UNIVERSITÀ DEGLI STUDI DI MILANO

PhD Course in Veterinary and Animal Science
Class XXIX

# EPIGENETIC INVESTIGATION ON CYCLIC ONSET OF REPRODUCTIVE ACTIVITY IN DAIRY GOATS

(Final Thesis)

PhD Candidate: Stefano Frattini
R10509

Tutor: Prof. Giulio Pagnacco

Academic Year 2016

**INDEX**

**FIGURES**

**TABLES**

**LIST OF ABBREVIATIONS**

CDS - Coding Sequence
CGIs - CpG Islands
CHI - Goat chromosome
DEGs - Differentially Expressed Genes
DIO2 - type 2 Deiodinase Gene
DMGs - Differentially Methylated Genes
DMRs - Differentially Methylated Regions
DNMTs - DNA methyltransferases
ELISA - Enzyme-Linked ImmunoSorbent Assay
EWAS - Epigenome-Wide Association Studies
FDR - False Discovery Rate
FL - flanking region
FSH - Follicle-Stimulating Hormone
GnRH - Gonadotropin-Releasing Hormone
HMRs - Highly methylated regions
HPG axis - Hypothalamic-Pituitary-Gonadal axis
LD - Long Days
LH - Luteinizing Hormone
MBD-seq - Methylated DNA Binding Domain sequencing
*MTNR1A* - Melatonin Receptor 1a
P4 - Progesterone
PcG - Polycomb Group
RRBS - Reduced Representation Bisulfite Sequencing
SD - Short Days
SNP - Single Nucleotide Polymorphism
T3 - Triiodothyronine
T4 - Thyroxine
TSH - Thyroid Stimulating Hormone
TSHB - Thyroid Stimulating Hormone beta subunit
TSS – Transcription Start Site
TTS - Transcription Termination Site

# 1 - Abstract

At our latitude and under natural conditions, goats show autumn heats with calvings in spring. Individual variability in the onset of puberty and reproductive activity could hide a genetic and epigenetic variability that can allow a selection of goats more prone to deseasonalization. As a consequence, breeders could easily spread milk production throughout the year. The aim of the project was to analyze epigenetic profiles responsible for the reproductive seasonality in goat species. Considering the lack of information in literature for the species, we decided, first of all, to focus our effort on the characterization of the epigenetic profile (involving DNA methylation and microRNA) of two key organ in reproduction performances: hypothalamus and ovary. This basic research gave a first picture of the epigenome of goat species, representing a starting point for future researches in this field. In the three year of my PhD we gathered biological samples of interest for the project and we set up a pipeline for the epigenetic analysis (for DNA methylation and smallRNA) of the data produced. We saw different DNA methylation patterns across the goat genome between two organs, hypothalamus and ovary, and even a marked difference in the same organ (hypothalamus) between two physiological stages, before and after reaching puberty. In particular, the comparison between transcriptome and methylome in hypothalamus and ovary showed that a higher level of methylation is not accompanied by a higher gene suppression. We discovered interesting differences in the expression of microRNA in hypothalamus and ovary. We confirmed microRNAs already known on databases and we discovered some new, through structure homology with other related species. A selection of differentially expressed miRNAs was identified among hypothalamus, pituitary and ovary. In addition, 44, 55, and 69 miRNAs identified in pituitary, hypothalamus and ovary respectively, were specifically up-regulated in only one organ. Then we focused our efforts on the characterization of the epigenetic factors that could drive to puberty in the species. Genome wide DNA methylation analysis revealed 664 Differentially Methylated Genes (DMGs) among pubertal goats and their control. The crossing of DMGs and Differentially Expressed Genes (DEGs) among pubescent goats and their control, revealed new 8 genes, as potential marker for puberty onset. Among these, *ATG16L1* (autophagy related 16 like 1), *LRP5* (low-density lipoprotein receptor-related protein 5) and *MERTK* (proto-oncogene tyrosine-protein kinase MER) are suitable examples of puberty-related genes. According to False Discovery Rate threshold, it seems that miRNA population does not play a crucial role in the reaching of puberty in goat species in the organ analyzed.

Furthermore, we conducted a SNP discovery on three key genes for photoperiod response (Melatonin Receptor 1 - *MTNR1A*, type 2 Deiodinase Gene - *DIO2*, Thyroid Stimulating Hormone beta subunit - *TSHB*), confirming and identifying new genetic polymorphisms. We found a novel non-synonymous polymorphism in the caprine *DIO2* gene and we confirmed the high variability of the *MTNR1A* gene, discovering a new SNP bringing to a silent mutation.

The first genome-wide comparison among organs of DNA methylation in *Capra hircus* has been generated. The DNA methylome in the hypothalamus and ovary of goat offers a valid base for studying the involvement of epigenetic modifications in goat reproduction performances. By confirming that the

neuroendocrine control of female puberty entail mechanisms subjected to epigenetic regulation, our results provide insight into genes and pathways involved in the system used by the hypothalamus to control the initiation of mammalian puberty. As such, they are consistent with the concept that the pubertal process depends not only on genetic determinants, but also on developmentally regulated changes in epigenetic information.

# 2 - Introduction

## 2.1 Epigenetics in livestock

The term “epigenetics” was coined half a century ago, as a fusion of the words “epigenesis” and “genetics”, to describe the mechanisms of cell fate commitment and lineage specification during animal development (Waddington 1959). Today, the “epigenome” is generally used to describe the global, comprehensive view of sequence-independent processes that modulate gene expression patterns in a cell, tissue or organ. As an example, the epigenome of a cell is the complete collection of epigenetic marks, such as DNA methylation, histone tail modifications (acetylation, methylation, ubiquitylation, etc.), chromatin remodeling and other molecules that can transmit information through gene regulation such as non-coding RNA species (e.g., microRNAs and long non-coding RNAs), that exist in a cell at any given point in time (Rakyan et al. 2012). In fact, epigenetic mechanisms encompass processes that regulate gene expression at the transcriptional and post-transcriptional levels with resultant effects on the phenotype without changes on the DNA sequence. Accumulating evidence shows that epigenetic marks influence gene expression and phenotypic outcome in livestock species. However, epigenetic research activities on farm animal species are currently limited, partly due to lack of recognition, funding and a global network of researchers. Therefore, considerable less attention has been given to epigenetic research in livestock species in comparison to extensive work in humans and model organisms. Elucidating the epigenetic determinants of animal diseases and complex traits may represent one of the principal challenges to use epigenetic markers for further improvement of animal productivity (Ibeagha-Awemu & Zhao 2015).

Increasing documentations indicate that two or more epigenetic mechanisms may interact to control gene expression. Remarkably, miRNAs are involved in global DNA hypo-methylation through their targeting of DNA methyltransferases (DNMTs) in 3’untranslated regions of genes (Hu et al. 2014). Epigenetic modifications in mammals have essential and important roles in genome reprogramming and in the expression of genes that control growth and development. Recent findings indicate that epigenetic factors play critical roles in the expression of imprinted genes, cellular processes and the development of muscle tissue in livestock species (Bartolomei 2009). Despite the fact that reprogramming is required to remove epigenetic signatures acquired during development or imposed by the environment, several reports indicate that environmental influences such as exposure to chemicals, nutrition and maternal behavior cause modifications in gene expression that persist throughout life and may be transmitted to the next generation (Anway & Skinner 2006; Kaminen-Ahola et al. 2010). However, such epigenetic markers need to be stable or show evidence of being transgenerationally inherited. Alternatively, understanding how certain triggers, such as diets or management strategies modulate epigenetic marks could lead to improved management practices for increased productivity. A major limitation to livestock epigenomics research is the insufficient recognition of the importance of epigenomic contributions to the emergence of livestock phenotypes of economic importance and disease traits.

Epigenetic mechanisms could be the missing yet uncovered players in the expression of complex animal production traits and disease etiology and are likely responsible for a portion of the heritability that is not accounted for in existing genetic assessment schemes. Concerted efforts from several research groups led to the production of the epigenome maps of different cell types in humans (Rivera & Ren 2013). The use of these maps has furthered our understanding of the contribution of the epigenome to different biological processes. The importance of producing the epigenome maps of the different cell types, organs and tissues of livestock species is no longer contestable (Ibeagha-Awemu & Zhao 2015). Recently, Couldrey and Cave (Couldrey & Cave 2014) reviewed available tools that could be used to assess DNA methylation levels in livestock species and concluded that a great deal of work is required before present technologies can find wide applications in animals. Limited technical knowhow in informatics management of large scale data generated by new sequencing technologies and platforms may slowdown epigenomics discovery in livestock species. Therefore, tools for livestock epigenetics research are needed to drive discovery and application.

It has known that the phenotype, at the basis of the breed improvement progress, results from interaction of the genotype, epigenotype, and environmental/other factors. Therefore, both genomic and epigenomic information might have been unintentionally applied in animal breed improvement all along. Given overwhelming evidence that epigenetic marks contribute to the appearance of different phenotypes in livestock species, it is probable that the primary goal over the next decade will be to accelerate epigenetic research to enable the understanding of how epigenetic marks influence livestock phenotypes under different conditions. It is only then that epigenomic information can complement genomic information and provide a better understanding of the forces that shape livestock phenotypes and directional application in breed improvement and management practices. The association of epigenetic marks with different phenotypic outcomes in livestock species is pointing to the notion that unexplained phenotypic variation in livestock traits could be due to epigenetic factors. It is advantageous to accelerate research on how epigenetic marks influence livestock complex diseases and production traits under different conditions and include this information to increase animal productivity and sustainability (Ibeagha-Awemu & Zhao 2015).

### 2.1.1 DNA methylation

DNA methylation has a key role in regulating gene expression. The addition of a methyl group to the 5-carbon position of a cytosine is managed by a class of enzymes known as DNA-methyltransferases (DNMT) with S-adenosyl-methionine as the methyl donor. Of utmost importance are DNMT1, DNMT3a, and DNMT3b because the enzymatic activity of DNMT1 maintains DNA methylation during DNA replication, while DNMT3a and DNMT3b are responsible for *de novo* methylation (Jin et al. 2011). The connection between methylation and gene expression is intricate, with high levels of gene expression frequently associated with low levels of methylation in the promoter region of genes (Jones 2012). Methylation within genes has been hypothesized to influence gene expression levels by reducing the

rate of transcriptional elongation, despite the fact that the cellular mechanism responsible for the link between gene-body methylation and gene activity remains unclear (Medvedeva et al. 2014; Szyf 2012; Lou et al. 2014). In fact, DNA methylation patterns can be inherited and influenced by environment, diet and aging (Jaenisch & Bird 2003). It has been also shown that methylation plays a key role in X-chromosome inactivation (Cotton et al. 2015), differentiation and development of tissues and imprinting of genes (Straussman et al. 2009; Smith & Meissner 2013), while aberrant DNA methylation is implicated in many types of disease, including cancer (Rodríguez-Paredes & Esteller 2011; Szyf 2012). Moreover, hypo-methylation, and not hyper-methylation, was more likely to be associated with the tissue-specific functions. Still, it remains unclear how the gene body tissue specific DMRs may function as regulators of gene expression, and this question should be addressed in the future epigenetic studies (Lokk et al. 2014). An overview of the effects caused by CpG methylation is shown in figure 1.

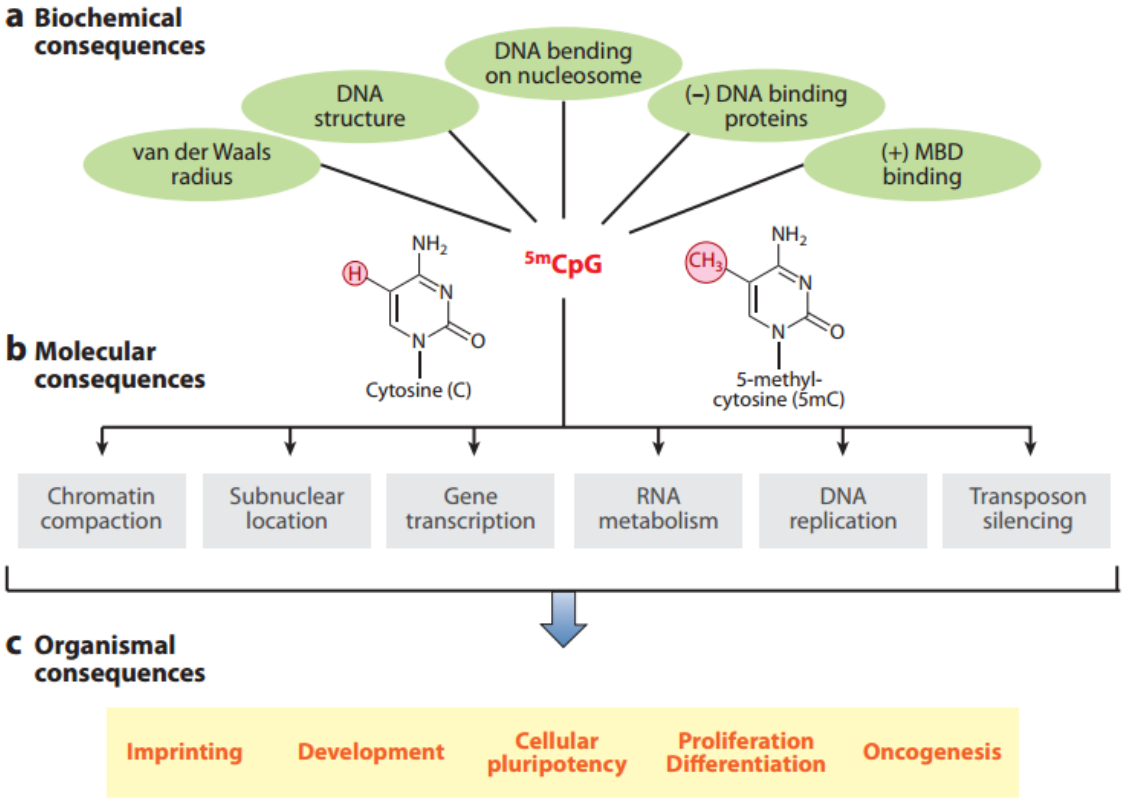


**Figure 1** - Overview of how methylation of cytosines influences biological processes. (a) The methyl group on cytosine can exert biochemical changes to DNA, which are either direct (van der Waals radius, DNA structure, DNA bending on nucleosomes) or indirect (enabling/disabling the association of protein binding). (b) The biochemical modifications serve as molecular signals for chromosome function via various means. (c) The combination of these molecular signals determines the phenotype (development, physiology, and pathology) of an organism (Franchini et al. 2012).

In mammals, methylation is commonly found at the 5-carbon position of cytosines and is predominantly observed at CpG dinucleotides. It occurs in GC-rich regions called CpG islands (CGIs), often associated with the 5' end of genes and is a potential gene marker (Smith & Meissner 2013). Since CpG islands represent an important gene feature and are expected to be broadly used in gene prediction, gene feature analysis, and epigenetic studies, their investigation is very important (Han & Zhao 2009). In early 2000s, the ability to measure levels of DNA methylation in the genome was mainly limited to qualitative analysis of single or small numbers of CpG sites at which restriction enzymes cleaved DNA (Galm et al. 2002). The last 15 years have seen a rapid development of a wide array of tools for measuring DNA methylation. Today, there exists an entire tool box of techniques that can be used depending on resources available, resolution required and the number of CpG sites to be analysed. We are no longer limited by a lack of ability to measure DNA methylation but, rather, where to look in the genome to identify epigenetic mechanisms controlling phenotypes of interest and/or how to interpret the vast quantities of high-quality data that is readily generated through Next Generation Sequencies (NGS). Interpretation of data is further complicated by the need for sodium bisulphite conversion or other techniques to allow DNA methylation to be measured. The ability to target specific epigenomics regulatory elements may provide powerful tools for manipulating gene regulation. Epigenome-wide association studies (EWAS) from mouse and human researches have been performed on complex diseases. DNA methylation is correlated over tissue-specific blocks of CpG sites spanning up to 1 kb. Knowledge of this block structure for different tissues and cell types is improving the selection of CpG sites for EWASs as new methylome maps become available. EWAS can complement large- scale genome-wide association studies (GWAS) which are on- going for livestock animals (Rakyan et al. 2012).

Another prerequisite for understanding the function of DNA methylation is the localization of its distribution in the genome. Methyl-CpG binding domain protein sequencing (MBD-seq) is a cost efficient method to investigate genome-wide methylation patterns with high efficiency and resolution (Lan et al. 2011). Genome-wide DNA methylation studies of many livestock species have been recently reported (Couldrey et al. 2014; Su et al. 2014; Yuan et al. 2016; Hu et al. 2013; Lee et al. 2014). These studies utilize different sequencing methodologies to explore a variety of tissue specific methylation associated with economically important traits. Methylomes were explored in sheep and pigs using reduced representation bisulfite sequencing (RRBS) while MeDIP-seq was utilized in horses, cows and chickens. A recent review summarizes evidence associated with epigenetic modifications and variable phenotypic outcomes in livestock species (Ibeagha-Awemu & Zhao 2015). Unexplained phenotypic variation in economically important traits, like lipid synthesis, milk production, growth and development could be influenced by epigenetic factors, such as DNA methylation. These studies indicate that the spectrum of methylation levels and patterns in livestock is very broad (Bird 2002), although, methylation of goat species remains to be better explored.

There are evidences that epigenetic mechanism of transcriptional repression, operating in the neuroendocrine brain, influences the timing of female puberty and the adulthood seasonal cyclicity (Gabory et al. 2011). The reproductive function and puberty onset are also sensitive to nutritional

influences during embryonic development. The processes through which maternal environment affects reproductive function in the offspring suggest that epigenetic modifications are an important link. Further studies are needed to better understand the mechanisms involved, identify the crucial critical periods, and prevent or treat the adverse effects (Dupont et al. 2012).

### 2.1.2 Noncoding RNA: microRNA

Noncoding RNAs (ncRNAs) accomplish a remarkable variety of biological functions. They regulate gene expression at the levels of transcription, RNA processing, and translation. They protect genomes from foreign nucleic acids. They can guide DNA synthesis or genome rearrangement. These RNAs are divided into multiple families based on their sizes and biogenesis pathways, and act as RNA-protein complexes in regulating gene expression (Cech & Steitz 2014). Long ncRNAs (lncRNAs), are members of one ncRNA family, which size exceed 200 nucleotides (Mercer et al. 2009). LncRNAs make up the largest portion of the mammalian non-coding transcriptome. Antisense lncRNAs may regulate the expression and function of genes with key roles in sex determination in the gonad, such as *forkhead box L2* (Georges et al. 2014). A direct involvement of lncRNAs in the regulation of sexual maturation and HPG axis was observed in a precocious puberty model in rat. The premature activation of the hypothalamic-pituitary-gonadal axis was reported to be related to Maternally Expressed Gene 3 *MEG3* dysregulation, encoding for a lncRNA of approximately 1700 nucleotides. Small ncRNAs (sncRNA) are a subcategory of ncRNAs less than 200 nt long. This group includes many infrastructural or regulatory RNAs (transfer RNA - tRNAs, small nuclear RNA - snRNAs, small nucleolar RNA - snoRNAs, micro RNA - miRNAs, piwi interacting RNA - piRNAs, and others) that were grouped according to their genomic origins and biogenic processes. Various classes of ncRNAs have roles in promoting the mammalian sexual phenotype (McFarlane & Wilhelm 2009), and hundreds of siRNAs, miRNAs, piRNAs, and snoRNAs are reported to be expressed in the mouse testes and ovaries (Mishima et al. 2008).

Among all ncRNA families, microRNAs (miRNAs) assume an important epigenetic regulation mechanism. They are a group of single-stranded noncoding 21–24 nt RNAs that regulate targeted mRNA expression at a post-transcriptional level. They are involved in various aspects of eukaryotic biology including reproduction and development (Kuhn et al. 2008). It is estimated that miRNA genes may account for 2-5% of all mammalian genes and regulate the expression of up to 60% of protein-coding genes (Guo et al. 2010). Long hairpin precursors are generated in the nucleus (pre-miRNAs, ~70 nt) and cleaved out by the microprocessor Drosha endonuclease and cofactors. The pre-miRNAs are then exported to the cytosol where they are cleaved by the Dicer protein, releasing the loop and a duplex consisting of the mature -5p and -3p miRNA (Winter et al. 2009). The miRNAs are subsequently incorporated in the miRNA-induced silencing complex (miRISC) so that they can act on their targets. Approximately half of vertebrate miRNAs are located in intergenic regions, referred to as intergenic miRNAs. The other half are intragenic, localized in introns and exons (about 40% and 10%, respectively) of protein-coding transcription units. Intragenic miRNAs tend to be coexpressed with their host genes. Host genes and their resident miRNAs have been considered to have synergistic effects. Indeed, genes

highly correlated in expression with an intragenic miRNA gene have been found to be more likely predicted as miRNA targets (Tsang et al. 2007). To even more complicate the picture every miRNA could act in the regulation of several mRNA targets. MiRNAs seem to be crucial in the regulation of seasonal estrus, although further analyses are needed to better understand their role. They have been observed to have an active role in modulating the hypothalamic-pituitary-gonadal axis (HPG axis). Lin28/let-7 in hypothalamus influence the correct timing of puberty (Sangiao-Alvarellos et al. 2013). Lin28a and Lin28b expression decreases across the pubertal transition and Lin28b expression in the ovary may affect the timing of puberty and/or follicular development (Grieco et al. 2013). Genome-wide miRNA expression has been examined in the gonad of mice (*Mus musculus*) (Mishima et al. 2008) cattle (*Bos taurus*) (Huang et al. 2011), pig (*Sus scrofa*) (Li et al. 2011) and also in goat (*Capra hircus*) (Zhang et al. 2013). Future works could be performed to characterize the expression of ovarian and hypothalamic miRNAs at different stages of reproduction and in different breeds of goat, or in specific cell lines derived from ovarian tissues (Zhang et al. 2013; Ling et al. 2014). Moreover, to date, only 267 caprine miRNAs have been identified and listed in the miRBase database (Release 21), the database of reference of all the microRNA. All these miRNAs refer to only one publication in which researchers analyzed the testis of Saanen goats (Wu et al. 2014). No further information is available for the species. Even many studies reported the repertoire of microRNAs in ovary and testis, details of how miRNAs expression vary in HPG axis remain still not investigated.

