	5.0	0.90	24,500,963
			7	25	ACTGAT(A)	4.1	0.90	20,089,481
		Restricted	1	12	CTTGTA(A)	3.0	0.91	14,836,023
			2	14	AGTTCC(G)	3.0	0.91	14,979,444
			3	3	TTAGGC(A)	2.2	0.91	10,926,581
			4	10	TAGCTT(A)	3.7	0.91	17,921,173
		Overfed	1	9	GATCAG(A)	4.5	0.90	21,904,593
			2	4	TGACCA(A)	2.2	0.91	11,086,560
			3	1	ATCACG(A)	2.6	0.91	12,564,537
			4	8	ACTTGA(A)	5.7	0.91	27,966,870
	Lung	Control	1	5	ACAGTG(A)	3.4	0.90	16,897,583
			2	27	ATTCCT(T)	4.1	0.90	20,199,107
			3	2	CGATGT(A)	3.8	0.90	18,491,810
			4	20	GTGGCC(T)	4.8	0.90	23,499,521
			5	15	ATGTCA(G)	5.4	0.90	26,421,784
			6	16	CCGTCC(C)	6.2	0.89	29,926,965
			7	19	GTGAAA(C)	5.0	0.90	24,366,852
		Restricted	2	21	GTTTCG(G)	6.9	0.90	33,525,561
			4	18	GTCCGC(A)	2.8	0.90	13,964,369

RNA-Seq Run number	Tissue	Treatment	Sample ID	Adapter number	Adapter sequence	Mean % Reads Identified (Post Filtering)	% Mapped Reads	Number of Mapped Reads
Run 2	Lung	Overfed	1	20	GTGGCC(T)	12.9	0.89	60,526,905
			2	25	ACTGAT(A)	4.6	0.91	22,129,781
			3	22	CGTACG(T)	3.0	0.91	14,345,916
			4	16	CCGTCC(C)	6.5	0.90	30,242,283
		Restricted	1	21	GTTTCG(G)	5.4	0.90	25,508,281
			3	23	GAGTGG(A)	2.9	0.90	13,539,473
	Kidney	Overfed	1	19	GTGAAA(C)	1.6	0.91	7,591,885
			2	15	ATGTCA(G)	4.6	0.91	21,765,983
			3	13	AGTCAA(C)	1.5	0.91	7,026,844
			4	18	GTCCGC(A)	3.8	0.91	17,920,199
		Restricted	3	14	AGTTCC(G)	5.1	0.90	24,124,589
Run 3	Kidney	Control	1	21	GTTTCG(G)	1.0	0.91	4,836,961
			2	16	CCGTCC(C)	3.9	0.90	17,858,838
			3	23	GAGTGG(A)	4.4	0.91	20,386,772
			4	25	ACTGAT(A)	5.2	0.90	24,242,852
			5	15	ATGTCA(G)	5.4	0.91	25,146,483
			6	18	GTCCGC(A)	3.8	0.90	17,692,978
			7	19	GTGAAA(C)	4.5	0.90	20,656,829
		Restricted	1	22	CGTACG(T)	4.5	0.91	20,732,144
			2	27	ATTCCT(T)	4.1	0.90	18,835,286
			4	20	GTGGCC(T)	1.1	0.91	4,992,172

3.4.6 Additional RNA-seq datasets

To analyze X dosage compensation across tissue types, female and male specific tissues, and embryos, two additional RNA-seq datasets were downloaded from Sequence Read Archive (SRA) (http://www.ncbi.nlm.nih.gov/sra) under the accession numbers PRJNA254105[152] and PRJEB6169[151]. The additional datasets include day 14 embryos (PRJNA254105), adult and juvenile heart, brain, liver, biceps, rumen, and female and male specific tissues (PRJEB6169). Female specific tissues include cervix, ovarian follicles, ovary, uterus, and corpus luteum. Male specific tissues include testes and epididymis. To normalize within each dataset, the mRNA level of each gene was estimated by transformed transcripts per kilobase million (TPM) and was quantified using IsoEM (version 1.1.4).

3.4.7 RNA-seq data trimming and mapping

Sequence adapter and quality trimming were done using Sickle v1.33 with the parameters Q score ≥ 30 and length ≥ 20 (-q30, -l20). After filtering, read quality was checked using FastQC v0.11.3.. Filtered RNA-seq reads from fetal day 135 tissues were aligned to the sheep reference genome Oar_v4.0 using Hisat2 v2.0.5[155] (Table 5).

3.4.8 RNA-seq data assembly and Dosage compensation calculation

Aligned reads for each tissue from our study and two sheep online public datasets were assembled using IsoEM v1.1.4 to estimate gene expression in transcripts per kilobase per million (TPM). Only genes with TPM>1 were selected for later analysis, and were later log 2-transformed. Gene expression chromosome-wide distributions for 20,549 genes were isolated by chromosome and then plotted in R. The relative X expression (RXE) was calculated for X-linked genes (x) and autosomal genes (a) by following formula:

$$RXE = \log_2\left(\frac{x}{a}\right) = \log_2 x - \log_2 a$$

An *RXE* greater than or equal to 0 represents up-regulation of X and dosage compensation. An *RXE* between 0 and -1 indicates X up-regulation, but incomplete dosage compensation. When *RXE* is equal to -1 dosage compensation is nonexistent.

3.4.1.9 Gene ontology of X-linked genes

A Gene Ontology (GO) classification was conducted on the combined list of expressed X-linked genes found in sheep brain, lung, and kidney for all treatments using DAVID 6.8 [156,157]. Specifically, a functional annotation enrichment analysis revealed GO terms and major functional categories. Major Gene Ontology terms are presented with Benjamani-Hochberg adjusted P-values.

3.5 RESULTS

3.5.1 Dosage compensation in sheep

Here we present the first comprehensive study of dosage compensation in the sheep using RNA sequencing data of day 135 fetal sheep somatic tissues. The addition of two other RNA seq datasets allows for a more precise study of dosage compensation in male and female specific organs and within more varied somatic tissues. Relative X expression (RXE) is calculated for all tissues to standardize and compare dosage compensation.

3.5.2 Dosage compensation in ovine major organs

The major organs: heart, liver, muscle, rumen, placenta, and day 14 embryos displayed incomplete dosage compensation. The mean RXE values had a range of -0.19 to -0.05 in the major organs evaluated, with an overall average RXE of -0.12 respectively (Figure 1A). No significant difference was observed between male and female organ tissues. Dosage appears to be more highly compensated in the brain. Average RXE had a range of -0.12 to 0.16 in brain cerebrum, cerebellum, hypothalamus, and pituitary (Figure 1B). All brain tissues, with the exception of cerebellum displayed complete dosage compensation. Compared to the other major organs studied, brain had the highest overall average RXE of 0.01.

3.5.3 Dosage compensation in ovine female specific tissues

Incomplete dosage compensation was observed in juvenile and adult female specific cervix, ovarian follicles, ovary, uterus, and corpus luteum. The overall RXE varied slightly between juvenile and adult female tissues with an average RXE of -0.19 for juvenile tissues and an average RXE of -0.15 for adult tissues respectively (Figure 2A). The combined juvenile and adult female specific tissues ranged in average RXE from -0.32 to -0.03.

3.5.4 Dosage compensation in ovine male specific tissues

Two male specific tissues, testes and epididymis were studied. An interesting pattern was observed, as there was low dosage compensation in the testes at an average RXE of -0.84 and high compensation of 0.32 in the epididymis (Figure 2B). The average RXE of the two male specific tissues was -0.33. This pattern has been observed in drosophila, mouse, human, and rat testes[26].

3.5.5 Dosage compensation and maternal nutrition

Dosage compensation was investigated for changes due to maternal over or undernutrition. Fetal control tissues, both male and female, displayed incomplete dosage compensation. The mean RXE values had a range of -0.14 to -0.06 in brain, lung, and kidney, with an overall average RXE of -0.10 respectively (Figure 3A). Fetal tissues from the overfed treatment displayed incomplete dosage compensation. Mean RXE values ranged from -0.09 to -0.07 in the three tissues. Overall average X: autosome expression ratio of overfed fetal tissues was -0.08 (Figure 3B). Fetal tissues from the restricted treatment displayed incomplete dosage compensation. Mean RXE values ranged from -0.13 to -0.09 in the three tissues. Overall average X: autosome expression ratio of restricted fetal tissues was -0.11 (Figure 3C). The pattern of dosage compensation remained fairly consistent between treatment groups.

3.5.6 X-linked genes in ovine somatic tissues

The mean number of expressed X-linked genes was calculated separately for each tissue type in the control, restricted, and overfed groups (Table 1). This allows for an even comparison between the effect of maternal nutrition on the number of X-linked genes expressed in a single tissue. The mean number of expressed X-linked genes for the control group was 459 ± 3.8 in the brain, 442.7 ± 2.7 in the kidney, and 429.3 ± 1.0 in the lung. Significance was tested using the Kruskal-Wallis one-way ANOVA in the IBM SPSS software. The mean did not change significantly based on maternal nutrition. The ten most highly expressed X-linked genes were investigated in female and male control brain, kidney, and lung. X-linked genes with the highest expression in fetal brain, lung, and kidney displayed tissue specificity as only four genes were consistent among all three tissues.