Nowadays it is clear that phenotypic complexity is not just a simple matter of Mendelian genetics, and we realize that our comprehension of the mechanisms involved is far from complete. A fundamental role of epigenetics is assessed for adulthood reproductive seasonality. The processes of modulating reproductive transitions probably involves epigenetic molecular regulation that alters temporal and spatial patterns of gene expression. Epigenetic imprinting, resulting from genome-environment interactions, can further affect the following generations. Although direct effects of epigenetic modulation on annual cycles are unknown, perinatal exposure to various photoperiods or temperatures influenced circadian, endocrine or thermoregulatory traits later in life that could be highly relevant for seasonal timing. This suggests an urgent need for further studies (Helm et al. 2013). The goal now is to expedite research in epigenetic processes to provide a better understanding of the underlying mechanisms governing full phenotypic determination for the overall animal health and productions (Triantaphyllopoulos et al. 2016).

## 2.2 Reproduction in mammals

Reproduction is the pivotal event of whatever livestock activity. The HPG axis controls throughout the reproductive life cycle pathways that are responsible for the development, puberty, adulthood and reproductive systems (Maffucci & Gore 2009). Information emanating from a variety of external cues (e.g. visual, auditory, tactile, olfactory) is conveyed into the central nervous system and converges on the hypothalamus. The hypothalamus is the portion of the anterior end of the diencephalon that lies below the hypothalamic sulcus and in front of the interpeduncular nuclei. The hypothalamus is the

ventral-most part of the diencephalon. It extends from the rostral limit of the optic chiasm to the caudal limit of the mammillary bodies. Three rostral to caudal regions are distinguished in the hypothalamus that correspond to three prominent features on its ventral surface: 1) The supraoptic or anterior region at the level of the optic chiasm, 2) the tuberal or middle region at the level of the tuber cinereum (also known as the median eminence—the bulge from which the infundibulum extends to the hypophysis), and 3) the mammillary or posterior region at the level of the mammillary bodies (Figure. 1). The hypothalamus controls the release of 8 major hormones by the hypophysis, the food and water intake, sexual behavior and reproduction and daily cycles in physiological state and behavior. Furthermore, it is involved in temperature regulation and in mediation of emotional responses. The pituitary gland, or hypophysis, consists of two major subdivisions, the anterior and the posterior lobe. The posterior lobe is made up of neural tissue and is connected to the rest of the brain via the infundibular stem, or pituitary stalk. Thus, there is a direct neural link between the posterior pituitary and the brain. The anterior lobe of the pituitary (or adenohypophysis) is further subdivided into the pars distalis, pars intermedia and pars tuberalis. The pars tuberalis surrounds the infundibular stem like a cuff and extends upwards to lie beneath the portion of the median eminence. The anterior pituitary communicates with the brain by a vascular connection, the hypothalamo-hypophyseal portal system (Austin & Short 1982).

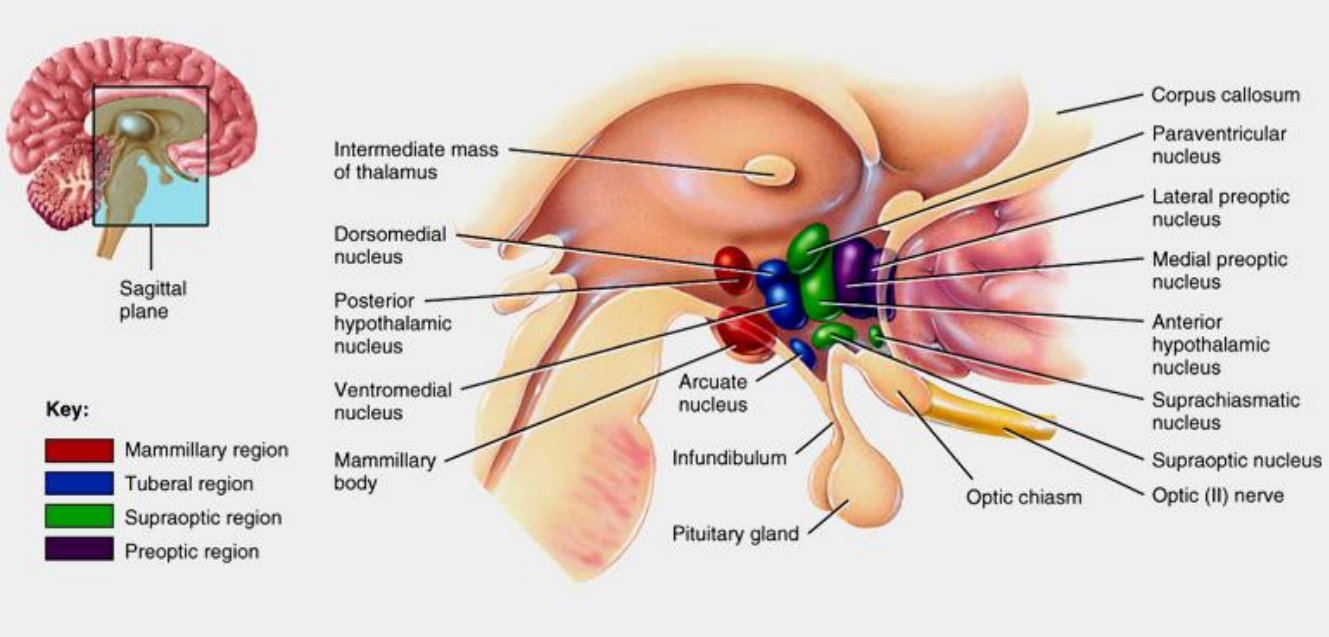


**Figure 2** - Graphical representation of pituitary and hypothalamus in humans. An indication of the nuclei that compose hypothalamus is shown.
Source: http://humanbrainfacts.org/hypothalamus.php

The anterior pituitary secretes six hormones, including follicle-stimulating hormone (FSH) and luteinizing hormone (LH). Pituitary FSH is essential for development and maintenance of ovarian follicles in single and multiple ovulating species (Kaneko et al. 1991a). The production and/or secretion

of FSH is suppressed by Inhibin (INHBE), a glycoprotein hormone produced by granulosa cells of ovarian follicles, through negative feedback at pituitary level (Burger et al., 2008). The LH is important in studies of ovarian activity since its pre-ovulatory peak is responsible for the follicular wall rupture and ovulation in some species. Both FSH and LH play a fundamental role also in the onset of puberty in mammals.

### 2.2.1 Puberty

Puberty is the culmination of a complex series of maturational events that lead to the completion of sexual and somatic development and the acquisition of reproductive competence (Tena-Sempere 2013). In female, puberty is characterized by the manifestation of oestrus and ovulation. This is a developmental process with genetic drivers conserved among species (Matzuk & Lamb 2002).

Puberty is a product of increased activity of the hypothalamic-pituitary-gonadal axis leading to production of gonadal steroids and other growth-associated hormones (McCarthy 2013b). Gonadotropin-Releasing Hormone (GnRH) from the hypothalamus regulates the pituitary gonadotrope production of follicle stimulating hormone and luteinizing hormone, which in turn reaches the gonad and promotes steroidogenesis (Figure 3).

During pubertal development, GnRH secretion modulates low-level irregular pattern into a regular pattern of pulsatile secretion, which is fundamental in initiating this process (Plant et al. 2006). The increased GnRH secretion at puberty is determined by a cascade of hormonal events. During the prepubertal period an inhibitory neuronal system suppresses GnRH release and during the subsequent maturation of the hypothalamus this prepubertal inhibition is removed, allowing the adult pattern of pulsatile GnRH secretion. However, interruption of inhibition proves insufficient for induction of puberty; there is also the need for an accelerator, which must include fine-tuned temporal control of GnRH neurons in what is referred to as the GnRH pulse generator, so that luteinizing hormone is released from the pituitary at the appropriate frequency and amount (Windsor-Engnell et al. 2007). In this regard, during the past decade evidence has accumulated suggesting GnRH secretory activity is modulated by a specific glial-neuronal gene family which synthesizes adhesion/signaling proteins involved in the functional and structural integrity of bi-directional glial-neuronal communications (Srivastava et al. 2011). Furthermore, concerning the onset of puberty, the CpG system of transcriptional silencing seems to play a central role. For example, hypothalamic expression of chromobox 7 (*Cbx7*) and embryonic ectoderm development (*Eed*), two genes required for CpG action, decreased before the onset of puberty (Gil et al. 2004; Woo et al. 2010), and this change was associated with increased DNA methylation of their 5' flanking regions. Conversely, pharmacological inhibition of DNA methylation seems to prevent the pubertal increase in *Eed* and *Cbx7* DNA methylation, reverse the low peripubertal *Eed* and *Cbx7* mRNA levels to elevated early-juvenile values, and delay puberty. Hence, the pubertal process depends not only on genetic determinants, but also on developmentally regulated changes in epigenetic information (Simon & Kingston 2009).

The timing of puberty onset is an important phenotype for the livestock industry because late puberty has negative effects on reproduction rates and profitability.

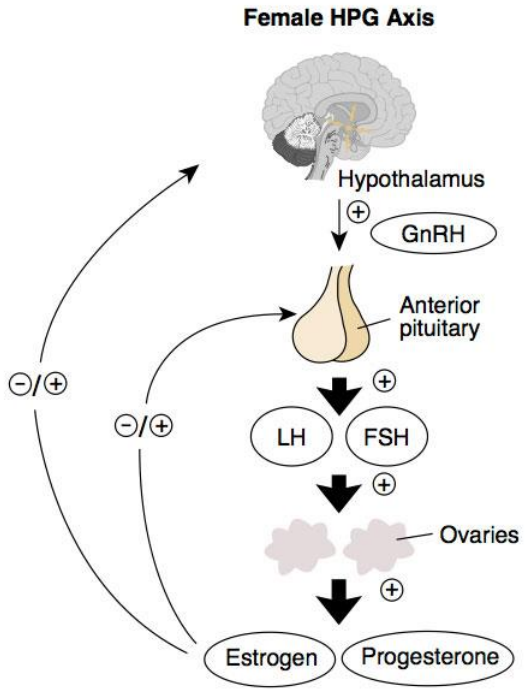


**Figure 3 -** Schematic representation of the Female Hypotalamus-Pituitary- Gonad Endocrine Axis (Hiller-Sturmhöfel & Bartke 1998). The hypothalamus secretes GnRH that acts on the pituitary gland stimulating its releasing of gonadotropins (i.e., LH and FSH).

### 2.2.2 The oestrus cycle

From the time of puberty, the female begins to express the oestrous cycle which, in goat, can vary from 17 to 25 days. The oestrous cycle consists of all morphological and physiological changes in the ovaries and genital tract leading to oestrus expression (phase of receptivity towards males) and ovulation and the preparation of the genital tract for copulation, fertilization and embryo implantation. During the breeding season, females can undergo several oestrous cycles successively and the number of successive cycles is dependent on the length of the breeding season and the breed of goat. The length of oestrous cycle is defined by the interval between two successive expressions of oestrus or two successive ovulations. While the average duration of the goat oestrous cycle is of 21 days (Figure 4), its length is highly variable.

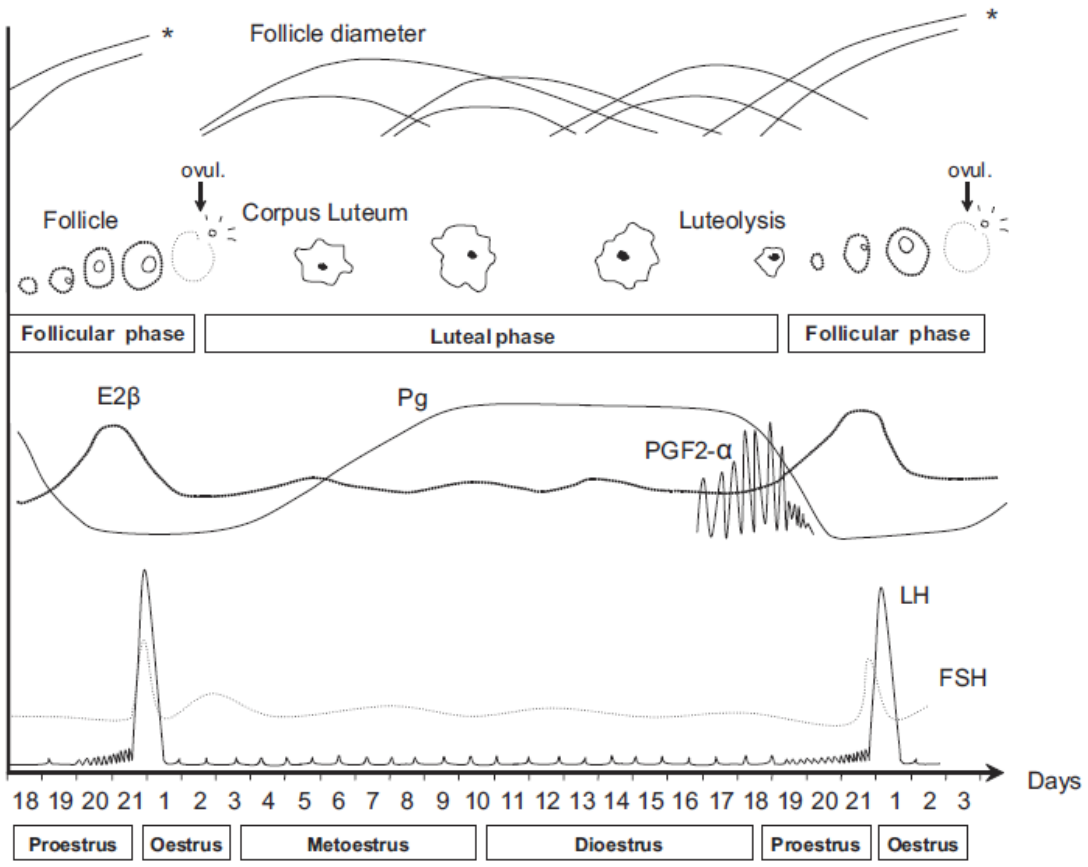


**Figure 4** - Pattern of follicle development, ovarian cycle and endocrine regulations in goat. Schematic representation of the different physiological events occurring during oestrous cycle. *Ovulatory follicle(s) (Fatet et al. 2011).

A study on Alpine goats during the breeding season recorded 77% cycles of normal duration, 14% short cycles (8 days in average) and 9% long cycles (39 days in average) (Baril et al. 1993). The relative high frequency of short cycles is characteristic of goats and increases when ovulation is induced either just before or during breeding season. This proportion can be modulated by environmental factors such as photoperiod and nutrition. During the oestrous cycle, ovaries undergo a number of morphological (follicular recruitment and growth), biochemical (follicle maturation) and physiological (endocrine regulations) changes leading to the ovulation. These cyclical changes in the gonads are referred to as the ovarian cycle.

The ovarian cycle is classically divided in two phases: the follicular phase and the luteal phase (Figure 4). The follicular phase corresponds to the wave of follicle development providing the ovulatory follicle and involves maturation of gonadotropin-dependent follicles until ovulation (terminal growth). During the follicular phase, FSH secreted by the pituitary gland stimulates follicular growth. A cohort of gonadotropin-dependent antral follicles of 2–3mm of diameter is recruited and follicles enter their terminal growth. Only 2–3 of these follicles reach 4mm diameter and are selected to enter the dominance phase. Under the influence of LH, they reach the pre-ovulatory stage (6–9 mm), while

subordinate follicles degenerate (follicular atresia). The increase in peripheral concentrations of oestradiol, secreted by bigger follicles, induces oestrous behaviour and acts as a positive retrocontrol on the gonadotropic axis. The consequent increase in GnRH secretion induces the pre-ovulatory LH surge which induces ovulation 20–26 h later and subsequently luteinization of follicular cells. The beginning of the follicular phase, before overt oestrous behaviour is observed, is also referred to as the proestrus. The oestrous phase includes events from overt oestrous behaviour to ovulation (Fig. 1). Both season and nutrition are known to affect the ovulation rate, especially in the Angora breed. Angora goats typically have a single ovulation but they may have two optimizing nutritional conditions. The luteal phase starts from the time of ovulation. About 5 days after the onset of oestrus, cells of the ovulating follicle turn into luteal cells and form the corpus luteum (CL). They secrete progesterone causing its concentrations to increase and remain at a high level (>1 ng/ml) during 16 days. During this luteal phase, gonadotropin-dependent follicular growth continues in a wave-like manner but progesterone inhibits ovulation. At the end of the luteal phase, 16–18 days after oestrus, prostaglandin F2 secreted by the non-gravid uterus induces the CL regression – called luteolysis – and the decrease of progesterone secretion. This gradually removes the inhibition of gonadotropic hormones secretion and a new follicular phase then commences (Baril et al. 1993). The luteal phase is also called the postoestrous period, which can be divided in metoestrus, when peripheral concentrations of progesterone begin to rise, and dioestrus, when peripheral concentrations of progesterone are high up to the start of luteolysis.

### 2.2.3 Photoperiod and seasonality

Domestic animals are broadly classified as seasonal and non-seasonal breeders depending on if they successfully mate only during certain times of the year or not. Non seasonal mammals, as cows, are not influenced by photoperiod and breed all through the year irrespective of the season or part of the year. Seasonal breeders, as goats, are those that have specified period of time in a year during which they actively breed. This is a survival strategy adopted by many mammals to ensure that the offspring are born at the most favorable season of the year (Vasantha 2015). This biological programming of births, or synchronization of reproductive response to appropriate environmental conditions, clearly leads to distinct advantages for the offspring, being born at the time of most suitable weather and maximal food availability during the early stages of life (Wood et al. 2006). Most of the organisms perceive the changes in the length of the day, defined as photoperiod, to distinguish seasons, because temperature and precipitation varies throughout each year and are unreliable when compared with the length of the day. This phenomenon is called “photoperiodism”. Among the various seasonally regulated phenomena, the mechanism of seasonal reproduction has been extensively studied. Animals that breed during the spring and summer are called long-day (LD) breeders, while others, such as goats and sheep, breed during fall and are called short-day (SD) breeders (Figure 5). These animals have a gestation period of approximately 5 months. Therefore, their offspring are born and raised during spring and summer, when the climate is moderate and food is abundant (Shinomiya et al. 2014).

Another example of short-day breeders are water buffaloes. Under Mediterranean latitudes, reproductive efficiency of buffaloes is usually negatively affected by increasing day-length. Buffaloes become sexually active in autumn. After a 10 months long gestation, they give birth to offspring in early autumn (Zicarelli et al. 1997). According to their origin, they mate and give birth in the rainy season, because wet conditions are the ideal habitat for the species (Perera 2011).

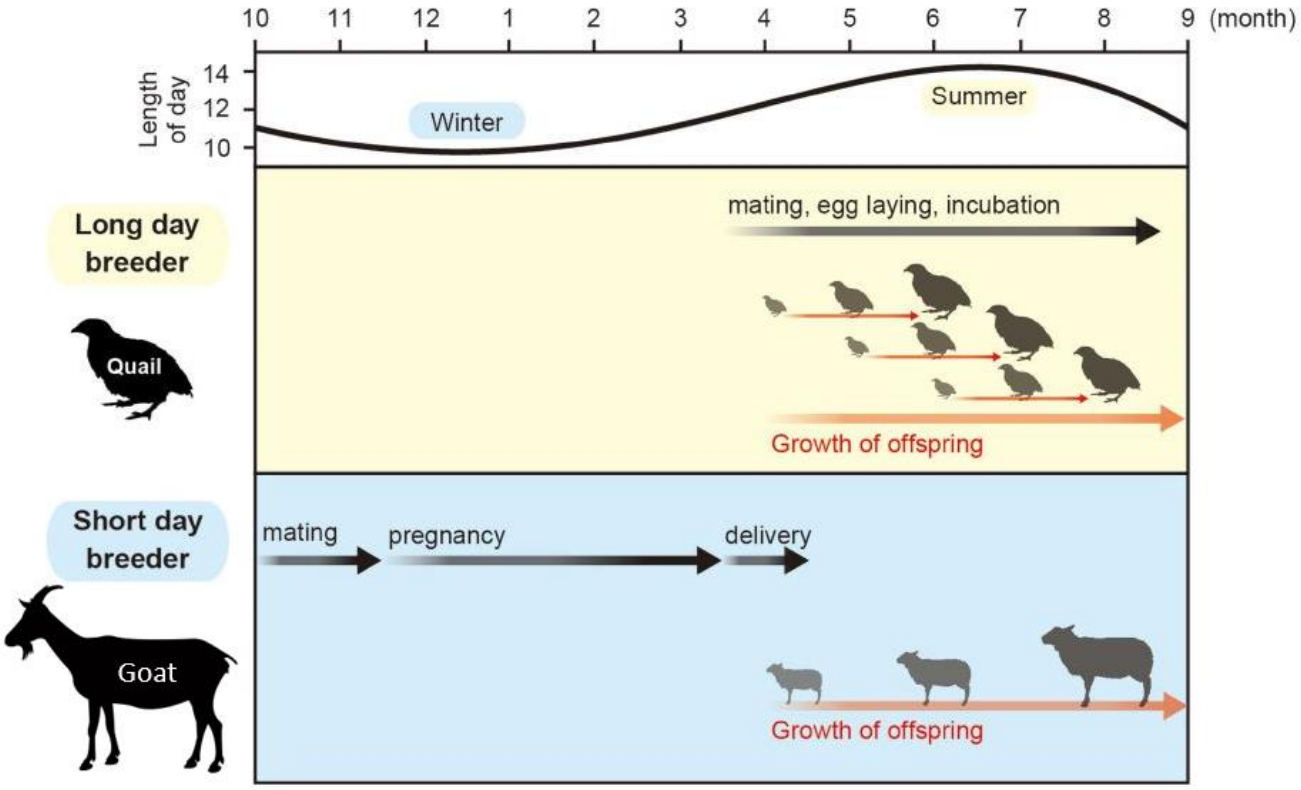


**Figure 5 -** Calendar of seasonal breeding animals. Seasonal breeder mate in a specific time of a year. Small animals with short gestation or incubation period mate in spring and summer, while large animals that have a 5-month gestation period mate in fall to give birth in spring (Shinomiya et al. 2014).

The annual photoperiod cycle provides a critical environmental signal, which entrains seasonal physiology (Dupré et al. 2008). Photoperiod regulates changes in the daily melatonin secretion by the pineal gland. The pattern of melatonin secretion provides photoperiodic information to cells within the brain that possess the relevant receptors and control reproductive function (Migaud et al. 2005). Nocturnal secretion of the pineal hormone melatonin reflects the seasonal changes in photoperiod and thereby provides the brain with an internal hormonal representation of external photoperiod changes.

### 2.2.4 *Capra hircus*

Goat is a fundamental livestock resource, particularly in rural areas where the breeding of more productive but less adaptable animals is difficult. However, *Capra hircus* is still considered a secondary

species and is poorly studied from an (epi)genetic point of view. The world's goat population reached approximately 1.2 billion in 2014 according to the latest FAO report (http://www.fao.org/faostat). Goats are widely distributed in developing regions, but less so in developed regions. In fact, Asia with 56 % (mainly China and India), Africa with 30 % (highest numbers in Nigeria and Kenya) and the Near and Middle East with 7 % (principally in Sudan and Yemen) account for the vast majority of the world's goats. There are also significant 3 % of the total population in Latin America and the Caribbean (mainly in Mexico and Brazil) and in Europe and the Caucasus (3 %; highest numbers in Turkey and Greece). The main change since 2005 has been a large increase in the African goat population, who contributes with the 4% of the global total rising, and with the 27 % in absolute terms of population size rising (FAO 2015). There is also a growing trend for goat milk and meat consumption over recent years. The increase has been even more marked than that of cattle, for both productions. In numbers, milk consumption per years, increased of more than 3 tonnes, as shown in Table 1. These increases are mainly due to the higher number of animals raised, than to a higher productivity. Understanding and mastering the genetics behind reproduction will help to improve productions.

| Product | 2004 | 2012 | Change |
|---|---|---|---|
| | tonnes | | % |
| Cattle meat | 58 093 900 | 63 288 600 | 9 |
| Chicken meat | 68 003 800 | 92 812 100 | 36 |
| Pig meat | 92 610 000 | 109 122 000 | 18 |
| Sheep meat | 7 836 070 | 8 470 310 | 8 |
| Goat meat | 4 382 020 | 5 300 340 | 21 |
| Turkey meat | 5 199 850 | 5 609 530 | 8 |
| Duck meat | 3 093 810 | 4 340 810 | 40 |
| Buffalo meat | 2 924 490 | 3 597 340 | 23 |
| Goose and guinea fowl meat | 1 945 640 | 2 803 720 | 44 |
| Rabbit meat | 1 419 250 | 1 833 840 | 29 |
| Horse meat | 765 229 | 750 747 | -2 |
| Camel meat | 380 947 | 524 390 | 38 |
| Donkey meat | 189 752 | 211 750 | 12 |
| Cattle milk | 529 669 000 | 625 754 000 | 18 |
| Buffalo milk | 76 872 600 | 97 417 100 | 27 |
| Goat milk | 14 368 000 | 17 846 100 | 24 |
| Sheep milk | 8 817 950 | 10 122 500 | 15 |
| Camel milk | 1 997 000 | 2 785 380 | 39 |

**Table 1** - Global output of animal-source foods (2004 and 2012) (FAO 2015).

Although most dairy goats are in developing countries, breeding programs are concentrated in Europe and North America. Genetic selection of dairy goats has resulted in considerable increases in yields and longer lactation lengths. The specialized dairy goat breeds used in developed countries therefore have higher genetic potential for milk production than breeds found in the developing world. Two of the most widely distributed dairy goat breeds are Saanen, and Alpine (Figure 6) (Manfredi et al. 2001).


**Figure 6** - Saanen and Alpine goats. Source: www.goatit.eu

### 2.2.5 Reproductive seasonality in goat

Goats are spontaneously ovulating, polyestrous animals, showing a seasonal pattern in reproductive activity related to the annual variations of photoperiod. In temperate regions, goats breed during fall and after a gestation period of approximately 5 months, they give birth to offspring in spring, with important differences in seasonality between breeds and locations.

Reproductive seasonality is a limiting factor, as both meat and milk industries are subjected to a growing demands for constant production (Fatet et al. 2011; FAO 2015). Hence, different strategies of breeding management, have been developed to meet the supply needs of consumers. At temperate latitude, out of season breeding can be achieved using strategies based on manipulation of the photoperiod (Chemineau et al. 2006). Following extended exposure to decreasing day length, animals become photo-refractory to the short day stimulus and will cease cyclic activity, unless a period of long day photo-stimulation is supplied. This is the main reason why photoperiod treatments are based on alternation of long and short days. First, the animals are subjected to long days (provided by artificial lighting in winter or by natural days in spring and summer) preparing them to respond to the stimulatory effects of subsequently administered short days (Malpaux et al. 1989). In addition, under field conditions, short day effects are further stimulated by melatonin implants. These induce high plasma concentrations of melatonin for 24 h every day, without suppressing the endogenous secretion of the pineal hormone during the night. Thereby, implants cause a short day-like response by lengthening the duration of the melatonin signal (Malpaux et al. 1997).

Another important management practice to increase the reproductive efficiency of goat is the synchronization of ovulation. This is crucial to the development of artificial insemination (AI) methods. Higher synchronous ovulations allow the protocols to be simplified and minimize costs by reducing the number of inseminations needed to optimize fertility. The most commonly used treatment consists of

a progestagen (i.e., fluorogestone acetate), eCG (Equine chorionic gonadotropin ), and cloprostenol (F2α prostaglandin analogue) (Corteel et al. 1988).

The list of the main hormonal treatments in goat species are shown in table 2. Several problems are associated with these hormones. Societal trends and European legislation (96/22/CE, 2003/ 74/CE, and 2008/97/CE) encourage producers to adopt practices that reduce or completely avoid the use of synthetic chemicals and hormones (Martin et al. 2004). Indeed, because eCG is purified from animal tissues, risks associated with the presence of pathogens in commercial eCG extracts remain. Progestagen residues in animal products and effluents could also act as hormonal disrupters and denote possible threats to human health. Regulation forbids milk to be sold for several days, whereas progestagen residues in milk remain higher than the maximal residue limits that are authorized. Moreover, hormonal treatment for the control of reproduction is forbidden in organic farms (2007/834/CE and 2008/889/CE), although the demand for organic products is increasing.

Strategies such as the “male effect” or photoperiodic treatments to control the timing of reproductive events have been proposed in order to achieve partial or total reduction of exogenous hormone use. The “male effect” consists in the introduction of males in a flock of seasonally anestrous females, which induces the activation of the secretion of LH and the consequent synchronized ovulation. AI protocols using the “male effect”, but excluding all classes of hormones or other exogenous substances, require estrus detection and many artificial inseminations (AIs) over at least 5 days after buck introduction. These protocols are neither practical nor economically viable and therefore need to be improved to reduce the number of inseminations needed for good fertility. The validation of such “zero-progestagen” protocols would also be an undisputed advantage for farmers, given the legislative constraints concerning the presence of progestagen residues in milk and the issues with organic farming (Pellicer-Rubio et al. 2016).

Photoperiodic treatments can induce sexual activity in males and females similarly to hormonal treatment in females. Clearly, the study of natural, hormone-free methods such as the male effect, photoperiodic treatments without melatonin, nutritional strategies and other innovative ways of controlling sexual activity are promising research topics. The development of hormone-free methods that allow out of season heat and artificial insemination will also contribute to faster genetic progress in the future (Fatet et al. 2011). This may be a long-term phenomenon, so it makes sense for animal producers in all countries to begin to move towards clean, green and ethical practices.

| Hormone | Pharmaceutical Form | Route of Administration | Dose | Season for Use | Time for Introduction of Male Animals into the Females | Optimum Male:Female Ratio | Comments |
|---|---|---|---|---|---|---|---|
| Progesterone | CIDR | Intravaginal × 12–14 d | 20–40 mg FGA | All year | 36–48 h after removal of devices | 1:5 | Can be used in combination with eCG |
| Progestagen | Sponge | Intravaginal × 12–14 d | 20–40 mg FGA, 60 mg MPA | All year | 36–48 h after removal of sponges | 1:5 | Can be used in combination with eCG |
| | Melengestrol acetate | In feed × 8–14 days | 2.5 mg MGA split in 2 feedings | All year | 26–48 h after removal of feed | 1:5 | eCG injected 8 h after last feeding |
| eCG | Injectable solution | Intramuscular | 250–500 IU | All year | | | Must be used after administration of progesterone or progestagens |
| Prostaglandin or synthetic analogues | Injectable solution | Intramuscular/ Subcutaneous | 125 μg cloprostenol, 7.5 mg luprostiol | Breeding period | 48 h after administration | 1:10 | Two injections with a 10-d interval |
| Melatonin | Implant | Subcutaneous | Rams: 3×18 mg; Ewes: 18 mg | Outside of breeding season | 40 d after administration to ewes | 1:20 | Males separated from females for 45 d before introduction |

**Table 2** - Principal available hormonal treatments for the control of small ruminants' reproduction (Abecia et al. 2011).

Seasonality of ovulatory activity was observed in European sheep and goat breeds (LD breeder), whereas tropical breeds show almost continuous ovulatory activity. In a recent study it was shown that tropical breeds can be sensitive to temperate photoperiod. A groups of tropical adult Creole goats was maintained in light-proof rooms under simulated temperate (8 to 16 h of light per day) or tropical (11 – 13 h) photoperiods. Ovulatory activity was determined by blood progesterone assays for more than two years. Marked seasonality of ovulatory activity appeared in the temperate group of the female goats. The percentage of goats experiencing at least one ovulation per month dramatically decreased from May to September for the three years (0%, 27% and 0%, respectively). Tropical female goats demonstrated much less seasonality, as the percentage of animals experiencing at least one ovulation per month never went below 56%. These differences were significant. A marked seasonality in ovulatory activity of tropical Creole goats was observed when females were exposed to a simulated temperate photoperiod. An unexpected finding was that Creole goats exposed to a simulated tropical photoperiod also showed seasonality in their ovulatory activity (Figure 7). Such results indicate that the species is capable of showing seasonality under the photoperiodic changes of the temperate zone even though they do not originate from these regions (Chemineau et al. 2004).

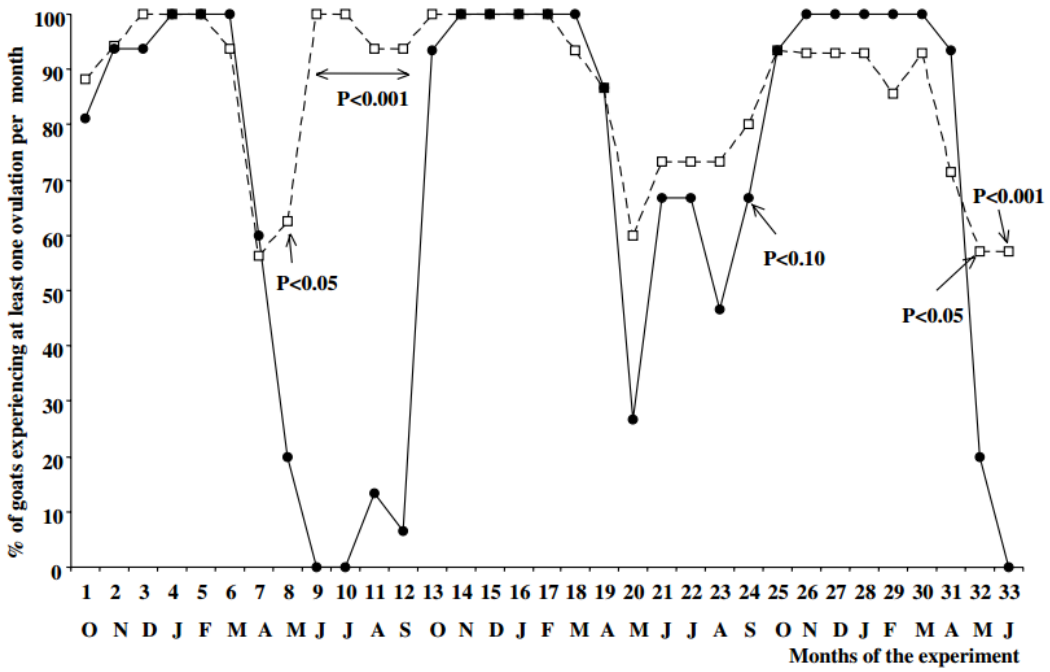


**Figure 7** - Ovulatory activity in Creole goats maintained under two different photoperiodic conditions. Solid line closed circles (-●-) refers to goats reared in Temperate photoperiod, dashed line open squares (-□-) to Tropical photoperiod (Chemineau et al. 2004).

### 2.2.6 A Molecular view on Reproductive seasonality

From a molecular point of view, photoperiodic control of reproductive patterns is mediated through circadian rhythmic secretions of melatonin by the pineal gland (epiphysis) during darkness, which influences the gonadotropin-releasing hormone (GnRH) pulse generation and the hypothalamic–pituitary–gonadal feedback loop. The MT1 receptor of the melatonin, coded by the gene *MTNR1A* (*Melatonin Receptor 1a*), binds melatonin in a complex that regulates the pulsatility of hypothalamic GnRH. This hormone, in turn, regulates the secretion of gonadotropins, as follicle-stimulating hormone (FSH) and luteinizing hormone (LH), secreted by the pituitary gland (or hypophysis). Pituitary FSH is essential for development and maintenance of ovarian follicles in single and multiple ovulating species (Kaneko et al. 1991b). The LH is important in studies of ovarian activity since its pre-ovulatory peak is responsible for the follicular wall rupture and ovulation (Terzano & Terzano 2012). Both FSH and LH play a fundamental role also in the onset of puberty in mammals. This is determined by the increased activity of the hypothalamic-pituitary-gonadal axis leading to production of gonadal steroids and other growth-associated hormones (McCarthy 2013a). During pubertal development, GnRH secretion transforms from low-level irregular pattern to a pattern of regular, pulsatile secretion, which is critical for the initiation of this process (Plant et al. 2006). The timing of puberty is controlled by many genes. The *KISS1* gene has received considerable recent attention. This gene codes for Kisspeptin (Kp), a small RF-amide peptide from a phylogenetically diverse family of peptides that have critical roles in the

control of reproduction, food intake and energy expenditure (Ebling & Luckman 2008). The neuropeptide kisspeptin stimulate GnRH neurons to initiate puberty. KISS1 neurons are localized to the arcuate nucleus and the anteroventral periventricular nucleus, in the hypothalamus. A fundamental role of Kp neurons is assessed also for adulthood reproductive seasonality. In fact, during the non-breeding period kisspeptin expression in the arcuate nucleus is markedly reduced (McCarthy 2013a). Kp appears implicated in virtually all facets of reproduction, from the initiation of puberty to the daily and seasonal control of reproduction (Beltramo et al. 2014).

Several studies investigated also the polymorphism in the *MTNR1A* gene in different seasonal species, but results are controversial. One single nucleotide polymorphism (SNP), associated with seasonal reproductive activity in buffalo (Carcangiu et al. 2011), was not confirmed in another population (Zetouni et al. 2014). Associations were found also in sheep (Carcangiu, Mura, et al. 2009) and even goat (Carcangiu, Vacca, et al. 2009), but a study performed on another goat breed did not find any relationship between *MTNR1A* polymorphisms and seasonal reproduction (Lai et al. 2013). It is well known that thyroid hormone is also essential for the maintenance of seasonal reproductive changes in a wide number of mammals. The locally generated bioactive thyroid hormone, T3 (triiodothyronine), regulates seasonal GnRH secretion. Moreover, T3 administered exogenously under short-day conditions was observed to mimic the effects of long-day conditions on gonadal growth in hamsters (Barrett et al. 2007). The thyroid hormone-activating enzyme that converts the T4 (thyroxine) to bioactive T3 is encoded by the type 2 deiodinase gene (DIO2) (Bernal 2002). The expression of DIO2 is influenced by the Thyroid stimulating hormone (TSH, also called thyrotropin). This consists of a common glycoprotein alpha subunit and a specific thyroid stimulating hormone beta subunit (TSHB) (Sairam & Li 1977). *TSHB*, mainly expressed by the pituitary, confers the specific biologic function to the hormone and it is strictly linked to the photoperiodic control and reproductive seasonality. Five mutations were found in goat *TSHB* gene with significant different distribution between seasonal and non-seasonal goat breeds (Huang et al. 2013). A graphic scheme of the photoperiodic signal transduction pathway in mammals is shown in figure 8.

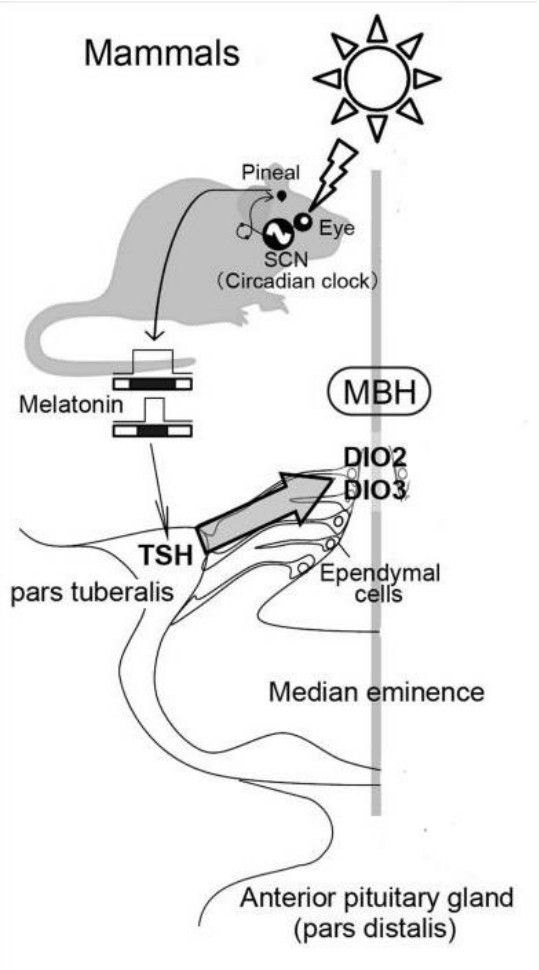


**Figure 8 -** Photoperiodic signal transduction pathway in mammals. Light stimulus is received by the eye and transmitted to the pineal gland via the circadian pacemaker, the suprachiasmatic nucleus (SCN). The duration of the pineal melatonin signal encodes the length of night and regulates TSH secretion in the pars tuberalis. The pars tuberalis TSH acts on TSH receptor expressed in the ependymal cells lining ventrolateral walls of the third ventricle (VIII) to induce DIO2 and reduce DIO3. Local thyroid hormone activation within the mediobasal hypothalamus (MBH) by DIO2/DIO3 switching plays a key role in the regulation of seasonal reproduction [15, modified].

Concerning circadian rhythms, the molecular machinery that regulates these processes comprises of a set of genes, known as “clock” genes, the products of which interact to generate and maintain the rhythms. Clock genes are strongly involved in the molecular mechanism that decodes the duration of the melatonin signal to produce a summer or winter phenotype (Lincoln 2006). The first negative feedback loop is a rhythmic transcription of period genes (*PER1*, *PER2*, and *PER3*) and chryptochrome genes (*CRY1* and *CRY2*). PER and CRY proteins form a heterodimer, which acts on the CLOCK/BMAL1 heterodimer to repress its own transcription. PER and CRY proteins are phosphorylated by casein kinase I epsilon (CKIepsilon), which leads to degradation and restarting of the cycle. The second loop is a positive feedback loop driven by the CLOCK/BMAL1 heterodimer, which initiates transcription of target

genes containing E-box cis-regulatory enhancer sequence. In recent years a new piece in the puzzle was added. In fact, several studies demonstrated the importance of epigenetic mechanisms in the estrogen regulation of anteroventral periventricular nucleus (*AVPV) Kiss1* expression to mediate estrogen positive feedback action to induce GnRH/LH surges (Tomikawa et al. 2012). Furthermore, the *GnRH* gene showed distinct epigenetic changes during embryonic development and across puberty (Kurian & Terasawa 2013).

# 3 - Aim

In this PhD project we investigated how epigenetics acts on cyclic onset of reproductive activity in goat. *Capra hircus* is an economically important livestock species that is globally distributed, especially in developing countries. It acts an indispensable part in the animal milk, meat and fiber industry (Xu et al. 2013). A better understanding of the (epi)genetic variability of puberty and reproduction cyclicity in the species will contribute for a selection of animals less sensitive to the photoperiod, with a consequent natural all-year milk production. Therefore, we explored Next Generation Sequencing (NGS) epigenetic data to find out a proper pipeline, both at molecular and at bioinformatic level, trying to fill the lack of information on epigenomic data in goat species.