The four genes in common were thymosin beta 4, X-linked (TMSB4X), ribosomal protein L10 (RPL10), ribosomal protein L39 (RPL39), and ribosomal protein S4, X-linked (RPS4X). Large gene overlap between the three treatments revealed that these genes are highly expressed irrespective of treatment. TMSB4X, RPL10, and RPS4X had the highest expression in the lung while RPL39 had the highest expression in both the lung and kidney. Thymosin beta 4, X-linked

(TMSB4X) is in a highly conserved class of small proteins found in immune tissues, where it functions in wound healing, anti-inflammation, cell survival, and apoptosis[158]. Rengaraj et al observed that rat TMSB4X is expressed at an intermediate level in brain and kidney and chicken TMSB4X is expressed at almost equal measure in the brain and is 1.76-fold lower in the kidney[158]. Ribosomal protein L10 (RPL10) encodes a ribosomal protein that is part of the large ribosomal subunit and participates in ribosome function and biogenesis[159]. RPL10 has been studied in bovine blastocyst formation and it was found that RPL10 expression is higher *in vivo* at the 8-cell stage than *in vitro*[160]. Ribosomal protein L39 (RPL39) also encodes a ribosomal protein that is part of the large ribosomal subunit. Mutations in RPL39 lead to the initiation and metastasis of tumors, which have been studied in breast and lung cancer[161]. Ribosomal protein S4, X-linked (RPS4X) encodes the ribosomal protein S4 that is part of the small ribosomal subunit. This protein is also encoded by ribosomal protein S4, Y-linked (RPS4Y). RPS4X has been observed to have similar expression in bovine somatic cell nuclear transfer (SCNT) embryos and SCNT freemartin embryos[162].

In control female and male brain, the top ten highly expressed X-linked genes were consistent, but had different expression measured in TPM based on fetal sex (Figure 4). The gene PLP1 (Proteolipid Protein 1) had the highest expression in the brain at 2422.5 ± 362.2 TPM and 2117.8 ± 1044.5 respectively for males and females. NGFRAP1 (Nerve Growth Factor Receptor-Associated Protein 1) aka BEX3 (Brain Expressed X-Linked 3) had the lowest expression at 653.2 ± 84.7 TPM and 607.2 ± 8.2 TPM respectively for males and females. The genes GDI1 (GDP Dissociation Inhibitor 1), TSPAN7 (Tetraspanin 7), GPM6B (Glycoprotein M6B), and SYP (Synaptophysin) were also among the most highly expressed.

In control female and male kidney, seven of the top ten highly expressed X-linked genes were consistent, but had different expression measured in TPM based on fetal sex (Figure 5). Female control kidney included the genes CAPN6 (Calpain 6), S100G (S100 Calcium Binding Protein G), and SSR4 (Signal Sequence Receptor Subunit 4), while male control kidney included the genes PGRMC1 (Progesterone Receptor Membrane Component 1), BGN (Biglycan), and PGK1 (Phosphoglycerate Kinase 1). The gene RPL10 had the highest expression in the kidney at 3212.3 ± 281.3 TPM and 2977 ± 206.5 TPM respectively for males and females. PGK1 had the lowest expression in males at 340.2 ± 52.8 TPM and SSR4 had the lowest expression in females at 387.5 ± 18 TPM. The genes RPL36A (Ribosomal Protein L36a), SAT1 (Spermidine/Spermine N1-Acetyltransferase 1), GPC3 (Glypican 3) were also among the most highly expressed.

In control female and male lung, nine of the top ten highly expressed X-linked genes were consistent, but had different expression measured in TPM based on fetal sex (Figure 6). Female control lung included the gene GPC3 (Glypican 3), while male control lung included the gene FLNA (Filamin A). In addition, the gene MSN (Moesin) was not among the top ten highly expressed genes in brain and kidney. The gene RPL10 had the highest expression in the lung at 3443 ± 180.8 TPM and 4068 ± 371.7 TPM respectively for males and females. FLNA had the lowest expression in males at 400 ± 20 TPM and GPC3 had the lowest expression in females at 387.2 ± 23.2 TPM.

3.5.7 Gene Ontology analysis of X-linked genes

The total number of unique expressed X-linked genes in control, restricted, and overfed fetal day 135 brain, kidney, and lung was found by compiling files of all expressed X-linked genes in each tissue sample with a TPM>1 and filtering out the duplicates. This yielded a combined list of 513 expressed X-linked genes in brain, kidney, and lung from all of the treatment groups that were used for gene ontology analysis. The gene ontology terms, negative regulation of microtubule depolymerization and synapse had the lowest P-values of 7.43E-03and 9.70E-03 respectively. The largest gene counts were attributed to zinc binding, intracellular, chromatin binding, and nucleotide binding. Out of the 513 expressed X-linked genes in this study, 14 genes reside in the ruminant pseudoautosomal region (Figure 7). The genes are P2RY8, DHRSX, ZBED1, CD99, XG, GYG2, ARSE, MXRA5, PRKX, NLGN4X, STS, PNPLA4, TBL1X, and GPR143.

3.6 DISCUSSION

To our knowledge, this is the first comprehensive study of dosage compensation in the sheep. We conclude that dosage compensation is present, but incomplete in sheep somatic tissues. This is consistent with dosage compensation studies in cattle that have reported incomplete dosage compensation in somatic tissues[163,164]. Ovine male and female major organs had similar RXE values. In addition, the average overall RXE did not differ much between the three nutritional treatment groups. Dosage compensation was not able to be investigated in the sheep placenta as caruncle and cotyledon tissues were difficult to separate completely.

We observed low dosage compensation in the male specific testes and high dosage compensation in the epididymis. This is consistent with the pattern of low dosage compensation in Drosophila testes and higher dosage compensation in the germline as well as with the low X:A ratio in mouse, rat, and human testes and spermatids[26]. Brain was determined to have the highest overall RXE compared to other somatic tissues and this has also been observed in Drosophila brain[62] and in mammals[26], specifically human, mouse, old world monkeys, opossum, platypus, and chicken[4]. Overall, a pattern of incomplete dosage compensation was observed in somatic tissues of day 135 fetal brain, lung, and kidney. Among the treatment groups, we calculated the highest average X: autosome expression ratio of -0.08 for brain, followed by -0.09 for kidney, and -0.11 for lung. The average RXE for all treatment groups ranged from -0.11 to -0.08 respectively.

The most highly expressed X-linked genes in fetal day 135 brain, kidney, and lung were found to differ in expression between males and females in the control group. Thymosin beta 4, X-linked (TMSB4X), ribosomal protein L10 (RPL10), ribosomal protein L39 (RPL39), and ribosomal protein S4, X-linked (RPS4X) were common in brain, kidney, and lung. TMSB4X, RPL10, and RPS4X had the highest expression in the lung while RPL39 had the highest expression in both the lung and kidney.

Our gene ontology results were consistent with other studies in sheep. A study of SNPs on the X chromosome in sheep important to artificial selection found gene ontology related to molecular functions, cellular components, and biological processes[28]. X chromosome genes found in sheep are linked to immune function, which has also been found in selection signatures in pigs[165]. Out of the combined list of 513 expressed X-linked genes in fetal day 135 brain, kidney, and lung, 14 genes reside in the pseudoautosomal region. These genes had low expression ranging from 1 TPM to 50 TPM. X-linked genes in the pseudoautosomal region are known as dosage-insensitive genes because they have a homologous gene on Y chromosome. Our study of X chromosome dosage compensation reveals incomplete dosage compensation in sheep somatic tissues. Potential future work to improve the current study includes an increased sample size and additional nutritional factors that can influence epigenetic mechanisms such as altering the protein content in the pregnant ewe’s diet. Future studies into dosage compensation of other tissues, X-linked genes in the ruminant pseudoautosomal region and X-specific regions are warranted.

Table 1. Mean number of expressed X-linked genes in control, restricted, and overfed day 135 fetal tissues

	Treatment			
	Control	Restricted	Overfed	P-value
Brain	459±3.8	453.5±2.2	455.3±3.2	0.66
Kidney	442.7±2.7	444.3±1.5	439.5±2.7	0.37
Lung	429.3±1.0	427.3±0.6	427±3.7	0.66

Table 2. Enrichment analysis of gene ontology (GO) terms for X-linked genes

GO terms	Count	% of Genes	P-value	Genes
negative regulation of microtubule depolymerization	4	0.8	7.43E-03	FGF13, MID1IP1, MID1, HDAC6
synapse	7	1.4	9.70E-03	GABRE, SLC9A6, GABRA3, NLGN4X, CASK, GLRA4, GABRQ
siRNA binding	3	0.6	1.99E-02	FMR1, MECP2, TLR7
sulfuric ester hydrolase activity	3	0.6	2.51E-02	STS, ARSE, IDS
cristae formation	3	0.6	2.76E-02	APOOL, TAZ, APOO
nucleoside metabolic process	3	0.6	2.76E-02	UPRT, PRPS2, PRPS1
intracellular	27	5.3	3.02E-02	SAT1, AVPR2, ZNF81, ASB11, ZNF75D, NXT2, AGTR2, ZNF182, PAK3, ERAS, DOCK11, GDI1, MCF2, ARHGEF6, BMX, MID1, RAB33A, ASB9, ZNF157, RPS6KA3, CD40LG, ARAF, SYTL5, NRK, HEPH, ARL13A, ZNF41
oxidoreductase activity	7	1.4	4.16E-02	GDI1, CYBB, MAOA, MAOB, F8, CHM, DHRSX
chromatin binding	13	2.5	4.28E-02	AR, MED12, POLA1, MECP2, HMGN5, RBMX, CITED1, ARX, ATRX, NONO, SMC1A, TBL1X, PHF8
receptor activity	5	1.0	5.52E-02	IGSF1, ATP6AP2, MED12, AMOT, MED14
positive regulation of exocytosis	3	0.6	5.54E-02	RAB9A, ATP6AP1, SYTL4
positive regulation of synapse assembly	5	1.0	5.72E-02	SLITRK2, SLITRK4, SRPX2, MECP2, NLGN3
ribonucleoside monophosphate biosynthetic process	2	0.4	5.86E-02	PRPS2, PRPS1
nucleotide binding	13	2.5	6.05E-02	RBM41, CSTF2, RBM3, PABPC5, POLA1, RBMX, NONO, ATP7A, HNRNPH2, UPF3B, HTATSF1, ZRSR2, RBM10
positive regulation of interleukin-4 production	3	0.6	6.34E-02	CD40LG, FOXP3, SASH3
catalytic activity	6	1.2	6.91E-02	PHKA2, SYN1, PHKA1, TKTL1, PCYT1B, ACSL4
growth cone	4	0.8	6.95E-02	FRMD7, USP9X, FMR1, FGF13
axon extension	3	0.6	7.18E-02	SLC9A6, USP9X, DCX
cytokine receptor activity	3	0.6	7.35E-02	IL2RG, IL13RA1, IL13RA2
zinc ion binding	33	6.4	7.96E-02	ZMAT1, APEX2, XIAP, FHL1, CA5B, RLIM, GATA1, USP27X, JADE3, MORC4, DMD, ZNF185, ZDHHC9, RNF128, RGN, LONRF3, KDM5C, AR, ZMYM3, SUV39H1, MID1, TAB3, MID2, ZDHHC15, PJA1, TEX13B, DRP2, PRICKLE3, ITGB1BP2, PHF8, RNF113A, HDAC6, PHF6
Rho guanyl-nucleotide exchange factor activity	5	1.0	8.06E-02	FGD1, MCF2, ARHGEF6, ARHGEF9, DOCK11
protein serine/threonine kinase activity	10	1.9	8.08E-02	SRPK3, IRAK1, RPS6KA6, RPS6KA3, PAK3, PDK3, ARAF, WNK3, PIM2, CDK16
signal transduction	10	1.9	8.41E-02	ARHGAP4, ARHGAP6, IGSF1, TENM1, STARD8, OPHN1, ARHGAP36, IL1RAPL2, OCRL, IL1RAPL1
regulation of RNA splicing	3	0.6	8.94E-02	PQBP1, AFF2, MBNL3
positive regulation of NF-kappaB transcription factor activity	6	1.2	8.95E-02	IRAK1, AR, CD40LG, IKBKG, EDA, MID2
GABA-A receptor activity	3	0.6	9.03E-02	GABRE, GABRA3, GABRQ
cell junction	7	1.4	9.58E-02	GABRE, PLXNA3, GABRA3, GLRA2, GRIA3, GLRA4, GABRQ

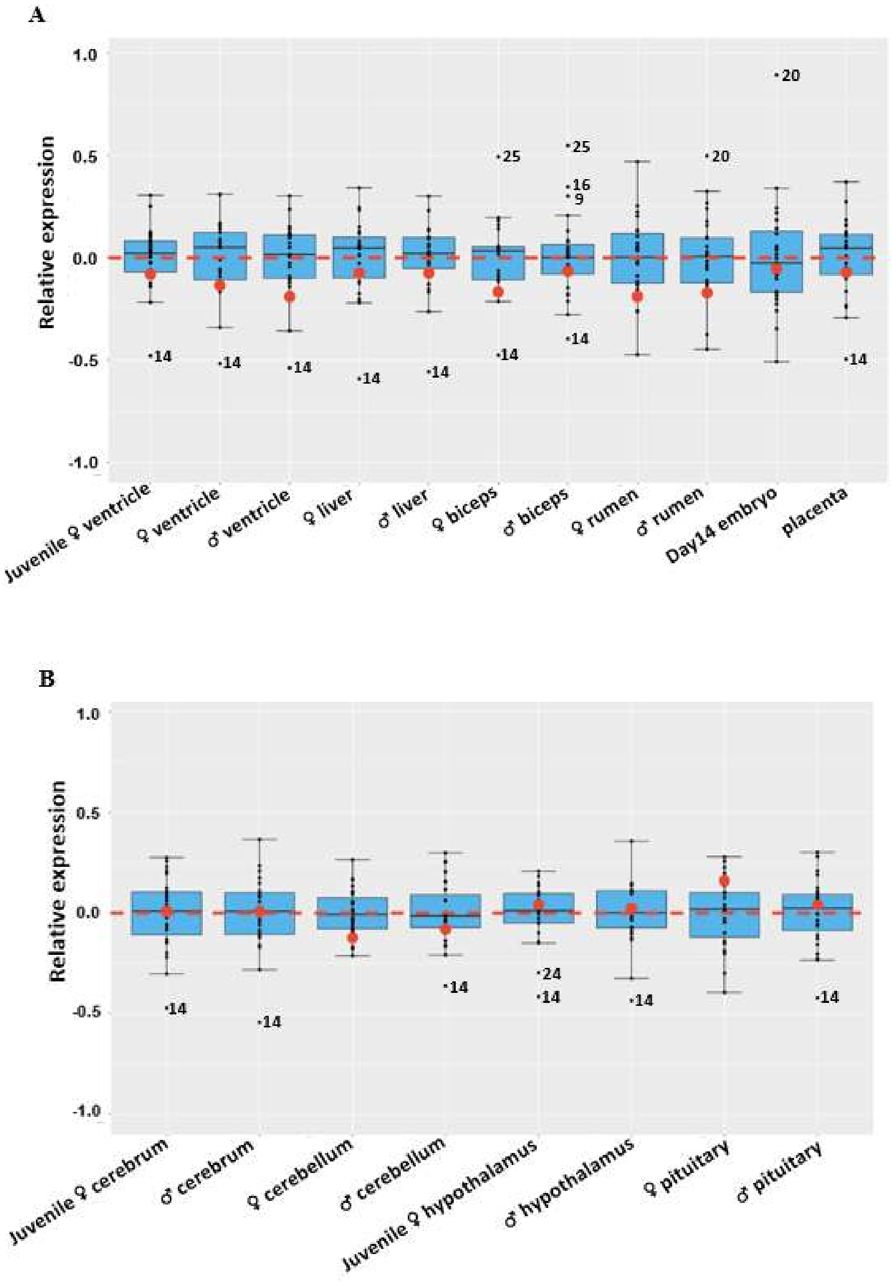


Figure 1 Boxplots of $\log_2$-transformed relative X chromosome expression (RXE) data in brain and major tissues. (A) RXE Boxplots of $\log_2$-transformed TPM values for major ovine tissues. **(B)** RXE Boxplots of $\log_2$-transformed TPM values for ovine brain tissues. Red dots represent the mean X chromosome expression for all replicate libraries within a treatment group. Black dots represent the mean expression for each autosome. The red dotted line represents complete dosage compensation. Equal X expression with the autosomal chromosomes is indicated by an RXE value of 0, while halved X expression relative to the other chromosomes is indicated by an RXE value of -1. The autosomes with a mean expression that falls outside of the quartiles of the boxplot are numbered.

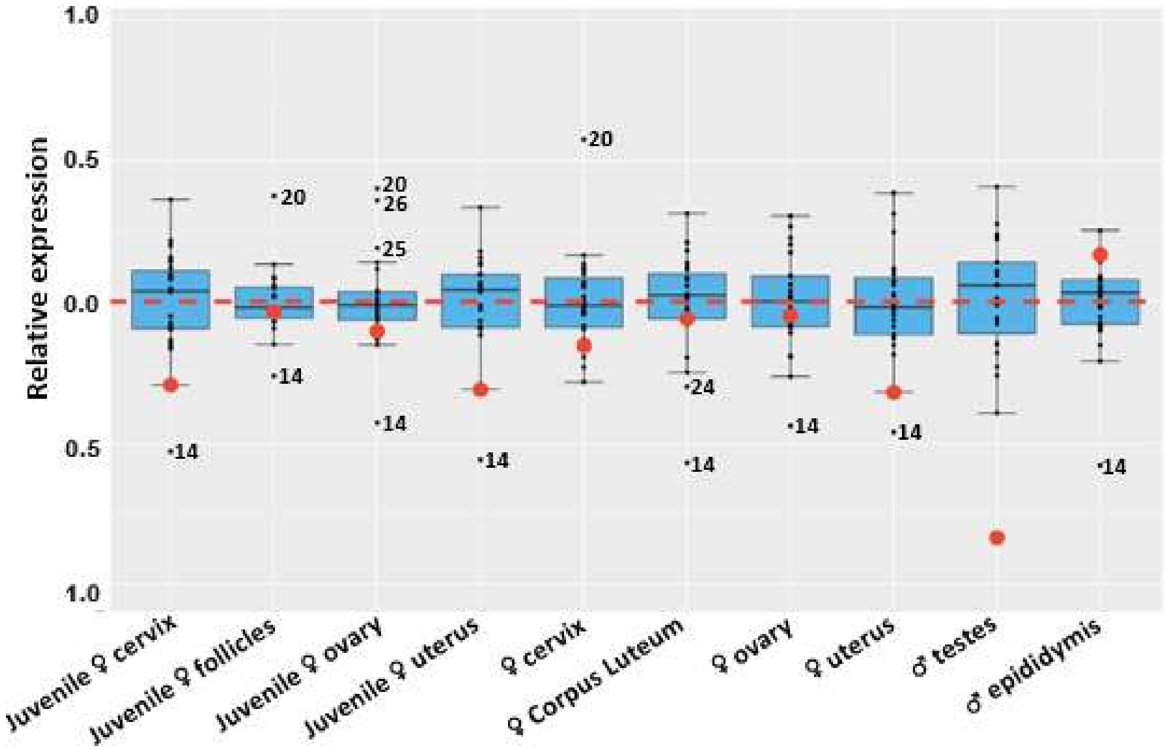


Figure 2 Boxplots of log$_2$-transformed relative X chromosome expression (RXE) data in female and male specific tissues. RXE Boxplots of log$_2$-transformed TPM values for ovine female specific and male specific tissues. Red dots represent the mean X chromosome expression for all replicate libraries within a treatment group. Black dots represent the mean expression for each autosome. The red dotted line represents complete dosage compensation. Equal X expression with the autosomal chromosomes is indicated by an RXE value of 0, while halved X expression relative to the other chromosomes is indicated by an RXE value of -1. The autosomes with a mean expression that falls outside of the quartiles of the boxplot are numbered.