Our specific aims for the PhD were:

- to optimize the bioinformatic pipeline and yields to analyze DNA methylation data;
- to point out a general view of the methylation level on each chromosome in two different reproduction-associated organs: hypothalamus and ovary;
- to explore miRNA population of key organs in the HPG axis: hypothalamus, pituitary and ovary;
- to characterize from an epigenetic point of view (DNA methylation and miRNA) the physiology that undergoes puberty in goats;
- to analyze the polymorphism of *MTNR1A*, *DIO2* and *TSHB* genes in a sample of goats subjected to the deseasonalisation practice. Together with the molecular analysis, ELISA tests were performed to evaluate the plasma level of melatonin, T3 and T4 hormones at the key steps for the reproductive cycles in order to couple molecular analysis with hormonal measurements.

Deepening the knowledge on (epi)genetic and physiologic aspects of seasonal reproduction will help the future development of goat breeding.

# 4 - Material and methods

## 4.1 Experimental design

The initial experimental design, approved by the ethic commission of the athenaeum, planned to study epigenetic pattern of 48 female goats belonging to the Alpine breed, born from regular spring calvings (between February and May 2014). In accordance with the ethic sentence, we also sacrificed 3 Adult Saanen goats, considered trials, to set up molecular and bioinformatic analyses. For the 48 Alpine goats, we established 6 checkpoints, as shown hereunder, in order to compare genetic and epigenetic patterns in different developmental stages of the selected goats. Each corresponds to a crucial moment of the animal reproductive life. The consistency and characteristics of each group are expresses above and summarized in figure 9:

Checkpoint 1 → 5 pre-pubertal goats
Checkpoint 2a → first 5 goats showing heat + 5 still non-ovulating goats as control
Checkpoint 2b → last 5 showing heat
Checkpoint 3 → out of season practice - long days phase (LD), 5 goats,
arresting the cyclic ovarian activity
Checkpoint 4a → out of season practice - short days phase (SD), first 5 showing "artificial" heat
Checkpoint 4b → out of season practice - short days phase (SD), last 5 showing "artificial" heat

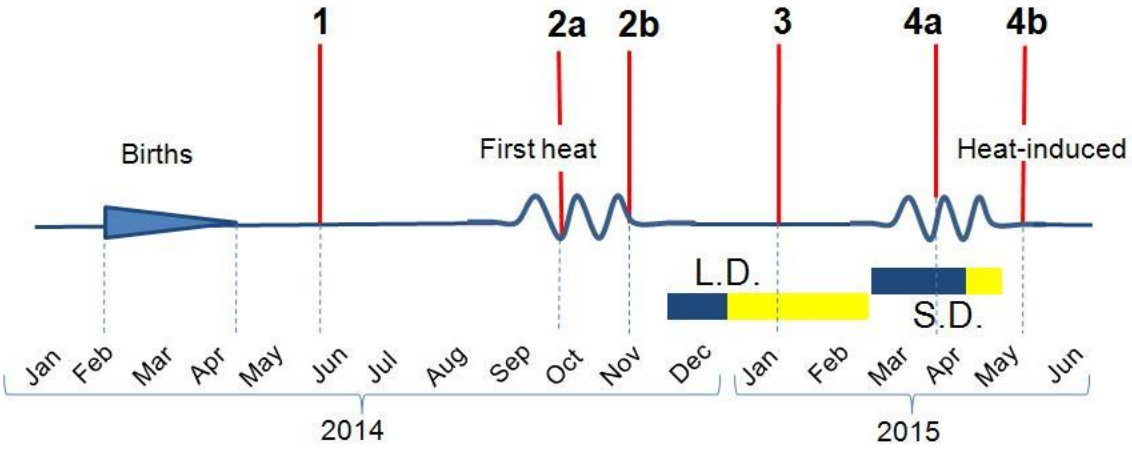


**Figure 9** - Timeline of the crucial checkpoints for the Alpine goats in the project. Each number corresponds to a checkpoint as expressed in the text.

In the on-field part, starting from the 1st of December 2014 (after checkpoint 2b), the remaining 28 animals were exposed to a classic out of season program, consisting of 3 months of light (Long day–LD in December, January and February 2015) followed by 2 months of darkness (Short day-SD in March and April 2015). During long days, goats received an intensity of at least 150 lux of artificial light for 1 to 2 hours a day, after the sunset and before the dawn. During short days, windows in the farm were darkened before the sunset and after the dawn.

## 4.2 Biological samples

In the on field part, for every Alpine goat involved in the experimentation we monitored:

- age
- weight
- progesterone levels
- Complete Blood Count (CBC)

Details on each parameter monitored are supplied hereunder.

The most suitable hormonal parameter chosen to point out the reproductive cyclicity was progesterone (P4). We set a threshold of 1 ng/ml of progesterone as an indicator of the presence of the corpus luteum. We weekly monitored the 48 goats involved in the study, measuring the progesterone levels in plasma, to point out their reproductive activity. Furthermore, we weekly recorded the weight of the animals to evaluate their development and to consider the effect of growth differences. All these activities allow us to choose the most suitable goats for each group. We performed a CBC on all the 48 goats to test their health status. For the same reason we performed a CBC before every checkpoint. Considering that we wanted to measure the response to the artificial photoperiod without any external perturbation, we decided to exclude as much artificial factors as possible. We excluded both the "male effect", do not exposing goats to buck, and the "melatonin effect", avoiding to apply a subcutaneous sponge that releases the hormone. In these natural conditions goats did not restart their cyclic ovarian activity after the SD exposure, and we could not fulfill samplings for checkpoint 4a and 4b.

The on field part ended in May, 2015. Then we collected several tissues and organs from each of the 20 sacrificed goats: hypothalamus, pituitary, epiphysis, ovaries, uterus, thyroid, adipose tissue (visceral and subcutaneous), adrenals. We collected the ovaries in formaldehyde for the histological analysis. Epigenetic analyses involved hypothalamus, pituitary and ovary. Histological analysis of the ovaries confirmed the clinical status of the animal. We decided to sample and to stock at -80$^{o}$C all the other tissues aforementioned because involved in the reproductive pathways; all the additional stored materials could be useful for future molecular analysis.

For the analyses reported in the present thesis, we collected biological samples from 3 Saanen goats and 20 Alpine goats and we sequenced the methylome and the miRNA population of the hypothalamus and other two key reproduction organs, ovary and pituitary. We performed an enrichment of the methylated fragment across the genome and we sequenced the first 15 samples. We optimized the yield of the enrichment of methylated fragments and of the miRNA library. We set up a pipeline for the bioinformatic analysis on 3 Saanen goats, which were used as trials for a complete description of some epigenetic features in the species. Furthermore, we investigated genetic and epigenetic differences that occurred in different physiological moments in goat development. The samples and the type of the analysis performed are described in table 3.

| Breed | Sample type | N of samples per organ | Analysis |
|---|---|---|---|
| **Saanen** | hypothalamus, ovary | 3 | DNA methylation |
| **Saanen** | hypothalamus, ovary, pituitary | 3 | microRNA |
| **Alpine** | hypothalamus | 15 | DNA methylation |
| **Alpine** | hypothalamus | 15 | microRNA |
| **Alpine** | blood | 20 | SNP discovery |
| **Alpine** | blood | 20 | ELISA |

**Table 3** - Summary of samples and analysis performed in the present work.

## 4.3 DNA methylation

For the three trial Saanen goats, biological samples of the hypothalamus and ovaries were collected from each goat. Samples were frozen in liquid nitrogen and ground to a fine powder, using mortar and pestle and stored at -80° C until DNA extraction. Genomic DNA from ovary was isolated using the commercial kit NucleoSpin® Tissue (Macherey-Nagel, Düren, Germany). For hypothalamus, phenol:chloroform genomic DNA extractions were performed. About 50 mg of tissue were resuspended in 300ul of TRIS EDTA lysis buffer (10mM Tris-HCl, 10mM EDTA, NaCl 250m, pH8), added with 15 µl of proteinase K (20 mg/ml) (Sigma-Aldrich., St. Louis, MO, USA) and 15 µl of SDS 10% and incubated at 56°C for 2h. Next, 25 µl of RNase A (20 mg/ml) (Sigma-Aldrich) was added to the suspension and incubated at 56°C for 30 min. DNA was extracted using an equal volume of 1:1 (v/v) phenol:chloroform and precipitated with 1 Vol. of cold Isopropanol. DNA was rinsed with 70% (v/v) cold ethanol, air dryed, resuspended in 30 µl of ultrapure water and stored at -80°C until use. DNA concentration and quality were estimated by PicoGreen® (Thermo Fisher, Waltham, MA USA) and by agarose gel electrophoresis. One µg of genomic DNA from each of the six samples (3 x ovary and 3 x hypothalamus) was sonicated to produce DNA fragments of about 350 bp lengths. Methyl-binding domain (MBD) enrichment was performed using the MethylMiner™ Methylated DNA Enrichment Kit (Invitrogen, Massachusetts, USA), following manufacture instruction.

The MethylMiner™ Kit is designed for the enrichment and fractionation of methylated double-stranded DNA (dsDNA) based on the degree of methylation. Methylated DNA is isolated from fragmented whole genomic DNA via binding to the methyl-CpG binding domain of human MBD2 protein, which is coupled to paramagnetic Dynabeads® M-280 Streptavidin via a biotin linker, as shown in figure 10. The methylated fragments can then be eluted as a single enriched population with a 1000 mM NaCl elution buffer. The high affinity of MethylMiner™ MBD-Biotin Protein for CpG-methylated DNA provides greater sensitivity than antibody binding, while the use of Dynabeads® provides for a simplified, streamlined workflow. The eluted methylated DNA is ready for downstream analysis, in our case a library preparation for high-throughput sequencing or a bisulfite conversion.

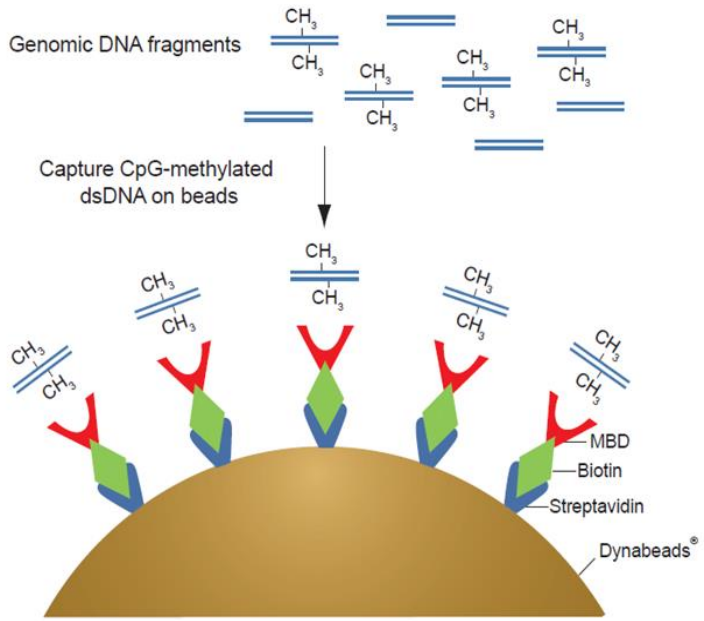


**Figure 10** – Grafical representation of how the MethylMiner™ Methylated DNA Enrichment Kit works (www.lifetechnologies.com)

Library construction for next generation sequencing was performed using the TruSeq® Nano Library prep Kit (Illumina, San Diego, CA, USA). The six samples were then used for cluster generation and subsequent sequencing on a single lane of Illumina Hiseq 2000. 100-base paired-end reads were generated. Raw data obtained from Illumina sequencers was first processed to clean out reads containing adapters, unknown or low quality bases and subsequently mapped to the goat reference genome (CHIR_1.0 GenBank assembly accession: GCA_000317765.1, ftp://ftp.ncbi.nih.gov/genomes/Capra_hircus). The algorithm BWA-SW (Li & Durbin 2010) was used to map sequences to the goat reference genome and the quality control of the sequencing was performed with FastQC (Andrews 2010). ChIPseeqer, a computational framework for the identification of peaks across the genome (Giannopoulou & Elemento 2011), was used for genome-wide detection of the methylation peaks according to the following threshold: t 2 -fold_t 1 -readlen 101 -fraglen 250. The uniquely mapped methylation data was then used for analysis of distribution in goat chromosomes and distribution in different components of the genome (promoters, 3' UTR, exon, intron, intergenic and distal regions). The analysis of the distribution of methylation peaks per chromosome was calculated dividing the number of peaks for the length of each chromosome and the total number of reads per sample per organ. Methylation densities in different components of the genome were compared by calculating the ratio of the length (in bp) of methylated peaks in a particular component to the total area of that region. The CpG islands were scanned by CpG cluster (Hackenberg et al. 2006). CGIs have been defined according to three criteria: length > 200 bp, ≥50% GC content, ≥ 0.6 of CpG

observed/expected. Bedtools v2.25.0 intersect has been used to calculate the overlapping (of at least 90% length/length) between CGIs and methylation peaks.

For the DNA methylation analysis in the hypothalamus of the 15 Alpine goats (belonging to checkpoints 1 and 2a) the same upstream processes has been applied. Organ were pulverized in liquid nitrogen and DNA isolated with NucleoSpin® Tissue (Macherey-Nagel). Methyl-binding domain (MBD) enrichment was performed using the MethylMiner™ Methylated DNA Enrichment Kit (Invitrogen), following manufacture instruction Libraries were generated using the TruSeq® DNA PCR-Free Library Preparation Kit (Illumina) with some modification. Then, we added a key passage in the last phase of the upstream analysis: a bisulphyte conversion of the MBD enrichment has been applied. Because of this new step, others bioinformatics softwares were used for downstream data analysis, as described hereunder. After adapters ligation, samples were converted with EpiTect® Plus Bisulfite Conversion (Qiagen, Venlo, Netherlands). DNA samples were treated with bisulfite mix and subsequently purified according to manufacturer's instruction, and finally PCR amplified with KAPA HiFi Uracil+ (Kapa Biosystems, MA, United States).

The methylation status of a DNA sequence can best be determined using sodium bisulfite. Incubation of the target DNA with sodium bisulfite results in conversion of unmethylated cytosine residues into uracil, leaving the methylated cytosines unchanged. Therefore, bisulfite treatment gives rise to different DNA sequences for methylated and unmethylated DNA. An example of the sequence a short DNA fragment is provided hereunder:

| | Original sequence | After bisulfite treatment |
|---|---|---|
| **Unmethylated DNA** | N-C-G-N-C-G-N-C-G-N | N-U-G-N-U-G-N-U-G-N |
| **Methylated DNA** | N-C-G-N-C-G-N-C-G-N | N-C-G-N-C-G-N-C-G-N |

From a chemical point of view, the conversion of an unmethylated cytosine into uracil is 3 steps reaction summarized in figure 11.

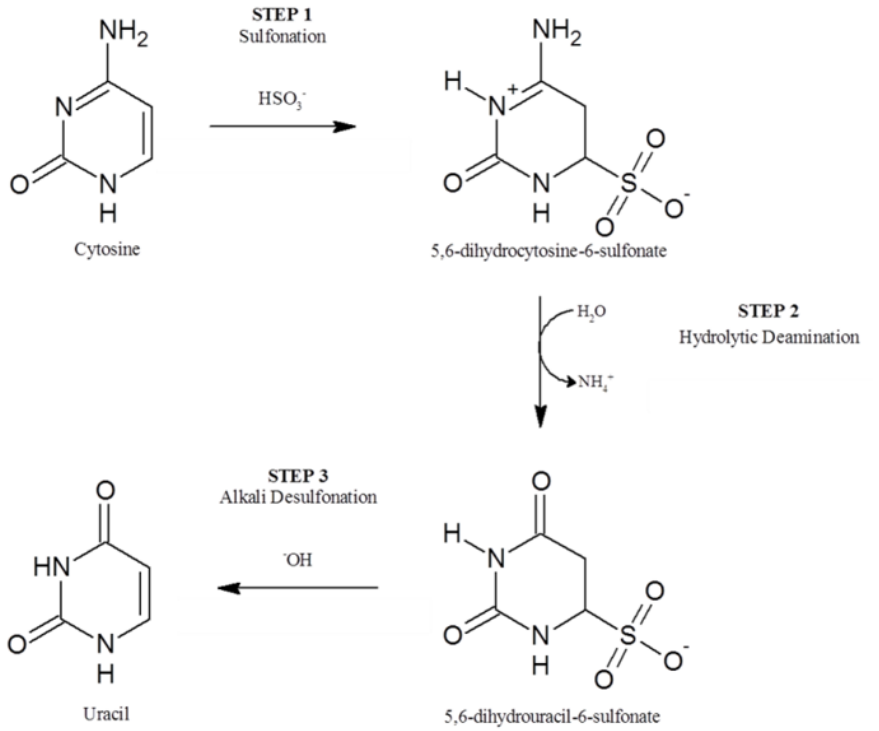


**Figure 11** – Graphical representation of the 3 steps chemical conversion cytosine → uracil. Source: www.epibeat.com

The most critical step for correct determination of a methylation pattern is the complete conversion of unmethylated cytosines. This is achieved by incubating the DNA in high bisulfite salt concentrations at high temperature and low pH. These harsh conditions lead to a high degree of DNA fragmentation and subsequent loss of DNA during purification. A purification step was necessary to remove bisulfite salts and chemicals used in the conversion process that inhibit sequencing. Common bisulfite procedures usually require high amounts of input DNA to compensate for DNA degradation during conversion and DNA loss during purification that often lead to low DNA yield, highly fragmented DNA, and irreproducible conversion rates. Coupling the bisulfite conversion with the DNA methylation enrichment allowed us to overcome this problem.

Whole Genome Bisulfite Sequencing libraries were used for cluster generation and subsequent sequencing on Illumina Hiseq 2000. Bismark software v.0.14.3 (Babraham Bioinformatics) was used to align each read to a bisulfite-converted goat genome and to simultaneously call for cytosine methylation. Seqmonk (Babraham Bioinformatics), a tool to visualize and analyze high throughput mapped sequence data, was used for bisulphite DNA methylation data analysis.

## 4.4 RNA sequencing

Transcriptome analysis was performed on the three Saanen goats and on the fifteen Alpine goats puberty-related.

Total RNA, for the RNAseq analysis, was extracted from each organ (hypothalamus and ovary) with Trizol (Invitrogen, Carlsbad, CA) and purified by NucleoSpin® miRNA kit (Macherey-Nagel, Germany), following the protocol in combination with TRIzol® lysis with small and large RNA in one fraction (total RNA). RNA concentration ng/µl and quality RNA Integrity Number RIN was determined Agilent 2100 Bioanalyzer (Santa Clara, CA). RNA extract was stored at -80°C until use.

Two µg of total RNA from each of samples (3 ovary and 3 hypothalamus for Saanen, 15 hypothalamus for Alpine) was used for library construction. Libraries were generated using the TruSeq® RNA Sample Preparation v2 Illumina kit according to manufacturer's instructions using poly(A) enrichment. The samples were then used for cluster generation and subsequent sequencing on a single lane of Illumina Hiseq 2000 (San Diego, CA). 100-base paired-end reads were generated.

Raw data was trimmed using Trimmomatic, a pre-processing tool for the trimming of paired-end data (Trimmomatic 2014). Sequences were aligned to the goat reference genome (CHIR_1.0 GenBank assembly accession: GCA_000317765.1, ftp://ftp.ncbi.nih.gov/genomes/Capra_hircus) using STAR_2.3.0 (Dobin et al. 2013). Subsequently, HTSeq-count (Anders et al. 2015) was utilized to count aligned sequences to coding regions. The software package EdgeR of Bioconductor (Robinson et al. 2009) was used to calculate differential expression in hypothalamus and ovary. Pathway analysis on the FDR more expressed and more methylated genes for both organs was performed through STRING 10.0 (Szklarczyk et al. 2015).

## 4.5 MicroRNA

For the three adult female Saanen goats and for the twenty Alpine goats analytical procedures has been as follow: we collected biological samples of the hypothalamus, pituitary and ovaries for each goat. The samples were frozen in liquid nitrogen and ground to a fine powder, using mortar and pestle and stored at -80° C until RNA extraction. Total RNA was isolated from each sample using Trizol (Invitrogen, Carlsbad, CA) and purified by NucleoSpin® miRNA kit (Macherey-Nagel, Germany), following the protocol in combination with TRIzol® lysis with small and large RNA in one fraction (total RNA). RNA concentration and quality was determined by Agilent 2100 Bioanalyzer (Santa Clara, CA). The isolated RNAs were stored at -80°C until use. SncRNA libraries were generated using the TruSeq Small RNA Library Preparation kits respectively according to manufacturer's instructions (Illumina). The libraries were then pooled together and purified on a Pippin Prep system (Sage Science, MA, USA) to recover two fractions: the 125nt and 167nt fraction 1 containing mature miRNAs. The quality and yield after sample preparation was measured with an Agilent 2200 Tape Station, High Sensitivity D1000. The libraries obtained were quantified by Real Time PCR with KAPA Library Quantification Kits (Kapa Biosystems, Inc. MA, United States). Libraries were sequenced on one Miseq desktop sequencer (Illumina) run with 50-base single reads. MiRNA detection and discovery was carried out with Mirdeep2 on Illumina high quality trimmed sequences. *Capra hircus* miRNAs available at MirBase (http://www.mirbase.org/) were used to accomplish known miRNA detection on the trimmed sequences. Known miRNAs from related species (sheep, cow and horse) available at MirBase were also

input into Mirdeep2 to support the individuation of novel miRNAs. The Mirdeep2 quantifier module was used to quantify expression and retrieve counts for the detected known and novel miRNAs. Differential expression analyses between the three organs were run with the Bioconductor edgeR package (Robinson et al. 2009). MiRNA cluster analysis was performed with Genesis to identify and visualize pattern within the dataset (Sturn et al. 2002). For miRNA validation, RNA samples isolated from each organ were retro-transcribed with miScript II RT Kit following to manufacturer's instructions (Qiagen, Inc., Valencia, CA USA). Quantitative Real Time RT-PCR was carried out on cDNAs with 7900HT Fast Real-Time PCR System (Applied Biosystems, Carlsbad, California, USA). Reactions were done in 10µl volumes containing 1 M of each primer, 2µl cDNA (see above), and 5µl 2× Power SYBR® Green PCR Master Mix (Applied Biosystems) according to manufacturer protocols. The primers used for chi-miR-141, chi-miR-7, chi-miR-9-5p and chi-miR-10a-5p quantification, were designed using miRprimer software (Busk 2014). For miR-124a1, and the reference small nucleolar snoRNA, C/D Box 95 SNORD95 quantification, specific miScript Primer were utilized in combination with miScript Universal Primer (Qiagen, Inc., Valencia, CA USA). Negative controls using water in place of sample were performed alongside each reaction. Reactions were run using the cycling parameters of 95°C for 10 min, 40 cycles of 95°C for 15 s, and 60°C for 1 min. Relative expression levels were calculated and significance for each treatment separately using the 2-Ct method (Livak & Schmittgen 2001).