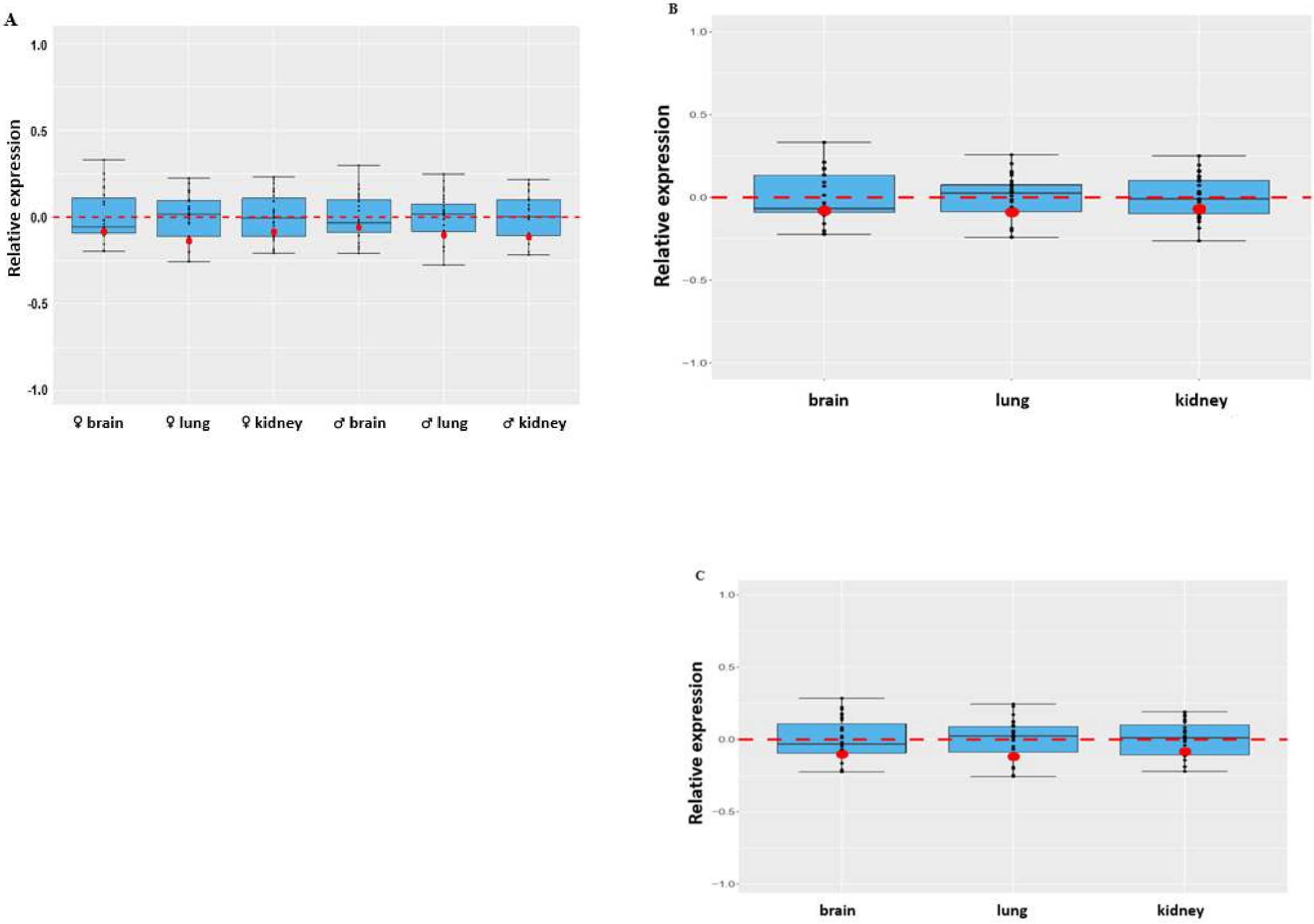


Figure 3 Boxplots of log$_2$-transformed relative X chromosome expression (RXE) data by nutritional treatment group. (A) Boxplots of log$_2$-transformed TPM values displaying the mean X chromosome gene expression relative to the mean gene expression of all autosome for the Control treatment group. **(B)** RXE Boxplots for the Overfed treatment group. **(C)** RXE Boxplots for the Restricted treatment group. Red dots represent the mean X chromosome expression for all replicate libraries within a treatment group. Black dots represent the mean expression for each autosome. The red dotted line represents complete dosage compensation. Equal X expression with the autosomal chromosomes is indicated by an RXE value of 0, while halved X expression relative to the other chromosomes is indicated by an RXE value of -1. The autosomes with a mean expression that falls outside of the quartiles of the boxplot are numbered.

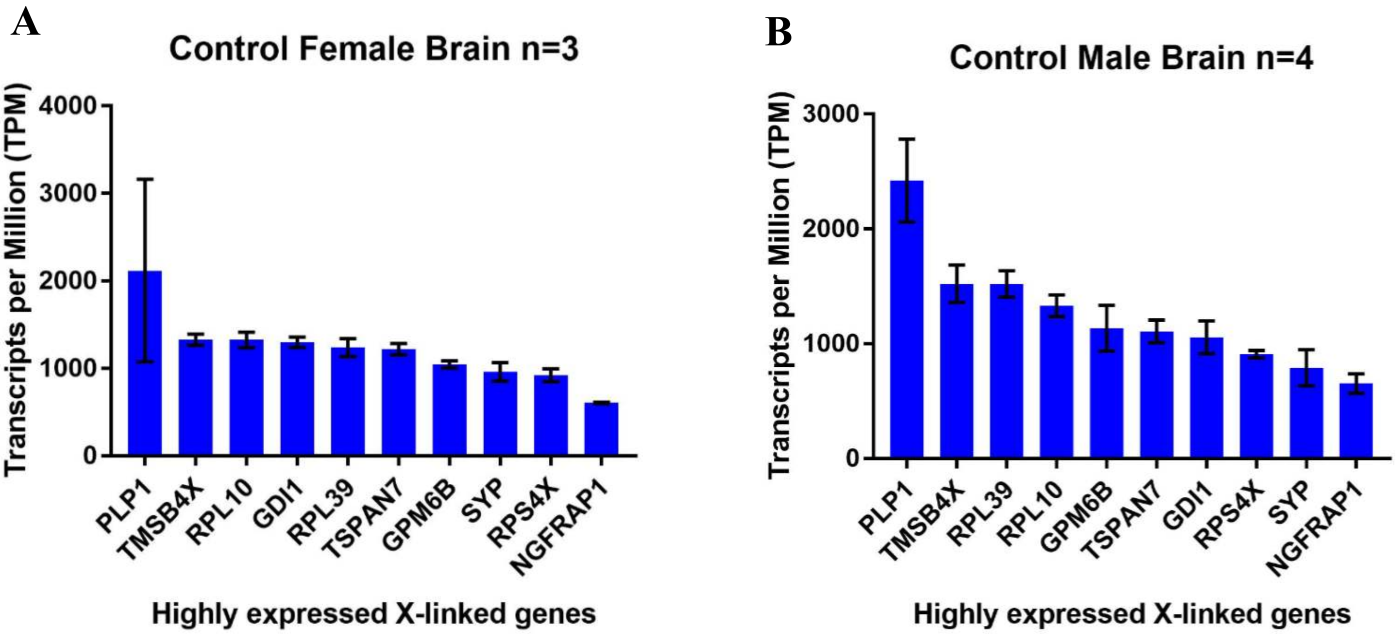


Figure 4. Top 10 X-linked genes expressed in control female and male fetal day 135 brain. (A) Top 10 X-linked genes expressed in control female brain. **(B)** Top 10 X-linked genes expressed in control male brain. Mean expression in TPM is plotted along with error bars for the standard error of the mean. The genes represented are PLP1 (Proteolipid Protein 1), TMSB4X (Thymosin Beta 4, X-Linked), RPL10 (Ribosomal Protein L10), GDI1 (GDP Dissociation Inhibitor 1), RPL39 (Ribosomal Protein L39), TSPAN7 (Tetraspanin 7), GPM6B (Glycoprotein M6B), SYP (Synaptophysin), RPS4X (Ribosomal Protein S4, X-Linked), and NGFRAP1 (Nerve Growth Factor Receptor-Associated Protein 1) aka BEX3 (Brain Expressed X-Linked 3).