## 4.6 Hormonal analysis

ELISA test (Enzyme-Linked ImmunoSorbent Assay) were conducted on 20 female Alpine goats. 7 chronological points were established, as shown hereunder, in order to compare the pattern of the hormones studied (melatonin, T3, T4) in the crucial moments of the reproductive cycle. These crucial moments were pointed out by monitoring the level, in plasma, of a fourth hormone, the progesterone (P4), considering a threshold of 1 ng/ml as an indicator of the presence of the corpus luteum. Based on all these indications the most suitable goats for each group were chosen. The consistency and characteristics of each group are expressed above:

*a*) 5 pre-pubertal goats.

*b*) first 5 goats showing heat + 5 still non-ovulating goats as control.

*c*) last 5 goats showing heat ("late-puberty" group).

The 5 animals selected for point *a* were monitored in the following steps:

*a1*) one week before the application of artificial Long Day.

*a2*) one week after the application of artificial Long Day.

*a3*) at the arrest of the cycle in response to the Long Day.

*a4*) one week before the lights off.

For the measurement of the plasmatic level of melatonin, T3 and T4 ELISA competitive test were performed. For melatonin Blue Gen Goat MT ELISA kit (BlueGene Biotech, Shangai, China) was used. For T3 and T4 Cusabio Goat T3 ELISA kit and Cusabio Goat T4 ELISA kit (distributed by Aurogene Srl). The kit used were characterized by different sensitivity. In particular:

- 1,0 pg/ml melatonin kit
- 0,5 ng/ml T3 kit
- 20 ng/ml T4 kit

## 4.7 SNP discovery

Blood samples were collected from the 20 goats selected from the on-field step (points *a*, *b*, *c* corresponding to checkpoint 1, 2a and 2b). DNA extraction was carried out from 200 µl of frozen whole blood using the commercial kit ReliaPrep Blood gDNA Miniprep System (Promega, Wisconsin, United States), according to the manufacturer's instructions. The concentrations obtained were around 50 ng/µl.

SNP detection was performed on *MTNR1A*, *TSHB* and *DIO2* genes by PCR amplification and sequencing. Amplified regions, primers and annealing temperatures are reported in the table 4. Primers for PCR reaction were drawn using the Primer3 software (Untergasser et al. 2012). The PCR reaction mix (20µl) comprised: 1µl of gDNA, 5X PCR Buffer (Promega), 5mM MgCl2, 0.4µl of each primer, dNTPs each at 5 mM, 0.2µl of Taq DNA Polymerase (Promega). PCR products were purified using the commercial kit Wizard® Genomic DNA Purification Kit (Promega) and sequenced in outsourcing (http://www.bmr-genomics.it/).

| GENE | Exon | PRIMER Forward | PRIMER Reverse | $T_m$ (°C) | Product size (bp) |
|---|---|---|---|---|---|
| ***MTNR1A*** | 1 | CATCTTGAAAGGCGTGGTGT | | 58.8 | 179 |
| | | | CTGCGTTCCTCAGCTTCTTG | 59.2 | |
| | 2 | GTGGTGAGCCTGGCAGTT | | 59.8 | 817 |
| | | | ATGGAGAGGGTTTGCGTTTA | 59.6 | |
| ***TSHB*** | 2 | TTTTGGCCTTGCATGTGGAC | | 59.6 | 106 |
| | | | GGTGGTGTTGATGGTTAGGC | 58.8 | |
| | 3 | GGCAAGCTGTTTCTTCCCAA | | 59.0 | 153 |
| | | | ACACTTGCCACACTTACAGC | 58.7 | |
| ***DIO2*** | 1 | TTTCTCCAACTGCCTCTTCC | | 59.4 | 149 |
| | | | GGAAGCTTTTCCAGATGCAG | 60.0 | |
| | 2 | CAGGTGAAACTGGGTGAAGA | | 58.7 | 169 |
| | | | AGCCGAAGTTGACCACCAGT | 62.1 | |
| | 3 | TTCAGCAAACTGGTGGAAGA | | 59.4 | 187 |
| | | | GCAAGGAGAAACGCTCCAG | 61.1 | |

**Table 4**- Primer sequences for SNP discovery

The alignment of sequences was carried out using the software BioEdit v7.2.5 (Ibis Therapeutics - http://www.mbio.ncsu.edu/bioedit). Chromatograms of the sequences has been visualized with the software Chromas Lite 2.4.4 (Technelysium Pty Ltd) to identified the presence of heterozygous peaks.

## 4.8 Statistical analysis

P-values in tables and figures were calculated with a Student's t-test, unless otherwise noted. The Jackknifing simulation test based on random selection of genes was performed using an in-house R script. For every match (4 in total), 100,000 simulations were tested.
For statistical descriptive analysis on concentrations ANOVA and non-parametric tests were performed using the statistical software JMP of the SAS Institute (JMP®9.0.2 SAS Institute.inc 2010).

For differential expression quasi likelihood (QL) F test method was used. The QL F-test is preferred as it reflects the uncertainty in estimating the dispersion for each gene. It provides more robust and reliable error rate control when the number of replicates is small.

Q-values are the name given to the adjusted p-values found using an optimized False Discovery Rate FDR approach

# 5 - Results

Considering the lack of information at the epigenome-wide level in the species at the moment, we decided to develop some trials for optimizing the pipelines for the analysis in the laboratory and most of all, for giving a first picture of the epigenetic landscape in the species. We analyzed epigenetic features from three adult female Saanen goats reared in the same farm and at the end of their productive career. We performed all the analyses planned for every checkpoint of our experimental plan, that dealt with puberty and seasonality in goat species. The preliminary results obtained were so promising that we decided to keep on analyzing data for a complete characterization of the DNA methylation and microRNA profile in *Capra hircus*. According to the original experimental plan, we evaluated the DNA methylation pattern in two specific organs, the hypothalamus and the ovary, which play a key role in the regulation of seasonality. Concerning the microRNA population analysis, we added another peculiar organ in the control of physiology reproduction, the pituitary. In the second part of this section (from paragraph 5.2 on), according with the original experimental plan, we analyzed DNA methylation pattern and miRNA population in the hypothalamus of Alpine goats across puberty. Furthermore, a SNP discovery on three candidate genes for reproductive seasonality and a hormonal characterization profiles were conducted.

## 5.1 Genome-wide DNA methylation in hypothalamus and ovary

Methyl binding domain sequencing was performed on two reproduction-associated organs (hypothalamus and ovary) of three Saanen goats. Following quality control and trimming, approximately 30 million reads per sample remained, of which 98% were mapped to the goat reference genome (CHIR_1.0 GenBank assembly accession: GCA_000317765.1) and subsequently analyzed (Table 1). A clonal fraction of sequence reads was removed from each sample, resulting in 65.7% (of the total mapped reads) uniquely mapped reads for the hypothalamus and 71.1% for the ovary, with a high quality read alignment against the goat reference genome version 1.0 (2.64 Gb). Highly methylated regions (HMRs), identified as peaks, were detected among uniquely mapped reads. A total of 382,850 methylation peaks were found in the hypothalamus and 413,010 in the ovary.

**Table 5** - Data generated by MBD-seq. Genome coverage is the percentage of bases mapped by genome-wide reads.

| Sample | Organ | Total number of raw sequence reads | Percentage of mapped reads in total reads | Percentage of genome coverage |
|---|---|---|---|---|
| **Goat 1** | Hypothalamus | 31,842,637 | 98.16 | 29.17 |
| | Ovary | 35,390,803 | 97.78 | 35.59 |
| **Goat 2** | Hypothalamus | 23,765,604 | 98.02 | 26.75 |
| | Ovary | 37,185,292 | 98.35 | 40.86 |
| **Goat 3** | Hypothalamus | 27,090,852 | 98.05 | 27.96 |
| | Ovary | 25,704,845 | 97.89 | 20.63 |

Peaks were distributed across all chromosomes and covered approximately 28% of the goat genome in the hypothalamus and 32 % in the ovary (Table 5). Of the 30 chromosomes in the goat genome, 27 chromosomes had a significantly ($P<0.05$) different number of methylation peaks between hypothalamus and ovary (Figure 12). Fifteen chromosomes were significantly more methylated in hypothalamus and twelve chromosomes has more methylation peaks in ovary. Goat chromosome (CHI) X showed the greatest difference in methylation between the two organs (P=4.45E-12), with a higher level of methylation in ovary. Only CHI 1, 17 and 29 did not show significant difference in methylation between the two organs.

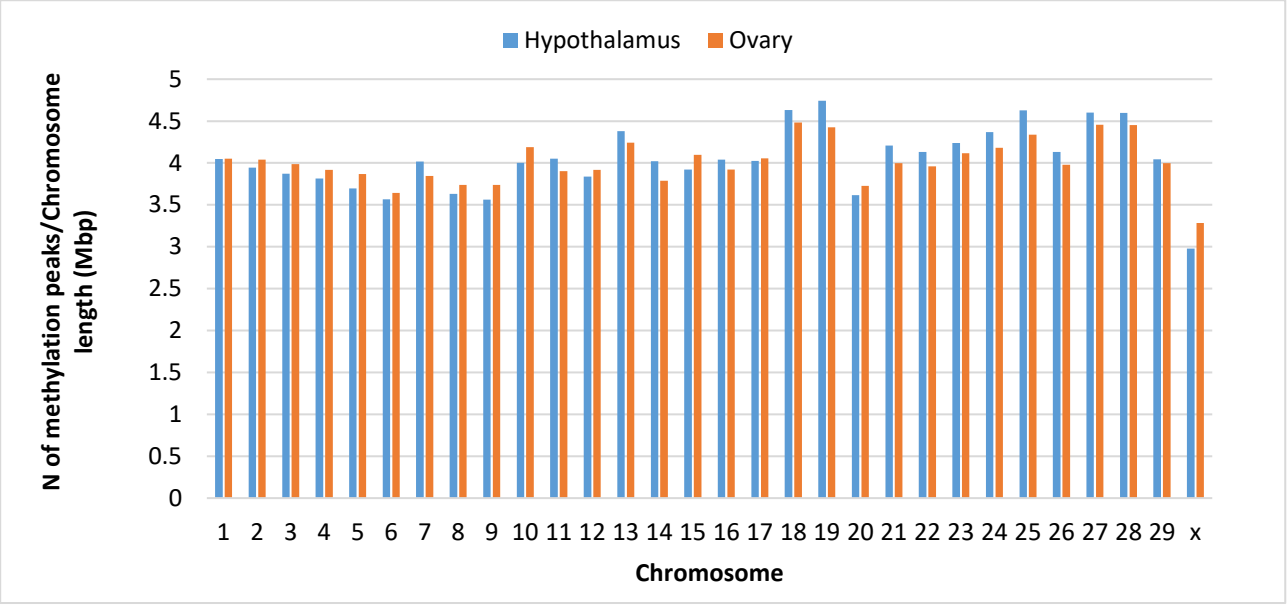


**Figure 12 -** Chromosome distribution of methylation peaks in goat hypothalamus and ovary. The number of methylation peaks per chromosome on the y axis is corrected by chromosome length in Mbp and by the total number of peaks per organ.

Methylation peaks identified in the two organs were connected to five genomic regions, according to where they occurred with respect to the CHIR1.0 reference annotation. The highest density of HMRs was observed in exons and promoters for both organs. Conversely, HMR density decreases in introns, downstream, distal and intergenic regions (Figure 13). Methylation distribution across genes and gene boundaries (2kb upstream region, intragenic and 2kb downstream region) was investigated. The DNA methylation level was found to increase sharply at the start of the coding sequence (CDS) and continued to increase until the end of the CDS. Subsequently, the DNA methylation level drops dramatically and remains steadily low within the 3' boundary (Figure 14). Some differences were observed between the two organs, with the hypothalamus showing a higher methylation level within the upstream and downstream regions with respect to the ovary.

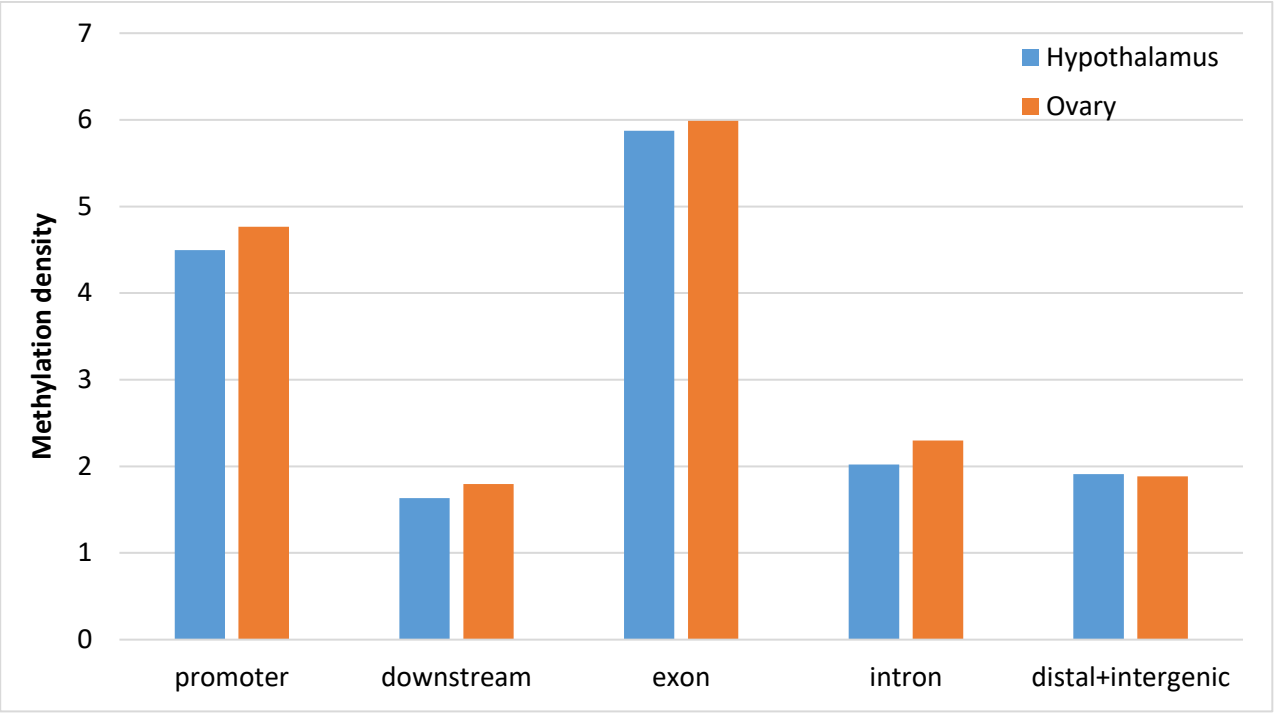


**Figure 13** - Methylation density in different genomic regions in hypothalamus and ovary. Methylation density within gene regions was calculated by dividing the peak length in the region by the average length (in bp) of the area in that genome and by the total number of peaks per organ.

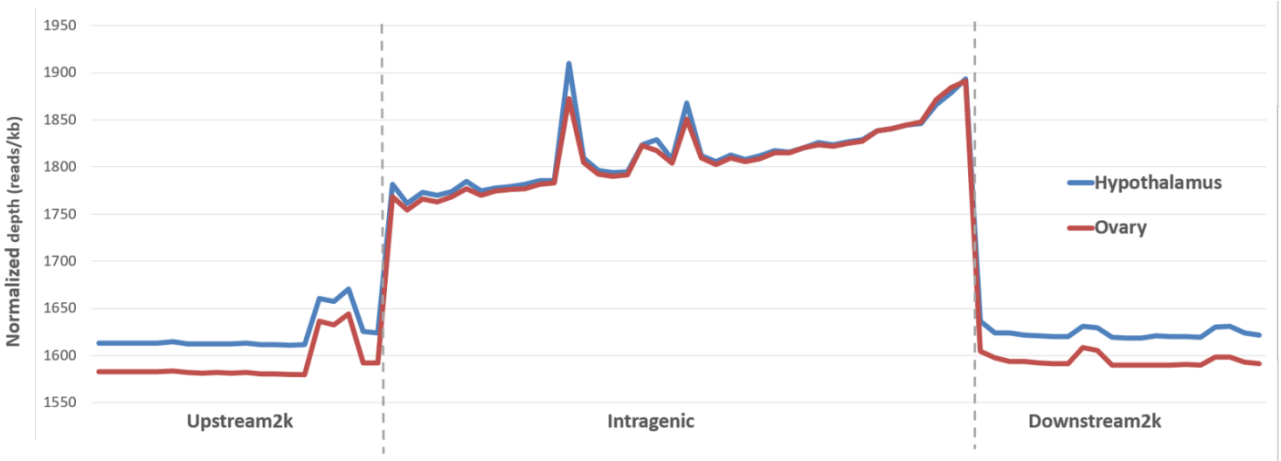


**Figure 14** - DNA methylation distribution in goat gene region. The DNA methylation profile in the gene region of hypothalamus and ovary was shown by the reads that were aligned on the unique locus in the genome. The regions including a 2 kb region upstream of the Transcription Start Site (TSS), the gene body from TSS to TTS, and a 2 kb region downstream of the Transcription Termination Site (TTS) have been considered different gene regions. In the upstream and downstream 2 kb regions, the regions were split into 20 non-overlapping windows (100 bp long), and the average alignment depth was calculated for each window. In the gene body, each gene was split into 40 equal windows, and the average alignment depth was calculated for each window. The y-axis is the average of the normalized depth for each window.

To investigate methylated CGIs the goat reference genome version 1.0 was used to identify CGIs. A total of 71,367 CGIs, representing 0.86% of the whole genome, were identified. Approximately 45.80% (n

=32,683) of the CGIs in hypothalamus were found among the HMRs and 46.17% (n = 32,952) in ovary. Therefore, a slightly higher CGIs methylation level (+0.37%) was observed in ovary compared to hypothalamus (Figure 15). Of the thirty chromosomes in the *Capra hircus* genome, thirteen show a significant difference in methylation (P<0.05) between hypothalamus and ovary. Among these, four chromosomes (X, 7, 11 and 15) show a higher DNA methylation level in the CGIs in the ovary and the remaining nine chromosomes (3, 14, 10, 21, 23, 5 and 25) show higher DNA methylation levels in hypothalamus. The greatest difference in CGIs methylation between the organs was observed on CHI X. *Capra hircus* chromosome 6 had the same number of methylated regions in CGIs in both tissues. Comparing the methylation level found in CGIs to that observed for CpGs, CHI X, which was less methylated considering the global level of DNA methylation, became one of the most methylated within only CGIs. Conversely, for CHI 18 and 19 the most methylated chromosome for both organs, showed two of the lowest level of methylated CGIs.

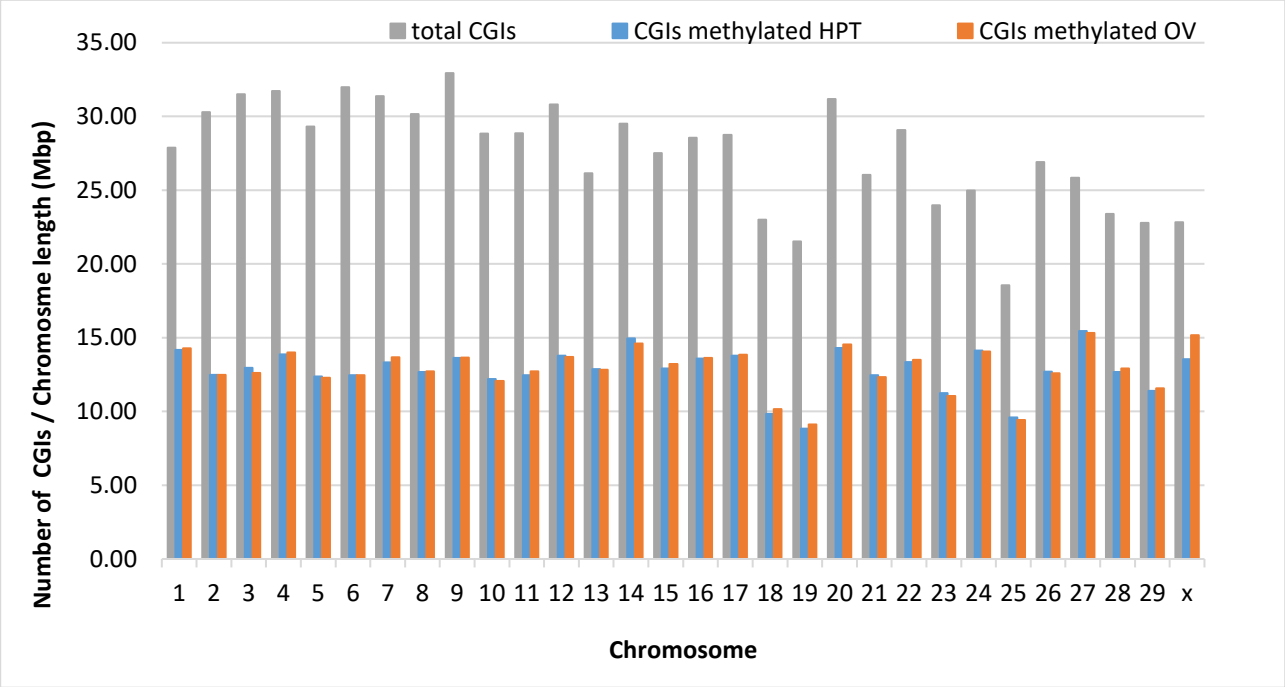


**Figure 15** - Genomic distribution of total and methylated CpG islands. Green bars show the total number of CGIs identified in the goat genome; blue and red bars indicate the number of the methylated CGIs in hypothalamus and ovary respectively. The number of CGIs per chromosome on the y axis is corrected by chromosome length in Mbp.

The difference in the methylation level between hypothalamus and ovary was measured considering the respective differentially methylated genomic regions. The analysis revealed 4,808 differentially methylated regions (DMRs) (FDR<0.05), among which, 1,547 were significantly more methylated in the hypothalamus and 3,261 in the ovary. A total of 2,651 DMRs were located within the gene body, gathered in 1,264 differentially methylated genes (DMGs) in ovary and 456 in hypothalamus. In addition, 74 highly methylated genes were in common between hypothalamus and ovary (with DMRs in gene bodies), for a total of 1,646 DMGs identified.

### 5.1.1 Differentially Methylated Genes and Differentially Expressed Genes

RNA-seq analysis was performed on the same samples to evaluate the variation of gene expression between the two organs. Approximately 330 million paired end (PE) reads were produced per sample which were assembled onto 13,686 unique genes identified from both organs. A total of 7,173 differentially expressed genes (FDR<0.05) were identified, of which 2,665 had significantly higher expression in the hypothalamus and 4,508 in the ovary.

DMGs and DEGs were compared in order to evaluate possible interactions between gene methylation and expression. Different lists of DEGs were selected based on three different FDR values (5,512 DEGs for FDR<0.01, 2,722 for FDR<0.0001, 1,424 for FDR <0.000001) and compared with DMGs list (1,646 genes) these comparisons identified 620, 349 and 215 shared genes between DEGs and DMGs, respectively. A simulation test based on random selection of loci (jackknifing) showed that the number of shared genes was higher than random expectation and the significance (P-value) was inversely proportional to the FDR values for the thresholds considered (Table 6).

To explore the potential regulatory role of methylation on gene expression, a list of DMGs and DEGs (1,646 DMGs with FDR<0.05 and 2,722 DEGs with FDR<0.0001) were divided accordingly to their level of gene methylation and expression, respectively. Thus, four subsets were obtained (hyper and hypo-methylated genes versus over and under-expressed genes) which were cross checked to obtain lists of shared genes (DEGs ∩ DMGs). In hypothalamus, a negative correlation was found between gene expression and hypo-methylation of 231 genes (P=1.29E-46). There was also a significant negative correlation (P=0.0167) between hyper-methylated genes and gene expression. However, considering positive correlations, hyper-methylated genes were not correlated with overexpressed genes. Conversely, the number of genes in common between hypo-methylated and under-expressed genes was less than half of the 100k simulation test (42 observed versus 103 expected). In this case, the significant correlation (P=1.97 E-10) suggested an opposite (negative) correlation between hypo-methylated genes and those under-expressed (Table 7).

| **DEGs thresholds** | **n° of DEGs** | **n° of DMGs (FDR<0.05)** | **DEGs ∩ DMGs** | **100K simulations* of ∩** | **± SD** | **P-value** |
|---|---|---|---|---|---|---|
| **FDR<0.01** | 5512 | 1646 | 620 | 496.41 | 17.67 | 9.55E-12 |
| **FDR<0.0001** | 2722 | 1646 | 349 | 245.22 | 13.78 | 1.92E-13 |
| **FDR<0.000001** | 1424 | 1646 | 215 | 128.24 | 10.41 | 3.20E-16 |

**Table 6 -** Intersections and simulations between DEGs and DMGs

***100k simulations** refer to the jackknifing test used for the random crossing of DEGs and DMGs.

| Hypothalamus | Obs. DMGs | Obs. DEGs | DEGs ∩ DMGs | 100k simulation of ∩ | SD | P-value |
|---|---|---|---|---|---|---|

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Positive Correlation | **Hyper DMGs/over DEGs** | 455 | 1138 | 34 | 30.46 | 5.25 | 0.318 |
| | **Hypo DMGs/under DEGs** | 1263 | 1583 | 42 | 103.44 | 9.38 | 1.97E-10 |
| Negative Correlation | **Hyper DMGs/under DEGs** | 455 | 1583 | 49 | 34.89 | 5.60 | 0.017 |
| | **Hypo DMGs/over DEGs** | 1263 | 1138 | 231 | 99.10 | 9.11 | 1.29E-46 |

**Table 7** - Correlation between DMGs and DEGs. Hypo and hypermethylated genes (FDR<0.05) overlapping with under and overexpressed genes (FDR<0.001) in the hypothalamus (DEGs ∩ DMGs). A jackknifing simulation test with relative P-values (observed vs random expectation) was performed for each subgroup.

Functional annotation analysis using STRING v10.0 was performed for shared genes between DMGs and DEGs in hypothalamus and ovary in order to investigate the pathways involved.

Biological pathway analysis for hypo-methylated genes that were overexpressed in hypothalamus, specified several pathways characterizing the organ. 114 significant pathways (FDR<0.0001) were found in hypo-methylated genes mainly concerning synaptic transmission and neuron morphogenesis and development, with almost 200 gene involved. Cellular component analysis in the same genes revealed 44 significant pathways (FDR<0.0001) affecting synapses, neuron and axons. The most significant pathways for both categories are shown in Table 8 and 9. No significant pathways (FDR<0.01) were detected for hyper-methylated and overexpressed genes for both biological pathways and cellular component analysis.

| **Biological Process (GO)** | | | |
|---|---|---|---|
| **pathway ID** | **pathway description** | **count in gene set** | **FDR** |
| **GO:0007399** | nervous system development | 69 | 4.77e-17 |
| **GO:0022008** | neurogenesis | 51 | 2.54e-12 |
| **GO:0048699** | generation of neurons | 48 | 2.1e-11 |
| **GO:0048731** | system development | 81 | 9.99e-11 |
| **GO:0007275** | multicellular organismal development | 83 | 1.78e-08 |
| **GO:0030182** | neuron differentiation | 35 | 1.36e-07 |
| **GO:0044700** | single organism signaling | 90 | 2.31e-07 |
| **GO:0048666** | neuron development | 31 | 2.31e-07 |
| **GO:0048856** | anatomical structure development | 80 | 2.31e-07 |
| **GO:0031175** | neuron projection development | 28 | 2.33e-07 |

**Table 8 -** Top 10 pathways of biological processes in hypo-methylated genes overexpressed in hypothalamus

On the other hand, considering under-expressed genes in hypothalamus (and consequently overexpressed in ovary) there were no significant pathways for hypo-methylated genes. 20 significant biological pathways (FDR<0.01) and 11 pathways for cellular component (FDR<0.01) were found for hyper-methylated genes, mainly concerning extracellular matrix, system and tissue development and collagen processes.

As for microRNAs (see below), we tried to analyze the DNA methylation profile of the pituitary. Despite suitable DNA concentration for downstream analysis, sequencing data we not readable after the Hi seq run. So we decided to focus our efforts on the hypothalamus of Alpine goats.

| Cellular Component (GO) | | | |
|---|---|---|---|
| **pathway ID** | **pathway description** | **count in gene set** | **FDR** |
| **GO:0045202** | synapse | 42 | 5.62e-19 |
| **GO:0043005** | neuron projection | 44 | 3.82e-16 |
| **GO:0097458** | neuron part | 49 | 5.46e-16 |
| **GO:0044456** | synapse part | 33 | 1.85e-14 |
| **GO:0042995** | cell projection | 54 | 2.4e-12 |
| **GO:0030054** | cell junction | 41 | 1.26e-10 |
| **GO:0030425** | dendrite | 26 | 1.26e-10 |
| **GO:0098794** | postsynapse | 23 | 4.03e-10 |
| **GO:0097060** | synaptic membrane | 19 | 5.68e-09 |
| **GO:0071944** | cell periphery | 89 | 9.34e-09 |

**Table 9** - Top 10 pathways of cellular components in hypo-methylated genes overexpressed in hypothalamus

## 5.2 MicroRNA population in hypothalamus, pituitary and ovary

In order to better understand molecular pathways involved in HPG axis regulation, we evaluate the transcriptomic and miRNA profile changes in hypothalamus, pituitary and ovary from 3 adult Saanen goats using Illumina high-throughput technology. In figure 16 we reported the profile of the total RNA fraction of the first 2 trial goats analyzed. The small RNA libraries obtained using the TruSeq technology of Illumina are shown in figure 17.

Miseq sequencing resulted in 12,592,015 total raw sequencing reads, with an average production of 1,399,112 (661,668 and 2,885,386) reads per sample. Bioinformatic analyses of sequenced products were performed using miRDeep2 software, identifying 760 miRNAs in total. These a small non-coding RNA molecule encoded by eukaryotic nuclear DNA function via base-pairing with complementary sequences within mRNA molecules, contributing to their silencing. Among the sequences analyzed, 402 were classified as already known chi-miRNA, 210 found for homology with miRNA from other species and 148 were predicted candidate miRNAs.

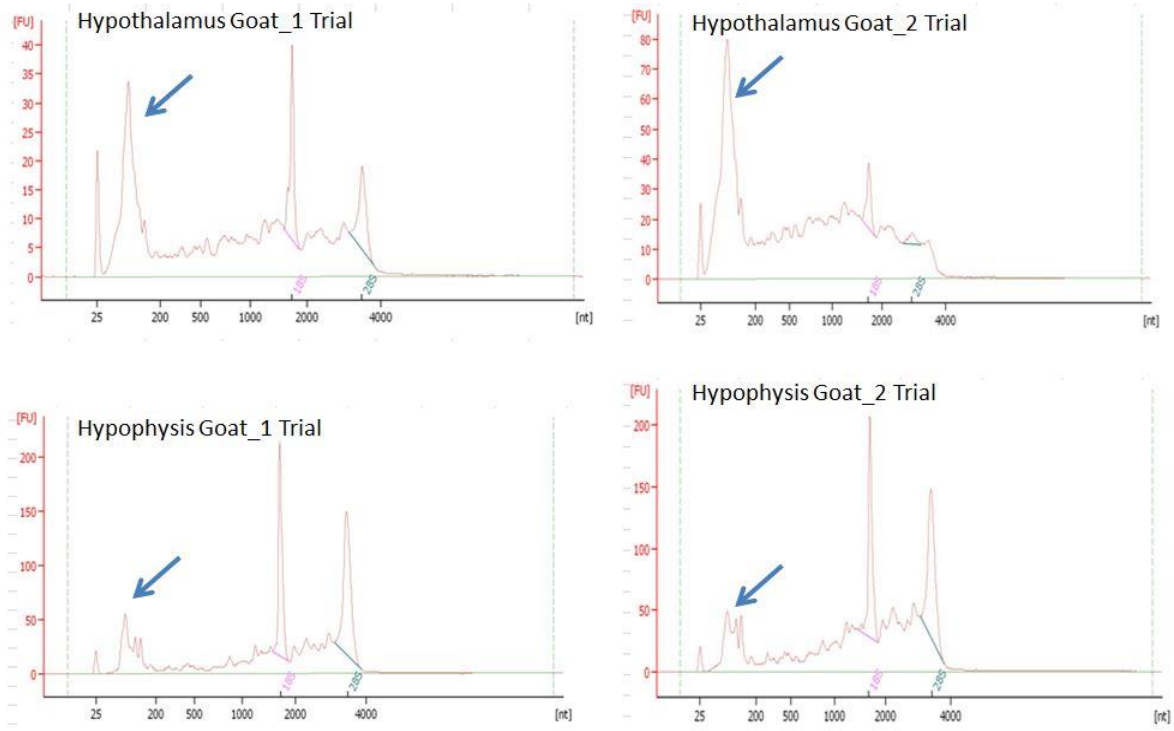


**Figure 16** - Total RNA profiles from hypothalamus and hypophysis of 2 trial goats. Arrows indicate the peak corresponding to smallRNAs fraction for each sample.

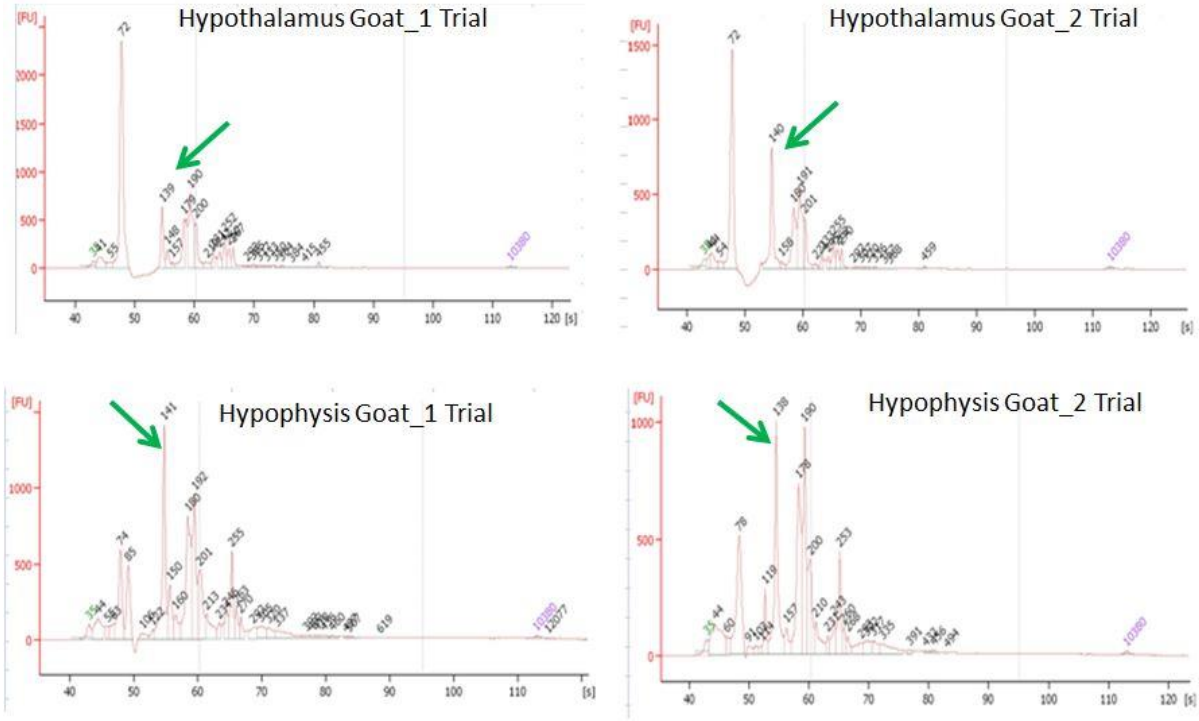


**Figure 17** – Small RNA libraries from hypothalamus and hypophysis (pituitary) of 2 trial goats. The number on the top of each peak indicates the length of the fragment (adapters included) in bp. Arrows indicate peaks corresponding to microRNA population.

After applying a stringent filtering approach (FDR<0.01), we identified 87, 70, 233 miRNAs that were differentially expressed in pituitary, hypothalamus and ovary respectively. The differential expression for each target organ was obtained using EdgeR software and applying the QL F-test, which compared the three groups. This was done by creating a matrix of contrasts, where each column (group) represented a contrast between the other two groups of interest.

In table 10 we reported a selection of miRNAs which are specifically up-regulated in only one organ, and consequently down-regulated in the other two. Applying this criteria, 44, 55, and 69 miRNAs were identified in pituitary, hypothalamus and ovary respectively.

| **Up-regulated MiRNAs** | | |
|---|---|---|
| **Pituitary** | **Hypothalamus** | **Ovary** |
| chi-let-7i-3p | chi-miR-105a | chi-let-7b-5p |
| chi-miR-136-5p | chi-miR-105b-3p | chi-let-7c-5p |
| **chi-miR-141** | chi-miR-105b-5p | chi-miR-1 |
| chi-miR-186-3p | chi-miR-1249 | chi-miR-10a-3p |
| chi-miR-200a | **chi-miR-124a** | **chi-miR-10a-5p** |
| chi-miR-200c | chi-miR-128-3p | chi-miR-10b-3p |
| chi-miR-216b | chi-miR-135b-5p | chi-miR-10b-5p |
| chi-miR-217-3p | chi-miR-147-3p | chi-miR-125b-5p |
| chi-miR-217-5p | chi-miR-187 | chi-miR-130a-3p |
| chi-miR-376b-3p | chi-miR-20b | chi-miR-133a-3p |
| chi-miR-376b-5p | chi-miR-219 | chi-miR-143-3p |
| chi-miR-376c-3p | chi-miR-326-3p | chi-miR-143-5p |
| chi-miR-376c-5p | chi-miR-330-3p | chi-miR-145-3p |
| chi-miR-376e-3p | chi-miR-330-5p | chi-miR-145-5p |
| chi-miR-376e-5p | chi-miR-338-3p | chi-miR-146a |
| chi-miR-379-3p | chi-miR-340-3p | chi-miR-155-5p |
| chi-miR-379-5p | chi-miR-346-5p | chi-miR-195-3p |
| chi-miR-381 | chi-miR-363-3p | chi-miR-195-5p |
| chi-miR-3959-5p | chi-miR-485-3p | chi-miR-196a |
| chi-miR-411a-5p | chi-miR-592 | chi-miR-196b |
| chi-miR-429 | chi-miR-874-5p | chi-miR-199a-3p |
| chi-miR-494 | chi-miR-9-3p | chi-miR-199a-5p |
| chi-miR-582-5p | **chi-miR-9-5p** | chi-miR-199b-3p |
| **chi-miR-7-5p** | Novel:bta-miR-105b | chi-miR-199b-5p |
| Novel:bta-miR-200a | Novel:bta-miR-1249 | chi-miR-199c-3p |
| Novel:bta-miR-200c | Novel:bta-miR-124a | chi-miR-199c-5p |
| Novel:bta-miR-205 | Novel:bta-miR-128 | chi-miR-202-3p |
| Novel:bta-miR-216a | Novel:bta-miR-1298 | chi-miR-202-5p |
| Novel:bta-miR-216b | Novel:bta-miR-135a | chi-miR-21-3p |
| Novel:bta-miR-217 | Novel:bta-miR-138 | chi-miR-214-3p |
| Novel:bta-miR-299 | Novel:bta-miR-147 | chi-miR-214-5p |
| Novel:bta-miR-376b | Novel:bta-miR-187 | chi-miR-21-5p |
| Novel:bta-miR-379 | Novel:bta-miR-219-3p | chi-miR-224-5p |
| Novel:bta-miR-381 | Novel:bta-miR-326 | chi-miR-2284a |
| Novel:bta-miR-411a | Novel:bta-miR-338 | chi-miR-25-3p |
| Novel:bta-miR-494 | Novel:bta-miR-346 | chi-miR-28-3p |
| Novel:bta-miR-582 | Novel:bta-miR-370 | chi-miR-28-5p |

| | | |
|---|---|---|
| Novel:bta-miR-7 | Novel:bta-miR-378d | chi-miR-320-3p |
| Novel:CM001726,1_4406 | Novel:bta-miR-448 | chi-miR-365-3p |
| Novel:CM001727,1_4754 | Novel:bta-miR-95 | chi-miR-449a-5p |
| Novel:CM001731,1_5621 | Novel:bta-miR-9-5p | chi-miR-449c |
| Novel:CM001739,1_7003 | Novel:CM001713,1_1132 | chi-miR-455-3p |
| Novel:eca-miR-376c | Novel:CM001713,1_1134 | chi-miR-455-5p |
| Novel:oar-miR-3959-5p | Novel:CM001717,1_2214 | chi-miR-483 |
| | Novel:CM001727,1_4643 | chi-miR-497-5p |
| | Novel:CM001728,1_4975 | chi-miR-92a-3p |
| | Novel:CM001730,1_5564 | chi-miR-92b |
| | Novel:CM001732,1_5836 | chi-miR-93-5p |
| | Novel:CM001736,1_6537 | chi-miR-99a-3p |
| | Novel:CM001739,1_6911 | chi-miR-99a-5p |
| | Novel:CM001739,1_6915 | Novel:bta-miR-1 |
| | Novel:CM001739,1_6918 | Novel:bta-miR-10a |
| | Novel:eca-miR-3548 | Novel:bta-miR-125a |
| | Novel:eca-miR-485-3p | Novel:bta-miR-143 |
| | Novel:eca-miR-592 | Novel:bta-miR-145 |
| | | Novel:bta-miR-146b |
| | | Novel:bta-miR-151-5p |
| | | Novel:bta-miR-155 |
| | | Novel:bta-miR-196a |
| | | Novel:bta-miR-199a-3p |
| | | Novel:bta-miR-21-5p |
| | | Novel:bta-miR-224 |
| | | Novel:bta-miR-2299-3p |
| | | Novel:bta-miR-31 |
| | | Novel:bta-miR-320a |
| | | Novel:bta-miR-365-3p |
| | | Novel:bta-miR-377 |
| | | Novel:CM001738,1_6736 |
| | | Novel:eca-miR-28-3p |

**Table 10** - Organ specific up-regulated miRNAs. In bold miRNAs tested with qRT-PCR.

Based on most significant differentially expressed miRNAs, a tree with clear distinction of three organs was generated by cluster analysis (Figure 18).