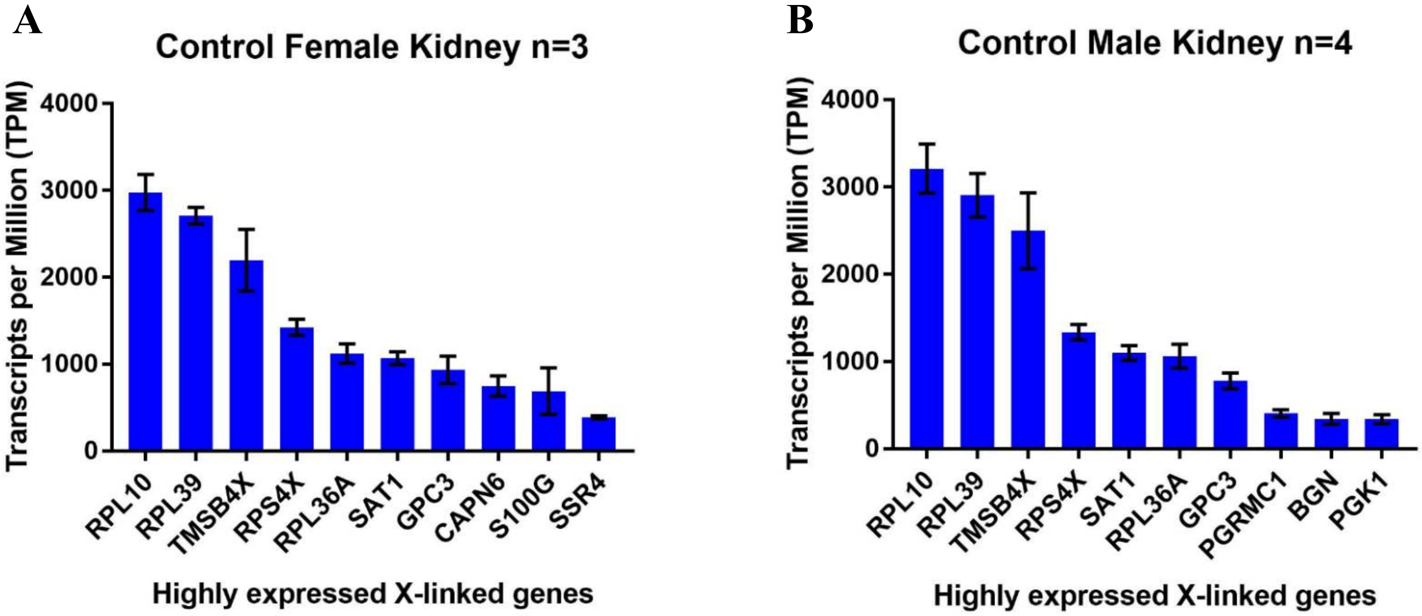


Figure 5. Top 10 X-linked genes expressed in control female and male fetal day 135 kidney. (A) Top 10 X-linked genes expressed in control female kidney. **(B)** Top 10 X-linked genes expressed in control male kidney. Mean expression in TPM is plotted along with error bars for the standard error of the mean. The genes represented are RPL10 (Ribosomal Protein L10), RPL39 (Ribosomal Protein L39), TMSB4X (Thymosin Beta 4, X-Linked), RPS4X (Ribosomal Protein S4, X-Linked), RPL36A (Ribosomal Protein L36a), SAT1 (Spermidine/Spermine N1-Acetyltransferase 1), GPC3 (Glypican 3), CAPN6 (Calpain 6), S100G (S100 Calcium Binding Protein G), SSR4 (Signal Sequence Receptor Subunit 4), PGRMC1 (Progesterone Receptor Membrane Component 1), BGN (Biglycan), and PGK1 (Phosphoglycerate Kinase 1).

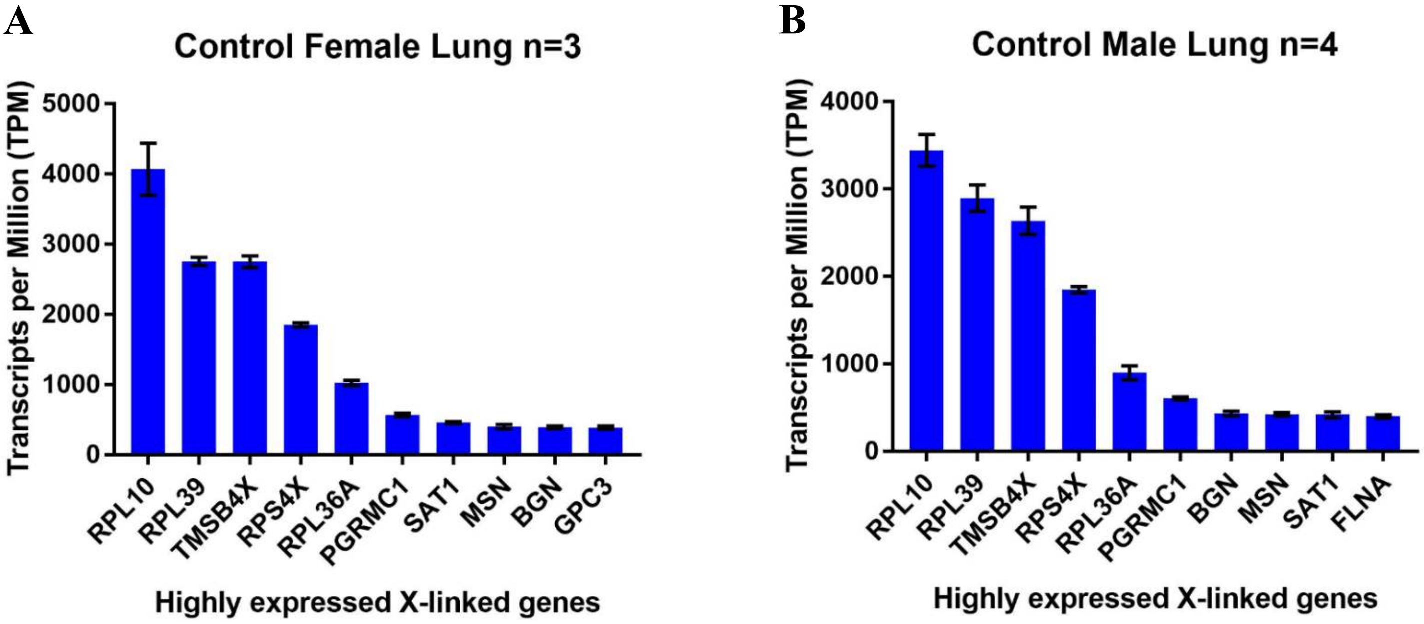


Figure 6. Top 10 X-linked genes expressed in control female and male fetal day 135 lung. (A) Top 10 X-linked genes expressed in control female lung. **(B)** Top 10 X-linked genes expressed in control male lung. Mean expression in TPM is plotted along with error bars for the standard error of the mean. The genes represented are RPL10 (Ribosomal Protein L10), RPL39 (Ribosomal Protein L39), TMSB4X (Thymosin Beta 4, X-Linked), RPS4X (Ribosomal Protein S4, X-Linked), RPL36A (Ribosomal Protein L36a), PGRMC1 (Progesterone Receptor Membrane Component 1), SAT1 (Spermidine/Spermine N1-Acetyltransferase 1), MSN (Moesin), BGN (Biglycan), GPC3 (Glypican 3), and FLNA (Filamin A).

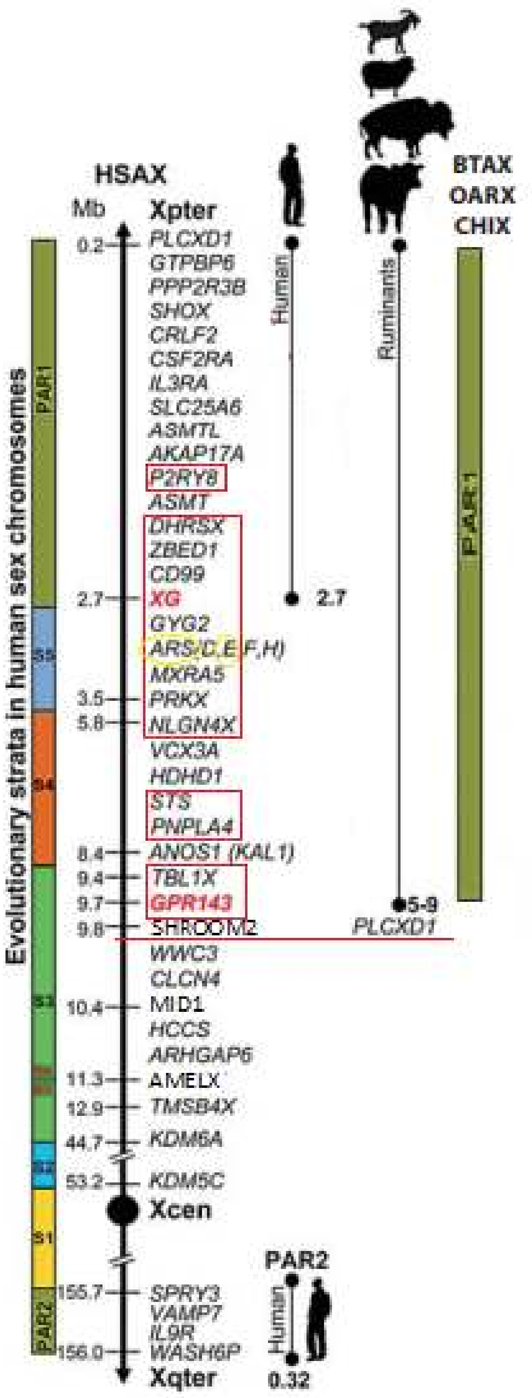


Figure 7 Expressed X-linked genes in sheep day 135 brain, kidney, and lung in the ruminant pseudoautosomal region. The Human X chromosome (HSAX) shows the organization of the human pseudoautosomal regions (PAR1 and PAR2) and X-specific regions evolutionary strata (S1-S5). Pseudoautosomal boundaries are shown in pink text. Human PAR1 is 2.7 Mb starting at the gene PLCXD1 and ending at the gene XG and human PAR2 is 0.32 Mb. The ruminant PAR starts at the gene GTPB6P and ends at the gene GPR143. The ruminant PAR is 5-9 Mb. The gene PLCXD1 is X-specific in ruminants and is not located in the pseudoautosomal region. Its location is marked by the red line. The 14 expressed X-linked genes in this study that are in the ruminant PAR are enclosed in red boxes. The genes are P2RY8, DHRSX, ZBED1, CD99, XG, GYG2, ARSE, MXRA5, PRKX, NLGN4X, STS, PNPLA4, TBL1X, AND GPR143.

3.7 REFERENCES

1. Ma Q, Liu X, Pan J, et al. Genome-wide detection of copy number variation in chinese indigenous sheep using an ovine high-density 600 K SNP array. *Sci Rep*. 2017;7(1):912-017-00847-9. doi: 10.1038/s41598-017-00847-9 [doi].

2. Jue NK, Murphy MB, Kasowitz SD, et al. Determination of dosage compensation of the mammalian X chromosome by RNA-seq is dependent on analytical approach. *BMC Genomics*. 2013;14:150-2164-14-150. doi: 10.1186/1471-2164-14-150 [doi].

3. Pessia E, Makino T, Bailly-Bechet M, McLysaght A, Marais GA. Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. *Proc Natl Acad Sci U S A*. 2012;109(14):5346-5351. doi: 10.1073/pnas.1116763109 [doi].

4. Julien P, Brawand D, Soumillon M, et al. Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. *PLoS Biol*. 2012;10(5):e1001328.

5. Deakin JE, Hore TA, Koina E, Graves JAM. The status of dosage compensation in the multiple X chromosomes of the platypus. *PLoS Genet*. 2008;4(7):e1000140.

6. Smith CA, Sinclair AH. Sex determination: Insights from the chicken. *Bioessays*. 2004;26(2):120-132.

7. Meyer BJ. Sex in the worm: Counting and compensating X-chromosome dose. *Trends in Genetics*. 2000;16(6):247-253.

8. Gelbart ME, Kuroda MI. Drosophila dosage compensation: A complex voyage to the X chromosome. *Development*. 2009;136(9):1399-1410. doi: 10.1242/dev.029645 [doi].

9. Deng X, Hiatt JB, Ercan S, et al. Evidence for compensatory upregulation of expressed X-linked genes in mammals, caenorhabditis elegans and drosophila melanogaster. *Nat Genet*. 2011;43(12):1179-1185.

10. Birchler JA, Riddle NC, Auger DL, Veitia RA. Dosage balance in gene regulation: Biological implications. *Trends in Genetics*. 2005;21(4):219-226.

11. Veitia RA, Bottani S, Birchler JA. Cellular reactions to gene dosage imbalance: Genomic, transcriptomic and proteomic effects. *Trends in Genetics*. 2008;24(8):390-397.

12. Roper RJ, Reeves RH. Understanding the basis for down syndrome phenotypes. *PLoS Genet*. 2006;2(3):e50.

13. Hassold T, Hunt P. To err (meiotically) is human: The genesis of human aneuploidy. *Nature Reviews Genetics*. 2001;2(4):280-291.

14. Tyler C, Edman JC. Down syndrome, turner syndrome, and klinefelter syndrome: Primary care throughout the life span. *Primary Care: Clinics in Office Practice*. 2004;31(3):627-648.

15. Ohno S. *Sex chromosomes and sex-linked genes.* Vol 1. Springer Science & Business Media; 2013.

16. Mank JE. Sex chromosome dosage compensation: Definitely not for everyone. *Trends in genetics*. 2013;29(12):677-683.

17. Lyon MF. Gene action in the X-chromosome of the mouse (mus musculus L.). . 1961.

18. Veitia RA, Veyrunes F, Bottani S, Birchler JA. X chromosome inactivation and active X upregulation in therian mammals: Facts, questions, and hypotheses. *J Mol Cell Biol*. 2015;7(1):2-11. doi: 10.1093/jmcb/mjv001 [doi].

19. Rens W, Wallduck MS, Lovell FL, Ferguson-Smith MA, Ferguson-Smith AC. Epigenetic modifications on X chromosomes in marsupial and monotreme mammals and implications for evolution of dosage compensation. *Proc Natl Acad Sci U S A*. 2010;107(41):17657-17662. doi: 10.1073/pnas.0910322107 [doi].

20. Mahadevaiah SK, Royo H, VandeBerg JL, McCarrey JR, Mackay S, Turner JM. Key features of the X inactivation process are conserved between marsupials and eutherians. *Current Biology*. 2009;19(17):1478-1484.

21. Okamoto I, Patrat C, Thépot D, et al. Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature*. 2011;472(7343):370-374.

22. Disteche CM. Dosage compensation of the sex chromosomes and autosomes. . 2016.

23. Xiong Y, Chen X, Chen Z, et al. RNA sequencing shows no dosage compensation of the active X-chromosome. *Nat Genet*. 2010;42(12):1043-1047.

24. He X, Chen X, Xiong Y, et al. He et al. reply. *Nat Genet*. 2011;43(12):1171-1172.

25. Lin F, Xing K, Zhang J, He X. Expression reduction in mammalian X chromosome evolution refutes ohno's hypothesis of dosage compensation. *Proc Natl Acad Sci U S A*. 2012;109(29):11752-11757. doi: 10.1073/pnas.1201816109 [doi].

26. Disteche CM. Dosage compensation of the active X chromosome in mammals. *Nat Genet*. 2006;38(1):47-53.

27. Kharchenko PV, Xi R, Park PJ. Evidence for dosage compensation between the X chromosome and autosomes in mammals. *Nat Genet*. 2011;43(12):1167-1169.

28. Zhu C, Fan H, Yuan Z, et al. Detection of selection signatures on the X chromosome in three sheep breeds. *International journal of molecular sciences*. 2015;16(9):20360-20374.

29. Ohno S. Patterns in genome evolution. *Curr Opin Genet Dev*. 1993;3(6):911-914.

30. Palmer S, Perry J, Ashworth A. A contravention of ohno's law in mice. *Nat Genet*. 1995;10(4):472-476.

31. Barry JS, Anthony RV. The pregnant sheep as a model for human pregnancy. *Theriogenology*. 2008;69(1):55-67.

32. Morrison JL. Sheep models of intrauterine growth restriction: Fetal adaptations and consequences. *Clinical and Experimental Pharmacology and Physiology*. 2008;35(7):730-743.

33. Peñagaricano F, Wang X, Rosa GJ, Radunz AE, Khatib H. Maternal nutrition induces gene expression changes in fetal muscle and adipose tissues in sheep. *BMC Genomics*. 2014;15(1):1.

34. Rose G, Mulder H, Thompson A, van der Werf J, van Arendonk J. Breeding objectives for sheep should be customised depending on variation in pasture growth across years. *animal*. 2015;9(08):1268-1277.

35. Godfrey KM, Barker DJ. Fetal nutrition and adult disease. *Am J Clin Nutr*. 2000;71(5 Suppl):1344S-52S.

36. Gupta V, Parisi M, Sturgill D, et al. Global analysis of X-chromosome dosage compensation. *Journal of biology*. 2006;5(1):1.

37. Lucchesi JC, Kelly WG, Panning B. Chromatin remodeling in dosage compensation. *Annu Rev Genet*. 2005;39:615-651.

38. Itoh Y, Melamed E, Yang X, et al. Dosage compensation is less effective in birds than in mammals. *Journal of Biology*. 2007;6(1):1.

39. Ellegren H, Hultin-Rosenberg L, Brunström B, Dencker L, Kultima K, Scholz B. Faced with inequality: Chicken do not have a general dosage compensation of sex-linked genes. *BMC biology*. 2007;5(1):1.

40. Knouse KA, Wu J, Whittaker CA, Amon A. Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proc Natl Acad Sci U S A*. 2014;111(37):13409-13414. doi: 10.1073/pnas.1415287111 [doi].

41. Prothero KE, Stahl JM, Carrel L. Dosage compensation and gene expression on the mammalian X chromosome: One plus one does not always equal two. *Chromosome research*. 2009;17(5):637-648.

42. Mori MA, Lapunzina P, Delicado A, et al. A prenatally diagnosed patient with full monosomy 21: Ultrasound, cytogenetic, clinical, molecular, and necropsy findings. *American Journal of Medical Genetics Part A*. 2004;127(1):69-73.

43. Graves JAM. Sex chromosome specialization and degeneration in mammals. *Cell*. 2006;124(5):901-914.

44. Filatov DA, Gerrard DT. High mutation rates in human and ape pseudoautosomal genes. *Gene*. 2003;317:67-77.

45. Galtier N. Recombination, GC-content and the human pseudoautosomal boundary paradox. *Trends in Genetics*. 2004;20(8):347-349.

46. Das PJ, Chowdhary BP, Raudsepp T. Characterization of the bovine pseudoautosomal region and comparison with sheep, goat, and other mammalian pseudoautosomal regions. *Cytogenet Genome Res*. 2009;126(1-2):139-147. doi: 10.1159/000245913 [doi].

47. Ercan S. Mechanisms of x chromosome dosage compensation. *Journal of genomics*. 2015;3:1.

48. Meyer BJ. Targeting X chromosomes for repression. *Curr Opin Genet Dev*. 2010;20(2):179-189.

49. Ercan S, Giresi PG, Whittle CM, Zhang X, Green RD, Lieb JD. X chromosome repression by localization of the C. elegans dosage compensation machinery to sites of transcription initiation. *Nat Genet*. 2007;39(3):403-408.

50. Graves JAM. Evolution of vertebrate sex chromosomes and dosage compensation. *Nature Reviews Genetics*. 2015.

51. Lin F, Xing K, Zhang J, He X. Expression reduction in mammalian X chromosome evolution refutes ohno's hypothesis of dosage compensation. *Proc Natl Acad Sci U S A*. 2012;109(29):11752-11757. doi: 10.1073/pnas.1201816109 [doi].

52. Davis TL, Meyer BJ. SDC-3 coordinates the assembly of a dosage compensation complex on the nematode X chromosome. *Development*. 1997;124(5):1019-1031.

53. Villeneuve AM, Meyer BJ. The role of sdc-1 in the sex determination and dosage compensation decisions in caenorhabditis elegans. *Genetics*. 1990;124(1):91-114.

54. Plenefisch JD, DeLong L, Meyer BJ. Genes that implement the hermaphrodite mode of dosage compensation in caenorhabditis elegans. *Genetics*. 1989;121(1):57-76.

55. Belote JM, Lucchesi JC. Male-specific lethal mutations of drosophila melanogaster. *Genetics*. 1980;96(1):165-186.

56. Takagi N, Abe K. Detrimental effects of two active X chromosomes on early mouse development. *Development*. 1990;109(1):189-201.

57. Okamoto I, Heard E. Lessons from comparative analysis of X-chromosome inactivation in mammals. *Chromosome research*. 2009;17(5):659-669.

58. Wang J, Yu R, Shete S. X-chromosome genetic association test accounting for x-inactivation, skewed x-inactivation, and escape from x-inactivation. *Genet Epidemiol*. 2014;38(6):483-493.

59. Escamilla-Del-Arenal M, da Rocha ST, Heard E. Evolutionary diversity and developmental regulation of X-chromosome inactivation. *Hum Genet*. 2011;130(2):307-327.

60. Lingenfelter PA, Adler DA, Poslinski D, et al. Escape from X inactivation of smcx is preceded by silencing during mouse development. *Nat Genet*. 1998;18(3):212-213.

61. Deng X, Berletch JB, Nguyen DK, Disteche CM. X chromosome regulation: Diverse patterns in development, tissues and disease. *Nat Rev Genet*. 2014;15(6):367-378. doi: 10.1038/nrg3687 [doi].