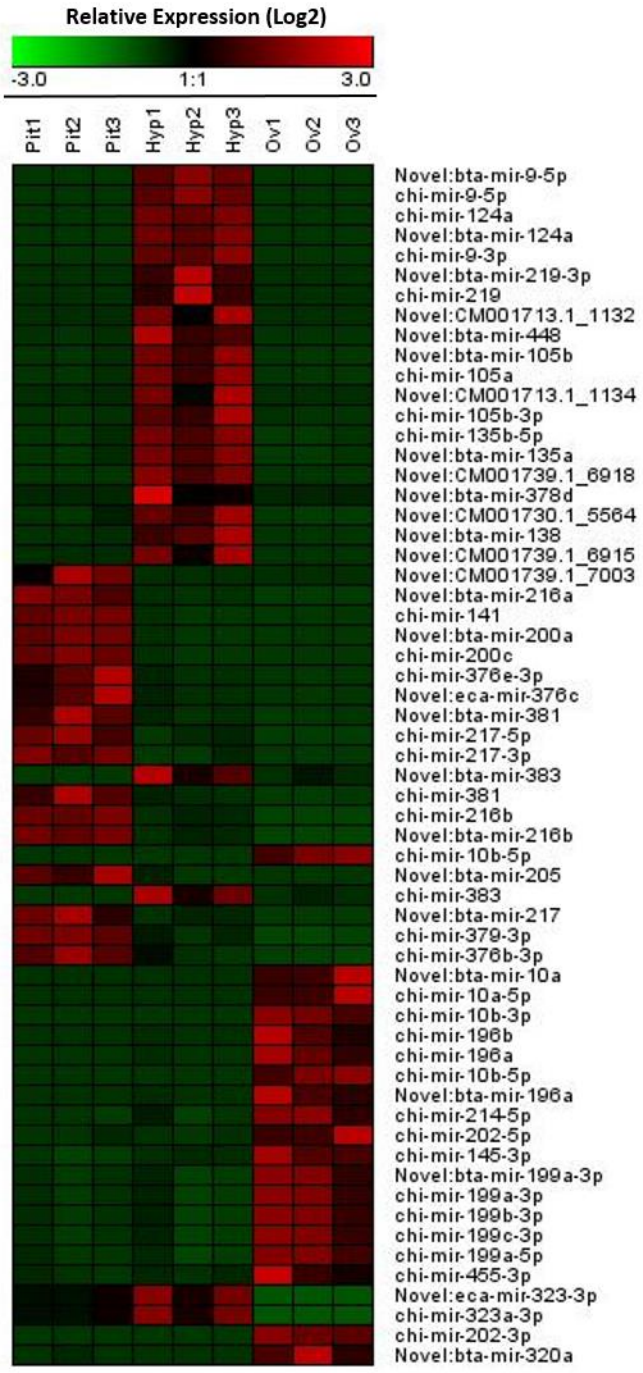


**Figure 18** - Hierarchical clustering of the first 20 most expressed miRNAs in pituitary, hypothalamus and ovary. Red indicates an increase in expression and green a decrease in expression relative to the mean expression of 60 miRNAs.

Differential expression of specific miRNA in each organ was verified by qRT-PCR. miR-141 and miR-7-5p, were higher expressed in pituitary, miR-9-5p and miR-124a in hypothalamus, whereas miR-10a-5p in the ovary. In each case, qPCR results confirmed RNA-seq data, as shown in figure 19.

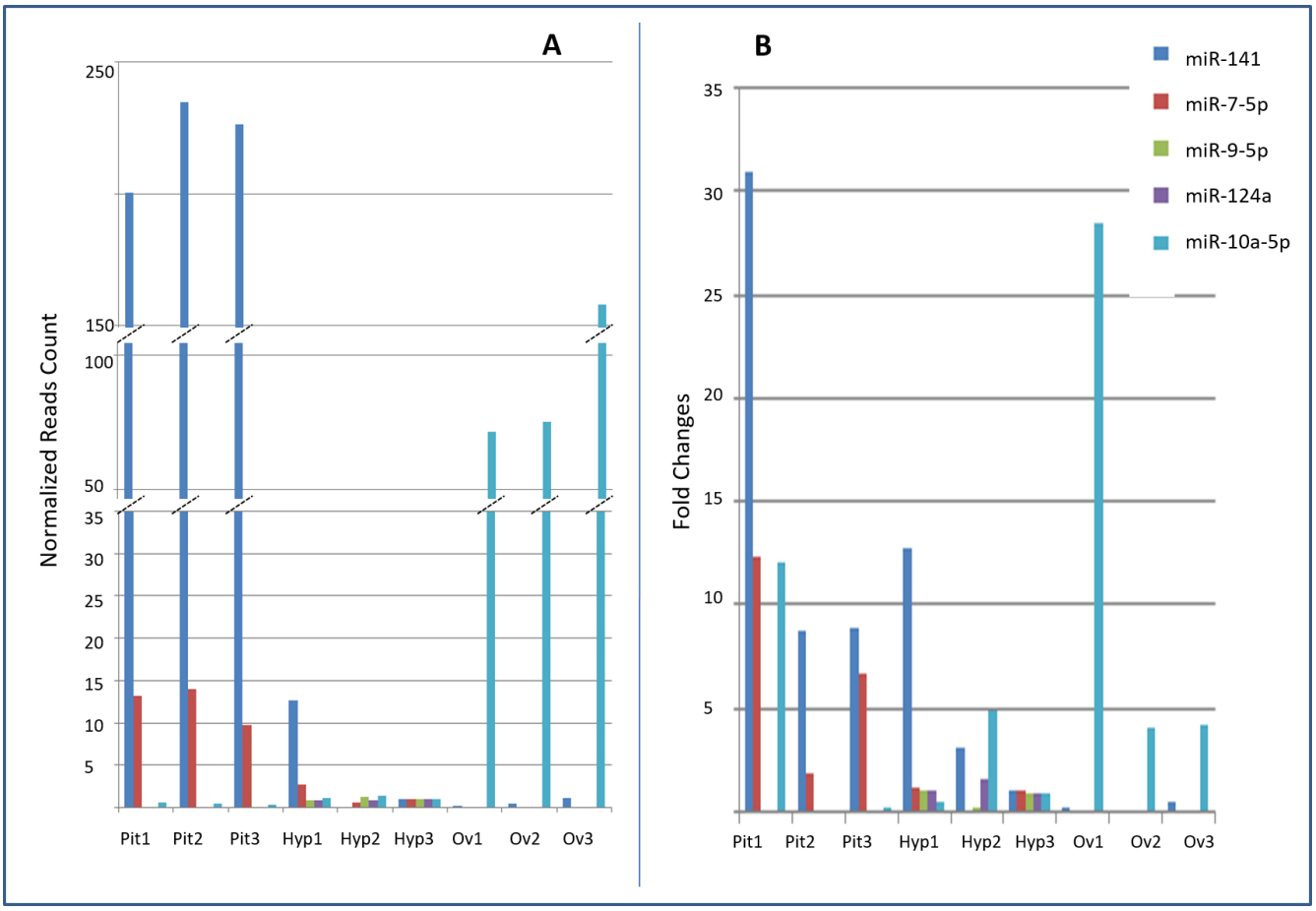


**Figure 19 -** MiRNA expression in Pituitary (Pit), Hypothalamus-(Hyp) and Ovary-(Ov). Values were determined by: A) small-RNA Sequencing; B) Real-Time PCR. In both cases, Hyp3 was used as reference for normalization.

## 5.3 DNA methylation pattern in the hypothalamus of pubertal goats

Once we characterized the methylome and the miRNA population of key-organs in the reproductive process, we moved on analyzing the same epigenetic features in the hypothalamus of specific group of goats across puberty. According with our plan, the experimental sample in the project was originally composed by 55 female Alpine goats.

We estimated a daily ponderal index with a linear regression, considering the weight and the age of our goats. In figure 20 we reported the growth and the growth curve estimated on the measurements recorded. The curve presents a high coefficient of determination, confirming that the variability of data is well explained by the relation:

Weight (kg) = 8.32287 + 0.09233* age (d) – 0.00020*(age (d) – 137.29) ^2

All parameters in the squared regression were significant.

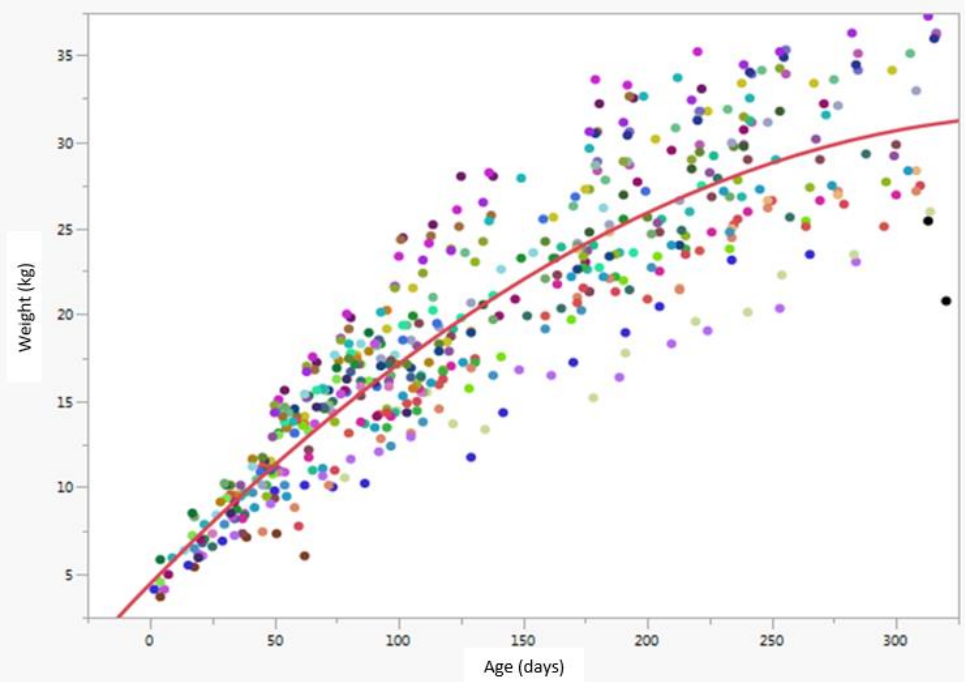


**Figure 20** - Ponderal index of the weight in kg goats (one color for each goat) by their age in days

We considered a threshold of 1 ng/ml of progesterone (P4) as an indicator of the presence of the corpus luteum. In checkpoint 1 all the five goats sacrificed showed a P4 < 0.6 ng/ml. Higher progesterone levels in the luteal phase witness that the heat has occurred in the animal. In checkpoint 2a, 5 out 10 goats showed a P4 concentration between 1.8 and 11.4 ng/ml. The remaining five goats had a P4 concentration < 0.4 ng/ml. In checkpoint 2b all the 5 sacrificed goats showed a P4 concentration between 6.3 and 12.5 ng/ml.

All these activities allow us to choose and sacrifice the most suitable goat for each group in the timetable (figure 1). A proper selection of the animals was fundamental for the downstream molecular analysis. Figure 21 shows an example of anatomical and histological pictures of a pre-pubertal goat and of a goat at puberty reached.

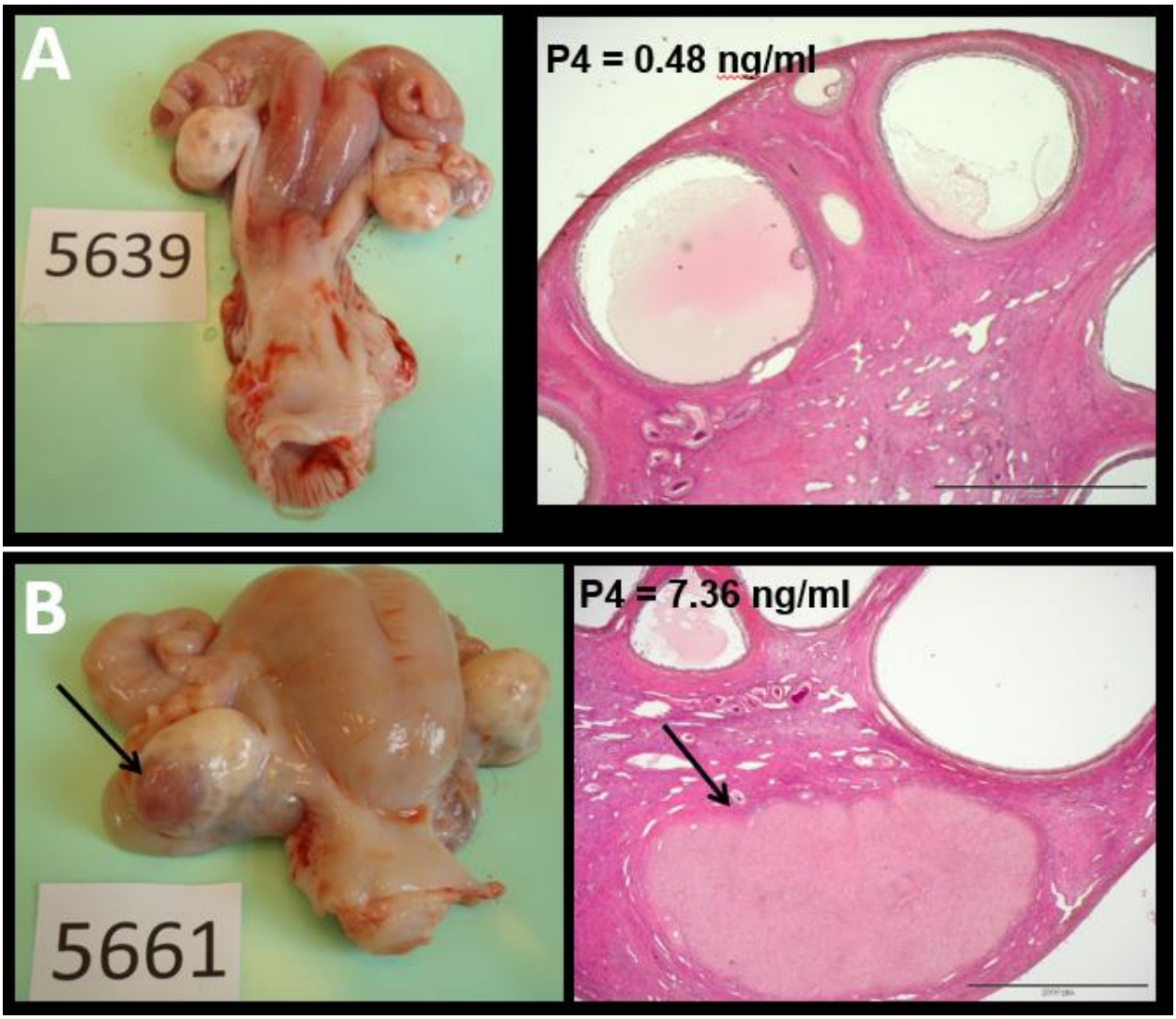


**Figure 21** - Macro-scale and micro-scale of anatomical analysis of uterus. A) Uterus of a pre-pubertal goat. Follicles growing on both ovaries B) Uterus of a goat at puberty reached. Arrows shows the corpus luteum in the left ovary. Progesterone levels are shown for both goats.

We evaluated the difference at the DNA methylation level in the hypothalamus of goats belonging to the first 2 checkpoints of our experimentation. The methylome of 15 Alpine goats, 5 at a prepubertal stage, 5 at their pubertal stage and 5 still prepubescent but at the same age/weight of those at the pubertal stage was analyzed to investigate the differences at the DNA methylation level behind these physiological changes. Details on the mean values for every group are shown in table 11.

In figure 22 it is provided an example of the library obtained for the samples above mentioned after the enrichment of methylated fragments and bisulfite conversion. A CpG methylation profile was obtained: the graph shows a population of methylated DNA fragments (about 603 bp length) in the hypothalamus of a goat at puberty reached.

| | | Age | | Weight | |
|---|---|---|---|---|---|
| | **n** | **Mean (days)** | **±SE** | **Mean (kg)** | **±SE** |
| **Pre-pubertal goats** | 5 | 92.6 | 7.87 | 14.48 | 3.05 |
| **Pubertal goats** | 5 | 229.8 | 7.63 | 29.47 | 2.20 |
| **Control** | 5 | 227.0 | 6.17 | 29.50 | 1.78 |

**Table 11** - Age and weight at the time of suppression for each group

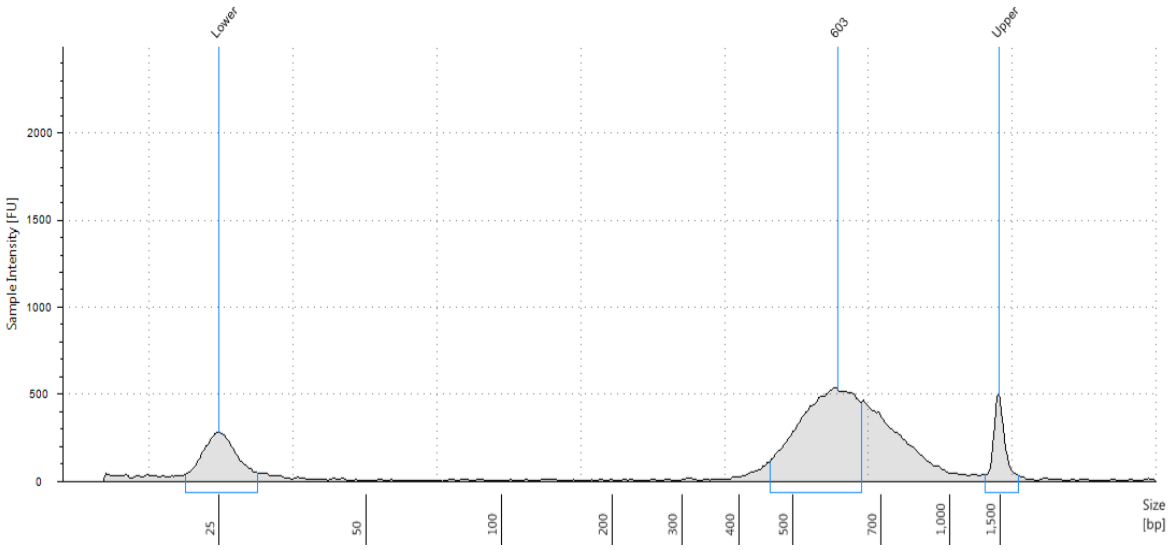


**Figure 22** - Quality and quantity of methylated DNA fragments in the library in hypothalamus of a goat from checkpoint 2a (Agilent Tape Station).