62. Huylmans AK, Parsch J. Variation in the X:Autosome distribution of male-biased genes among drosophila melanogaster tissues and its relationship with dosage compensation. *Genome Biol Evol*. 2015;7(7):1960-1971. doi: 10.1093/gbe/evv117 [doi].

63. Zechner U, Wilda M, Kehrer-Sawatzki H, Vogel W, Fundele R, Hameister H. A high density of X-linked genes for general cognitive ability: A run-away process shaping human evolution? *TRENDS in Genetics*. 2001;17(12):697-701.

64. Gregg C, Zhang J, Butler JE, Haig D, Dulac C. Sex-specific parent-of-origin allelic expression in the mouse brain. *Science*. 2010;329(5992):682-685. doi: 10.1126/science.1190831 [doi].

65. Wang PJ, McCarrey JR, Yang F, Page DC. An abundance of X-linked genes expressed in spermatogonia. *Nat Genet*. 2001;27(4):422-426.

66. Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JM. The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nat Genet*. 2008;40(6):794-799.

67. Zhang YE, Vibranovski MD, Landback P, Marais GA, Long M. Chromosomal redistribution of male-biased genes in mammalian evolution with two bursts of gene gain on the X chromosome. *PLoS Biol*. 2010;8(10):e1000494.

68. Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet*. 2004;36(6):642-646.

69. Talebizadeh Z, Simon SD, Butler MG. X chromosome gene expression in human tissues: Male and female comparisons. *Genomics*. 2006;88(6):675-681.

70. Goldammer T, Brunner RM, Rebl A, et al. A high-resolution radiation hybrid map of sheep chromosome X and comparison with human and cattle. *Cytogenet Genome Res*. 2009;125(1):40-45. doi: 10.1159/000207520 [doi].

71. Das PJ, Chowdhary BP, Raudsepp T. Characterization of the bovine pseudoautosomal region and comparison with sheep, goat, and other mammalian pseudoautosomal regions. *Cytogenet Genome Res*. 2009;126(1-2):139-147. doi: 10.1159/000245913 [doi].

72. Raudsepp T, Chowdhary BP. The eutherian pseudoautosomal region. *Cytogenet Genome Res*. 2015;147(2-3):81-94. doi: 10.1159/000443157 [doi].

73. Van Laere AS, Coppieters W, Georges M. Characterization of the bovine pseudoautosomal boundary: Documenting the evolutionary history of mammalian sex chromosomes. *Genome Res*. 2008;18(12):1884-1895. doi: 10.1101/gr.082487.108 [doi].

74. Hassanin A, Ropiquet A. Molecular phylogeny of the tribe bovini (bovidae, bovinae) and the taxonomic status of the kouprey, bos sauveli urbain 1937. *Mol Phylogenet Evol*. 2004;33(3):896-907.

75. Galloway SM, McNatty KP, Cambridge LM, et al. Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nat Genet*. 2000;25(3):279-283.

76. Davis GH, Dodds KG, Wheeler R, Jay NP. Evidence that an imprinted gene on the X chromosome increases ovulation rate in sheep. *Biol Reprod*. 2001;64(1):216-221.

77. Brockdorff N, Turner BM. Dosage compensation in mammals. *Epigenetics*. 2007:321-340.

78. Ozsolak F, Milos PM. RNA sequencing: Advances, challenges and opportunities. *Nature reviews genetics*. 2011;12(2):87-98.

79. Wolf JB, Bryk J. General lack of global dosage compensation in ZZ/ZW systems? broadening the perspective with RNA-seq. *BMC Genomics*. 2011;12(1):1.

80. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res*. 2008;18(9):1509-1517. doi: 10.1101/gr.079558.108 [doi].

81. Smith AM, Heisler LE, Mellor J, et al. Quantitative phenotyping via deep barcode sequencing. *Genome Res*. 2009;19(10):1836-1842. doi: 10.1101/gr.093955.109 [doi].

82. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet*. 2008;40(12):1413-1415.

83. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nature methods*. 2008;5(7):621-628.

84. Hillier LW, Reinke V, Green P, Hirst M, Marra MA, Waterston RH. Massively parallel sequencing of the polyadenylated transcriptome of C. elegans. *Genome Res*. 2009;19(4):657-666. doi: 10.1101/gr.088112.108 [doi].

85. Brawand D, Soumillon M, Necsulea A, et al. The evolution of gene expression levels in mammalian organs. *Nature*. 2011;478(7369):343.

86. Kharchenko PV, Xi R, Park PJ. Evidence for dosage compensation between the X chromosome and autosomes in mammals. *Nat Genet*. 2011;43(12):1167-1169.

87. Lin H, Halsall JA, Antczak P, O'Neill LP, Falciani F, Turner BM. Relative overexpression of X-linked genes in mouse embryonic stem cells is consistent with ohno's hypothesis. *Nat Genet*. 2011;43(12):1169-1170.

88. Albritton SE, Kranz AL, Rao P, Kramer M, Dieterich C, Ercan S. Sex-biased gene expression and evolution of the x chromosome in nematodes. *Genetics*. 2014;197(3):865-883. doi: 10.1534/genetics.114.163311 [doi].

89. Zhao S, Zhang B. A comprehensive evaluation of ensembl, RefSeq, and UCSC annotations in the context of RNA-seq read mapping and gene quantification. *BMC Genomics*. 2015;16(1):1.

90. Tarazona S, Garcia-Alcalde F, Dopazo J, Ferrer A, Conesa A. Differential expression in RNA-seq: A matter of depth. *Genome Res*. 2011;21(12):2213-2223. doi: 10.1101/gr.124321.111 [doi].

91. Barr ML, Bertram EG. A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. In: *Problems of birth defects.* Springer; 1949:101-102.

92. Kobayashi S. Live imaging of X chromosome inactivation and reactivation dynamics. *Dev Growth Differ*. 2017.

93. Avner P, Heard E. X-chromosome inactivation: Counting, choice and initiation. *Nature Reviews Genetics*. 2001;2(1):59-67.

94. de Mello, Joana Carvalho Moreira, de Araujo, Erica Sara Souza, Stabellini R, et al. Random X inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome. *PloS one*. 2010;5(6):e10947.

95. Wang X, Miller DC, Clark AG, Antczak DF. Random X inactivation in the mule and horse placenta. *Genome Res*. 2012;22(10):1855-1863. doi: 10.1101/gr.138487.112 [doi].

96. Soma M, Fujihara Y, Okabe M, Ishino F, Kobayashi S. Ftx is dispensable for imprinted X-chromosome inactivation in preimplantation mouse embryos. *Scientific reports*. 2014;4:5181.

97. Xue F, Tian XC, Du F, et al. Aberrant patterns of X chromosome inactivation in bovine clones. *Nat Genet*. 2002;31(2):216-220.

98. Dindot SV, Kent KC, Evers B, Loskutoff N, Womack J, Piedrahita JA. Conservation of genomic imprinting at the XIST, IGF2, and GTL2 loci in the bovine. *Mammalian genome*. 2004;15(12):966-974.

99. Wake N, Takagi N, Sasaki M. Non-random inactivation of X chromosome in the rat yolk sac. . 1976.

100. Latham KE. X chromosome imprinting and inactivation in preimplantation mammalian embryos. *TRENDS in Genetics*. 2005;21(2):120-127.

101. Huynh KD, Lee JT. Imprinted X inactivation in eutherians: A model of gametic execution and zygotic relaxation. *Curr Opin Cell Biol*. 2001;13(6):690-697.

102. Huynh KD, Lee JT. X-chromosome inactivation: A hypothesis linking ontogeny and phylogeny. *Nature Reviews Genetics*. 2005;6(5):410-418.

103. Ishido N, Inoue N, Watanabe M, Hidaka Y, Iwatani Y. The relationship between skewed X chromosome inactivation and the prognosis of graves' and hashimoto's diseases. *Thyroid*. 2015;25(2):256-261.

104. van den Berg, Ilse M, Laven JS, Stevens M, et al. X chromosome inactivation is initiated in human preimplantation embryos. *The American Journal of Human Genetics*. 2009;84(6):771-779.

105. Harper MI, Fosten M, Monk M. Preferential paternal X inactivation in extraembryonic tissues of early mouse embryos. *J Embryol Exp Morphol*. 1982;67:127-135.

106. Takagi N, Sasaki M. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. . 1975.

107. Merzouk S, Deuve JL, Dubois A, Navarro P, Avner P, Morey C. Lineage-specific regulation of imprinted X inactivation in extraembryonic endoderm stem cells. *Epigenetics & chromatin*. 2014;7(1):1.

108. Mak W, Nesterova TB, de Napoles M, et al. Reactivation of the paternal X chromosome in early mouse embryos. *Science*. 2004;303(5658):666-669. doi: 10.1126/science.1092674 [doi].

109. Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science*. 2004;303(5658):644-649. doi: 10.1126/science.1092727 [doi].

110. De La Fuente R, Hahnel A, Basrur PK, King WA. X inactive-specific transcript (xist) expression and X chromosome inactivation in the preattachment bovine embryo. *Biol Reprod*. 1999;60(3):769-775.

111. Brunner HG, Nelen M, Breakefield X, Ropers H, Van Oost B. Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *SCIENCE-NEW YORK THEN WASHINGTON-*. 1993;262:578-578.

112. Cases O, Seif I, Grimsby J, et al. Aggressive behavior and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA. *Science*. 1995;268(5218):1763-1766.

113. Ferreira A, Machado G, Diesel T, et al. Allele-specific expression of the MAOA gene and X chromosome inactivation in in vitro produced bovine embryos. *Mol Reprod Dev*. 2010;77(7):615-621.

114. Bermejo-Alvarez P, Rizos D, Lonergan P, Gutierrez-Adan A. Transcriptional sexual dimorphism in elongating bovine embryos: Implications for XCI and sex determination genes. *Reproduction*. 2011;141(6):801-808. doi: 10.1530/REP-11-0006 [doi].

115. Itoh Y, Arnold AP. X chromosome regulation of autosomal gene expression in bovine blastocysts. *Chromosoma*. 2014;123(5):481-489.

116. Chen Z, Hagen DE, Wang J, et al. Global assessment of imprinted gene expression in the bovine conceptus by next generation sequencing. *Epigenetics*. 2016(just-accepted):00-00.

117. Luciani J, Bézard J, Devictor-Vuillet M, Mauleon P. 3H-thymidine labelling pattern of preleptotene chromosome condensation stages in the foetal sheep ovary. . 1979;19(4B):1241-1250.

118. Bernardi M, Cotinot C, Payen E, Delouis C. Transcription of y-and x-linked genes in preimplantation ovine embryos. *Mol Reprod Dev*. 1996;45(2):132-138.

119. Zwingman T, Fujimoto H, Lai L, et al. Transcription of circular and noncircular forms of sry in mouse testes. *Mol Reprod Dev*. 1994;37(4):370-381.

120. Ao A, Erickson RP, Winston RM, Handysude AH. Transcription of paternal Y-linked genes in the human zygote as early as the pronucleate stage. *Zygote*. 1994;2(04):281-287.

121. Cao QP, Gaudette MF, Robinson DH, Crain WR. Expression of the mouse testis-determining gene sry in male preimplantation embryos. *Mol Reprod Dev*. 1995;40(2):196-204.

122. Fiddler M, Abdel-Rahman B, Rappolee DA, Pergament E. Expression of SRY transcripts in preimplantation human embryos. *Am J Med Genet*. 1995;55(1):80-84.

123. Zhao L, Zhao G, Xi H, Liu Y, Wu K, Zhou H. Molecular and DNA methylation analysis of Peg10 and xist gene in sheep. *Mol Biol Rep*. 2011;38(5):3495-3504.

124. Disteche CM, Filippova GN, Tsuchiya KD. Escape from X inactivation. *Cytogenet Genome Res*. 2002;99(1-4):36-43. doi: 71572 [doi].

125. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*. 2005;434(7031):400-404.

126. Pessia E, Engelstädter J, Marais GA. The evolution of X chromosome inactivation in mammals: The demise of ohno's hypothesis? *Cellular and Molecular Life Sciences*. 2014;71(8):1383-1394.

127. Yang F, Babak T, Shendure J, Disteche CM. Global survey of escape from X inactivation by RNA-sequencing in mouse. *Genome Res*. 2010;20(5):614-622. doi: 10.1101/gr.103200.109 [doi].

128. Disteche CM. Escape from X inactivation in human and mouse. *Trends in Genetics*. 1995;11(1):17-22.

129. Yen ZC, Meyer IM, Karalic S, Brown CJ. A cross-species comparison of X-chromosome inactivation in eutheria. *Genomics*. 2007;90(4):453-463.

130. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*. 2005;434(7031):400-404.

131. Wilson ME. Role of placental function in mediating conceptus growth and survival. *J Anim Sci*. 2002;80(E-Suppl_2):E195-E201.

132. Wu G, Bazer F, Wallace J, Spencer T. Board-invited review: Intrauterine growth retardation: Implications for the animal sciences. *J Anim Sci*. 2006;84(9):2316-2337.

133. Cooney CA, Dave AA, Wolff GL. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. *J Nutr*. 2002;132(8 Suppl):2393S-2400S.

134. Waterland R, Travisano M, Tahiliani K, Rached M, Mirza S. Methyl donor supplementation prevents transgenerational amplification of obesity. *Int J Obes*. 2008;32(9):1373-1379.

135. Lan X, Cretney EC, Kropp J, et al. Maternal diet during pregnancy induces gene expression and DNA methylation changes in fetal tissues in sheep. *Frontiers in genetics*. 2013;4:49.

136. Begum G, Stevens A, Smith EB, et al. Epigenetic changes in fetal hypothalamic energy regulating pathways are associated with maternal undernutrition and twinning. *FASEB J*. 2012;26(4):1694-1703. doi: 10.1096/fj.11-198762 [doi].

137. Greenwood P, Thompson A. Consequences of maternal nutrition during pregnancy and of foetal growth for productivity of sheep. *Recent advances in animal nutrition in Australia*. 2007;16:185-196.

138. Wu G, Bazer FW, Cudd TA, Meininger CJ, Spencer TE. Maternal nutrition and fetal development. *J Nutr*. 2004;134(9):2169-2172. doi: 134/9/2169 [pii].

139. Belkacemi L, Nelson DM, Desai M, Ross MG. Maternal undernutrition influences placental-fetal development. *Biol Reprod*. 2010;83(3):325-331. doi: 10.1095/biolreprod.110.084517 [doi].

140. Allen WR, Wilsher S, Turnbull C, et al. Influence of maternal size on placental, fetal and postnatal growth in the horse. I. development in utero. *Reproduction*. 2002;123(3):445-453.

141. Reed SA, Raja JS, Hoffman ML, Zinn SA, Govoni KE. Poor maternal nutrition inhibits muscle development in ovine offspring. *Journal of animal science and biotechnology*. 2014;5(1):1.

142. Jones AK, Gately RE, McFadden KK, Zinn SA, Govoni KE, Reed SA. Transabdominal ultrasound for detection of pregnancy, fetal and placental landmarks, and fetal age before day 45 of gestation in the sheep. *Theriogenology*. 2016;85(5):939-945. e1.

143. Raja J, Hoffman M, Govoni K, Zinn S, Reed S. Restricted maternal nutrition alters myogenic regulatory factor expression in satellite cells of ovine offspring. *animal*. 2016;10(7):1200-1203.

144. Hoffman M, Peck K, Forella M, Fox A, Govoni K, Zinn S. The effects of poor maternal nutrition during gestation on postnatal growth and development of lambs. *J Anim Sci*. 2016;94(2):789-799.

145. Pillai SM, Sereda NH, Hoffman ML, et al. Effects of poor maternal nutrition during gestation on bone development and mesenchymal stem cell activity in offspring. *PloS one*. 2016;11(12):e0168382.

146. Pillai S, Jones A, Hoffman M, et al. Fetal and organ development at gestational days 45, 90, 135 and at birth of lambs exposed to under-or over-nutrition during gestation. *Translational Animal Science*. 2017;1(1):16-25.

147. Hoffman M, Rokosa M, Zinn S, Hoagland T, Govoni K. Poor maternal nutrition during gestation in sheep reduces circulating concentrations of insulin-like growth factor-I and insulin-like growth factor binding protein-3 in offspring. *Domest Anim Endocrinol*. 2014;49:39-48.

148. Vickers MH. Early life nutrition, epigenetics and programming of later life disease. *Nutrients*. 2014;6(6):2165-2178.

149. Lillycrop KA, Phillips ES, Jackson AA, Hanson MA, Burdge GC. Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. *J Nutr*. 2005;135(6):1382-1386. doi: 135/6/1382 [pii].

150. Sinclair KD, Allegrucci C, Singh R, et al. DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. *Proc Natl Acad Sci U S A*. 2007;104(49):19351-19356. doi: 0707258104 [pii].

151. Jiang Y, Xie M, Chen W, et al. The sheep genome illuminates biology of the rumen and lipid metabolism. *Science*. 2014;344(6188):1168-1173. doi: 10.1126/science.1252806 [doi].

152. Brooks KE, Burns GW, Spencer TE. Peroxisome proliferator activator receptor gamma (PPARG) regulates conceptus elongation in sheep. *Biol Reprod*. 2015;92(2):42, 1-13.

153. Knights M, Hoehn T, Lewis P, Inskeep E. Effectiveness of intravaginal progesterone inserts and FSH for inducing synchronized estrus and increasing lambing rate in anestrous ewes. *J Anim Sci*. 2001;79(5):1120-1131.

154. National Research Council (US). Committee on Nutrient Requirements of Small Ruminants. *Nutrient requirements of small ruminants: Sheep, goats, cervids, and new world camelids.* 中国法制出版社; 2007.

155. Kim D, Langmead B, Salzberg SL. HISAT: A fast spliced aligner with low memory requirements. *Nature methods*. 2015;12(4):357-360.

156. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1-13. doi: 10.1093/nar/gkn923 [doi].

157. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*. 2009;4(1):44-57.

158. Rengaraj D, Hwang YS, Liang XH, Deng WB, Yang ZM, Han JY. Comparative expression and regulation of TMSB4X in male reproductive tissues of rats and chickens. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*. 2013;319(10):584-595.

159. Zanni G, Kalscheuer VM, Friedrich A, et al. A novel mutation in RPL10 (ribosomal protein L10) causes X-Linked intellectual disability, cerebellar hypoplasia, and Spondylo-Epiphyseal dysplasia. *Hum Mutat*. 2015;36(12):1155-1158.

160. Goossens K, Van Soom A, Van Poucke M, et al. Identification and expression analysis of genes associated with bovine blastocyst formation. *BMC developmental biology*. 2007;7(1):64.

161. Dave B, Granados S, Mai J, et al. *Identification of tumor initiating genes RPL39 and MLF2 that mediate lung metastasis through nitric oxide signaling and mesenchymal to epithelial transition.* 2013.

162. Jeon BG, Rho GJ, Betts DH, Petrik JJ, Favetta LA, King WA. Low levels of X-inactive specific transcript in somatic cell nuclear transfer embryos derived from female bovine freemartin donor cells. *Sex Dev*. 2012;6(1-3):151-159. doi: 10.1159/000334050 [doi].

163. Duan J, Jue N, Jiang Z, et al. 144 dosage compensation and x-linked gene expression in bovine in vivo and in vitro embryos. *Reproduction, Fertility and Development*. 2016;28(2):202-202.

164. Ka S, Ahn H, Seo M, Kim H, Kim JN, Lee H. Status of dosage compensation of X chromosome in bovine genome. *Genetica*. 2016:1-10.

165. Ma Y, Zhang H, Zhang Q, Ding X. Identification of selection footprints on the X chromosome in pig. *PLoS One*. 2014;9(4):e94911.