In a similar approach to that used for trial goats, in order to evaluate differentially methylated regions, Methylated DNA Binding Domain sequencing (MBD-seq) with enrichment of methylated DNA fragments coupled with bisulphite conversion and next generation sequencing (Hiseq2000 Illumina) were performed. An average of 21.75 million of reads (range 14.10 and 33.29 million of reads) were produced per sample and peaks corresponding to hyper-methylated regions were estimated using the software Seqmonk. The analysis showed that there was an increase in methylation before puberty.
Analyzing the global level of DNA methylation across the genome, pre-pubertal goats showed the highest percentage of C methylated in a CpG context, if compared to early-puberty goats (90.32% for pre-pubertal vs 85.66% for Early puberty goats). On the other hand, the percentage of C methylated for goats with similar age and weight to the pubertal ones, but still prepubescent, were slightly lower than their counterpart (85.00% for control vs 85.66% for Early puberty goats) (Table 12).

| | Pre-pubertal goats | | | | |
|---|---|---|---|---|---|
| | A1 | A2 | A3 | A4 | A5 |
| Sequence pairs analyzed: | 1,85E+07 | 2,01E+07 | 1,59E+07 | 1,79E+07 | 1,41E+07 |
| N of paired-end alignments unique | 1,03E+07 | 8,78E+06 | 8,25E+06 | 7,76E+06 | 7,21E+06 |
| Mapping efficiency: | 55.5% | 43.6% | 52.0% | 43.4% | 51.2% |
| Number of C's analyzed: | 4,46E+08 | 4,03E+08 | 3,60E+08 | 3,53E+08 | 3,14E+08 |
| Methylated C's in CpG context: | 4,57E+07 | 4,83E+07 | 3,81E+07 | 4,09E+07 | 3,33E+07 |
| Unmethylated C's in CpG context: | 5,03E+06 | 5,09E+06 | 4,11E+06 | 4,28E+06 | 3,61E+06 |
| C methylated in CpG context: | 90.1% | 90.5% | 90.3% | 90.5% | 90.2% |
| | **Early puberty goats** | | | | |
| | A6 | A8 | A10 | A12 | A14 |
| Sequence pairs analyzed: | 1,65E+07 | 1,55E+07 | 2,40E+07 | 2,50E+07 | 1,77E+07 |
| N of paired-end alignments unique | 9,11E+06 | 8,27E+06 | 1,35E+07 | 1,46E+07 | 9,50E+06 |
| Mapping efficiency: | 55.1% | 53.3% | 56.3% | 58.3% | 53.7% |
| Number of C's analyzed: | 3,81E+08 | 3,54E+08 | 5,71E+08 | 6,01E+08 | 4,02E+08 |
| Methylated C's in CpG context: | 3,62E+07 | 3,47E+07 | 5,00E+07 | 5,30E+07 | 3,58E+07 |
| Unmethylated C's in CpG context: | 5,08E+06 | 5,45E+06 | 8,85E+06 | 9,52E+06 | 6,64E+06 |
| C methylated in CpG context: | 87.7% | 86.4% | 85.0% | 84.8% | 84.4% |
| | **Control** | | | | |
| | A7 | A9 | A11 | A13 | A15 |
| Sequence pairs analysed: | 2,86E+07 | 2,79E+07 | 2,63E+07 | 2,48E+07 | 3,33E+07 |
| N of paired-end alignments unique | 1,42E+07 | 1,63E+07 | 1,39E+07 | 1,41E+07 | 1,97E+07 |
| Mapping efficiency: | 49.7% | 58.4% | 53.1% | 56.7% | 59.0% |
| Number of C's analysed: | 6,25E+08 | 6,85E+08 | 5,90E+08 | 5,85E+08 | 8,02E+08 |
| Methylated C's in CpG context: | 5,92E+07 | 6,09E+07 | 5,46E+07 | 4,84E+07 | 6,69E+07 |
| Unmethylated C's in CpG context: | 1,05E+07 | 9,39E+06 | 1,00E+07 | 8,80E+06 | 1,24E+07 |
| C methylated in CpG context: | 84.9% | 86.6% | 84.5% | 84.6% | 84.4% |

**Table 12 –** Final alignment report of DNA methylation efficiency with bisulphite treatment in 3 groups (Bismark output)

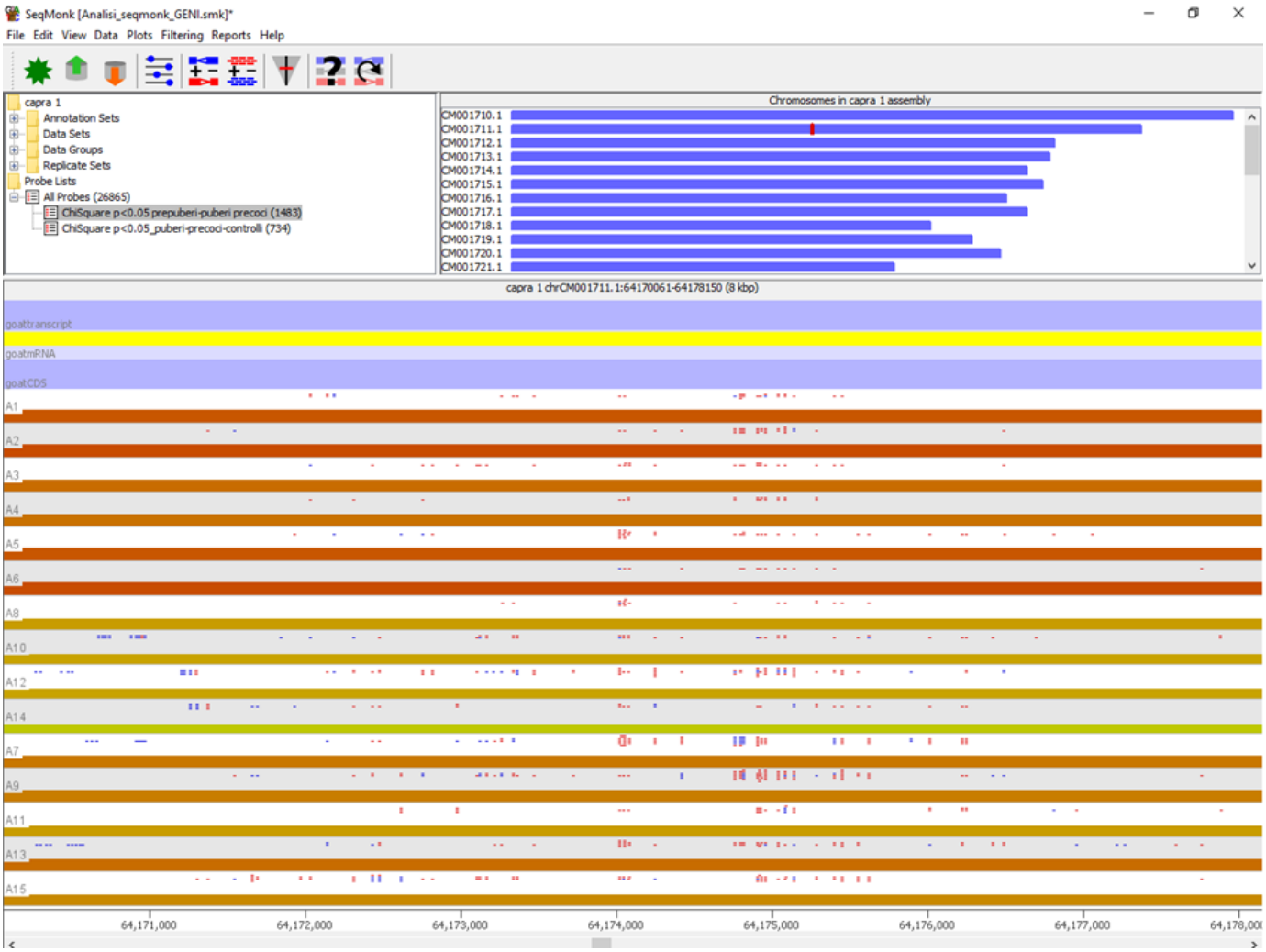


**Figure 23 -** Screenshot from SeqMonk of the DNA methylation profile (8kb window) of pre-pubertal goat (A1…A5), "early puberty" goats (A6, A8…A14) and their control (A7, A9…A15). The color of each line is proportional to the degree of methylation: red=intense methylation_green=weak methylation.

The DNA methylation pattern that characterizes a peculiar physiological status (in our case "puberty") is more thin and sensitive. Considering DMGs (Differentially Methylated Genes), we found 1,232 differentially methylated genes (Q-value<0.05) between pre-pubertal and pubertal goats. Comparing the group of pubescent goats and their sexually immature control, 664 DMGs were detected. Matching DMGs found between the two comparisons (1,232 ∩ 664 DMGs) revealed 232 genes in common. Pathway analysis of these genes highlighted, as significant biological processes (FDR<0.01), 65 pathways. As reported in table 13, embryonic skeletal system development and morphogenesis and regulation of nitrogen compound took on great importance. In addition, we decided to investigate the methylation pattern of the three candidate genes object of the SNP discovery (see paragraph 5.6), despite they did not result significant in the overall analysis of DMGs. According to what was observed in Table 12, pre-pubertal goats revealed the highest methylation level for all the three genes involved

in reproductive seasonality (*MTNR1A, TSHB* and *DIO2*). On the other hand, in this case, all the controls showed lower methylation level than goats that have reached puberty (Table 14).

| | **Biological Process (GO)** | | |
|---|---|---|---|
| **pathway ID** | **pathway description** | **count in gene set** | **FDR** |
| GO:0048706 | embryonic skeletal system development | 12 | 1.04e-05 |
| GO:0010467 | gene expression | 61 | 7.38e-05 |
| GO:0007389 | pattern specification process | 17 | 0.000181 |
| GO:0051171 | regulation of nitrogen compound metabolic process | 59 | 0.00019 |
| GO:0048704 | embryonic skeletal system morphogenesis | 9 | 0.000204 |
| GO:0016070 | RNA metabolic process | 53 | 0.000493 |
| GO:0034654 | nucleobase-containing compound biosynthetic process | 49 | 0.000493 |
| GO:0034641 | cellular nitrogen compound metabolic process | 68 | 0.000509 |
| GO:0010468 | regulation of gene expression | 56 | 0.000628 |
| GO:0006807 | nitrogen compound metabolic process | 70 | 0.000685 |

**Table 13** - Top 10 most significant pathways of Biological Processes involving shared genes of the two puberty related comparisons (pre-pubertal vs pubertal and pubertal vs control)

EED and CBX7, two genes related to puberty, confirmed the highest level of methylation observed in the prepubertal goats. Considering the comparison among pubertal goats and their control, EED showed a lower methylation level in controls, while in the gene body of CBX7 early pubertal goats had a DNA profile almost 6% less methylated than controls. The DNA methylation pattern in the 5' flanking regions (2 kb) of these two genes revealed a lower % of CpG methylated in the group of early pubertal goats if compared with the other two groups. However, the difference in methylation among pubertal goats and their control in the EED 5'fr resulted more marked than the slight difference in CBX7 5' fr (-33.02% vs -1.18% respectively).

| | ***MTNR1A*** | ***TSHB*** | ***DIO2*** | ***EED*** | ***CBX7*** | ***EED 5' fr*** | ***CBX7 5' fr*** |
|---|---|---|---|---|---|---|---|
| **Pre-pubertal goats** | 88.30 | 99.00 | 91.71 | 91.66 | 92.59 | 60.00 | 86.72 |
| **Early pubertal goats** | 83.18 | 88.00 | 88.05 | 89.71 | 77.87 | 57.33 | 78.64 |
| **Control** | 81.75 | 85.73 | 81.77 | 87.29 | 84.07 | 90.35 | 79.82 |

**Table 14** - Percentage of CpGs methylated within gene body of five candidate genes in reproductive seasonality and puberty. In particular, *MTNR1A*, *TSHB*, and *DIO2* are linked to reproductive seasonality, *EED* and *CBX7 (and their 5' flanking regions - fr*) to puberty.

### 5.3.1 DMGs and DEGs across puberty

For the same three groups of goat, we also investigated the expression profile. Checking the different expression profile of pre-pubertal goats and those who reached puberty, 855 Differentially Expressed Genes (DEGs) (FDR<0.05) were identified. Surprisingly, the group of pubescent goats versus their

control did not show any DEGs considering FDR<0.05 as threshold. Only considering a P-value<0.05, 161 genes had a different expression between the two groups of goats.
With the similar approach used for differentiate DNA methylation profile in ovary and hypothalamus (paragraph 5.1.1), we crossed DMGs and DEGs to extract a list of genes of particular interest from the epigenetic point of view. Intersections resulted in 197 common genes between DMGs and DEGs of pre-pubertal and pubertal goats. Pathway analysis of biological processes revealed a specific focus on brain functions, as in the case of synaptic plasticity, nervous system development and transmission of nerve impulse (Table 15).

| Biological Process (GO) | | | |
|---|---|---|---|
| **pathway ID** | **pathway description** | **count in gene set** | **FER** |
| GO:0031344 | regulation of cell projection organization | 16 | 0.0131 |
| GO:0048172 | regulation of short-term neuronal synaptic plasticity | 4 | 0.0131 |
| GO:0007399 | nervous system development | 36 | 0.0132 |
| GO:0016043 | cellular component organization | 64 | 0.0132 |
| GO:0007610 | behavior | 16 | 0.0133 |
| GO:0010975 | regulation of neuron projection development | 13 | 0.0133 |
| GO:0019226 | transmission of nerve impulse | 6 | 0.0133 |
| GO:0050767 | regulation of neurogenesis | 17 | 0.0133 |
| GO:0048699 | generation of neurons | 27 | 0.0146 |

**Table 15** - Top 10 most significant pathways of Biological Processes involving shared genes of DMGs and DEGs among pre-pubertal vs pubertal goats.

The same intersection among pubescent goats and their control, but considering a P-value<0.05 for DEGs, highlighted only 8 genes shared. Genes, sorted by statistical significance were (according to NCBI classification):

- *GTSE1* - G2 and S-phase expressed 1. The protein encoded by this gene, only expressed in the S and G2 phases of the cell cycle, colocalizes with cytoplasmic tubulin and microtubules. In response to DNA damage, the encoded protein accumulates in the nucleus and binds the tumor suppressor protein p53, shuttling it out of the nucleus and repressing its ability to induce apoptosis.
- *ATG16L1* - autophagy related 16 like 1. The protein encoded by this gene is part of a large protein complex that is necessary for autophagy, the major process by which intracellular components are targeted to lysosomes for degradation. Defects in this gene are a cause of susceptibility to inflammatory bowel disease type 10.
- *LRP5* - LDL receptor related protein 5. This gene encodes a transmembrane low-density lipoprotein receptor that binds and internalizes ligands in the process of receptor-mediated endocytosis. This protein plays a key role in skeletal homeostasis and many bone density related diseases are caused by mutations in this gene. Mutations in this gene also cause familial exudative vitreoretinopathy

- *MERTK* - MER proto-oncogene, tyrosine kinase. This gene is a member of a receptor kinase family and encodes a transmembrane protein with two fibronectin type-III domains, two immunoglobulin-like C2-type domains, and one tyrosine kinase domain. Mutations in this gene have been associated with disruption of the retinal pigment epithelium (RPE) phagocytosis pathway and onset of autosomal recessive retinitis pigmentosa (RP).
- *PCID2* - PCI domain containing 2. This gene encodes a component of the TREX-2 complex (transcription and export complex 2), which regulates mRNA export from the nucleus. This protein regulates expression of Mad2 mitotic arrest deficient-like 1, a cell division checkpoint protein. This protein also interacts with and stabilizes Brca2 (breast cancer 2) protein.
- *FOXN3* - forkhead box N3. This gene is a member of the forkhead/winged helix transcription factor family. Checkpoints are eukaryotic DNA damage-inducible cell cycle arrests at G1 and G2. Checkpoint suppressor 1 suppresses multiple yeast checkpoint mutations including mec1, rad9, rad53 and dun1 by activating a MEC1-independent checkpoint pathway.
- *PISD* - phosphatidylserine decarboxylase. The protein encoded by this gene catalyzes the conversion of phosphatidylserine to phosphatidylethanolamine in the inner mitochondrial membrane. The encoded protein is active in phospholipid metabolism and interorganelle trafficking of phosphatidylserine.
- *ITGB5* - integrin subunit beta 5. Integrins, heterodimeric trans-membrane matrix receptors, are major mediators of cell adhesion to extracellular matrix and extracellular matrix -induced intracellular signaling.

Pathway analyses for the 8 genes, did not reveal any significant interaction.

## 5.4 MicroRNA population in the hypothalamus across puberty

After proving that miRNA profile change significantly among organs, we investigated if, within the same organ, there were clusters of miRNAs characterizing puberty. So, we also produced a list of miRNAs found in the hypothalamus of 5 goats before puberty and 5 at puberty reached. The comparison revealed 39 miRNAs (P<0.001) differentially expressed in the hypothalamus. if the threshold became more stringent (FDR) there were no significant miRNA among the groups. In table 16 we reported a subset including only known *Capra hircus* miRNAs with the highest statistical significance.

| MiRNA | logFC | logCPM | LR | PValue | FDR |
|---|---|---|---|---|---|
| chi-let-7g-3p | 0.90021 | 5.33821 | 15.92711 | 6.58E-05 | 0.03248 |
| chi-let-7i-3p | 0.75807 | 6.11416 | 10.71539 | 0.00106 | 0.17707 |
| chi-miR-708-3p | 0.72180 | 9.09172 | 9.84220 | 0.00171 | 0.18133 |
| chi-miR-101-3p | 0.60408 | 12.91260 | 9.30510 | 0.00229 | 0.18789 |
| chi-miR-628-5p | -0.74692 | 4.90338 | 8.86990 | 0.00290 | 0.21453 |
| chi-miR-146a | -0.93512 | 9.21638 | 8.36652 | 0.00382 | 0.24594 |
| chi-miR-411b-3p | -0.67603 | 2.98083 | 7.24038 | 0.00713 | 0.37022 |
| chi-miR-21-3p | 0.70949 | 5.87927 | 7.16625 | 0.00743 | 0.37022 |
| chi-miR-1307-5p | 0.55538 | 9.73817 | 7.04855 | 0.00793 | 0.37022 |

| | | | | | |
|---|---|---|---|---|---|
| chi-miR-151-5p | -0.69330 | 12.80227 | 6.97828 | 0.00825 | 0.37022 |
| chi-miR-362-5p | 0.60859 | 4.33631 | 6.84966 | 0.00887 | 0.37022 |
| chi-miR-200c | -2.60433 | 9.47824 | 6.74980 | 0.00938 | 0.37022 |
| chi-miR-320-3p | 0.53102 | 9.63669 | 6.68368 | 0.00973 | 0.37022 |
| chi-miR-141 | -2.26510 | 12.72100 | 6.67897 | 0.00976 | 0.37022 |

**Table 16** – Known chi-miRNAs differentially expressed in the hypothalamus of goats belonging to "early puberty" and "pre-pubertal" group. If logFC value is positive the miRNA is highly expressed in "pre-pubertal" group, if negative it is highly expressed in the "early puberty" group. logCPM = average log counts per million over all samples; LR = likelihood ratio; FDR = False Discovery Rate.

As we did for methylation, we decided to analyze miRNAs population of "early puberty goats" with goats at the same age and weight but still prepubescent to exclude that the differences that we were observing were mainly due to different sizes of animals. We performed the sequencing of miRNAs population of the 5 "early puberty goats" and their 5 control. Considering a P-value <0.001, 12 miRNAs are differentially expressed in the two groups. Also in this case, if the threshold became more stringent (FDR) there were no significant miRNA among the groups (table 17).

| miRNA | logFC | logCPM | LR | PValue | FDR |
|---|---|---|---|---|---|
| Novel:CM001739,1_32789 | -2.71922 | 3.67034 | 14.97773 | 0.00011 | 0.1610 |
| Novel:CM001714,1_6322 | 4.11676 | -0.91695 | 13.20168 | 0.00028 | 0.2069 |
| Novel:CM001727,1_22373 | 1.79748 | -0.48047 | 11.97393 | 0.00055 | 0.2662 |
| Novel:CM001729,1_24811 | 0.83927 | 3.75177 | 7.46741 | 0.00628 | 1 |
| Novel:CM001716,1_10591 | -1.29498 | -0.34086 | 7.32597 | 0.00680 | 1 |
| chi-miR-183 | -2.89946 | 10.18273 | 7.19269 | 0.00732 | 1 |
| Novel:bta-miR-1584-3p | -1.39509 | -0.53199 | 7.17080 | 0.00741 | 1 |
| Novel:bta-miR-200 | -2.39583 | 12.59747 | 6.98830 | 0.00820 | 1 |
| Novel:CM001711,1_1861 | 1.44918 | 1.58265 | 6.96340 | 0.00832 | 1 |
| chi-miR-141 | -2.37398 | 12.72100 | 6.90578 | 0.00859 | 1 |
| Novel:CM001714,1_7660 | 0.91911 | 1.58688 | 6.83309 | 0.00895 | 1 |
| chi-miR-96 | -2.95954 | 6.25875 | 6.82723 | 0.00898 | 1 |

**Table 17** – miRNAs differentially expressed in the hypothalamus of goats belonging to "early puberty" and "control" group. If logFC value is positive the miRNA is highly expressed in "early puberty" group, if negative it is highly expressed in the "control". logCPM = average log counts per million over all samples; LR = likelihood ratio; FDR = False Discovery Rate.

## 5.5 Hormonal analysis for reproduction and photoperiod response

For a better phenotypic characterization of the animals object of the analysis, for the same 20 goats used for the SNP discovery, ELISA analysis were conducted on three key hormones in reproduction and development: melatonin, T3 (triiodothyronine) and T4 (thyroxine). Analysis was performed in four different topic moments of the seasonal photoperiod for the 5 goats at checkpoint 3. The four chronological moments corresponded to:

1. one week before the switching on of artificial lights

2. one week after the switching on of artificial lights
3. at the arrest of the cycle
4. the week before the switching off of the artificial lights

Further details are shown in the material and methods section, on paragraph 4.1.

For the 15 remaining animals ("early puberty", their control and "late puberty") the analysis was performed only once, as a function of the display of puberty. We found statistically significant differences over time between individuals for melatonin and T4 concentration. According to literature (Ikegami & Yoshimura 2012) T3 remained constant over time. Analyzing results among groups, we observed a statistical significant difference for melatonin concentration (P=0.037) between the group of precocious goats and their control. Furthermore, there is a significant difference between the groups "early-puberty" and "late-puberty" for T3 concentration (P=0.03).

We observed the higher variability of melatonin and T4 in all the 3 groups ("early-puberty", "control" and "late-puberty"), despite the fact that we did not observe any genetic variability among individuals (see paragraph 5.6). This phenotypic variability could be due to the epigenetic differences observed at the DNA methylation level.

Melatonin concentration was more variable compared to the other hormones investigated, as showed by the CV% value (Table 18). The Shapiro-Wilk test revealed that data obtained can be represented by a normal distribution except for the T3, due to the presence of an outlier, a pre-pubertal goat that was removed from the analysis of descriptive statistics. We found statistically significant differences over time between individuals for melatonin and T4 concentration (P<0.05). Given the small size of the sample analyzed and, in the case of the T3 hormone, a violation of the homoscedasticity, we performed the Wilcoxon test on the medians. We observed a positive correlation between T3 and T4 levels (r=0.52, P=0.007) as expected, and also between melatonin and T4 (r=0.44, P=0.027). Moreover, we observed the higher variability of melatonin and T4 in all the 3 groups ("early-puberty", "control" and "late-puberty"), despite the fact that we did not detect any genetic variability among individuals.

| Hormone | Mean | SD | Minimum | Maximum | N. animals | CV% |
|---|---|---|---|---|---|---|
| **Melatonin (pg/ml)** | 261.66 | 61.72 | 166.48 | 380.86 | 25 | 23.59 |
| **T3 (ng/ml)** | 1.12 | 0.08 | 0.97 | 1.27 | 25 | 7.30 |
| **T4 (ng/ml)** | 28.22 | 4.16 | 22.36 | 38.33 | 25 | 14.75 |

**Table 18 -** Descriptive statistics of three hormones investigated through ELISA test on five female goats. Animals were monitored along the artificial photoperiodic changes of a deseasonalisation procedure.

The level of the T3 was significantly lower at the arrest of the cycle, corresponding to the artificially increasing of day-length (checkpoint 3 described in materials and methods), with a mean value $\pm$ SD of 1.05 ± 0.05 ng/ml.

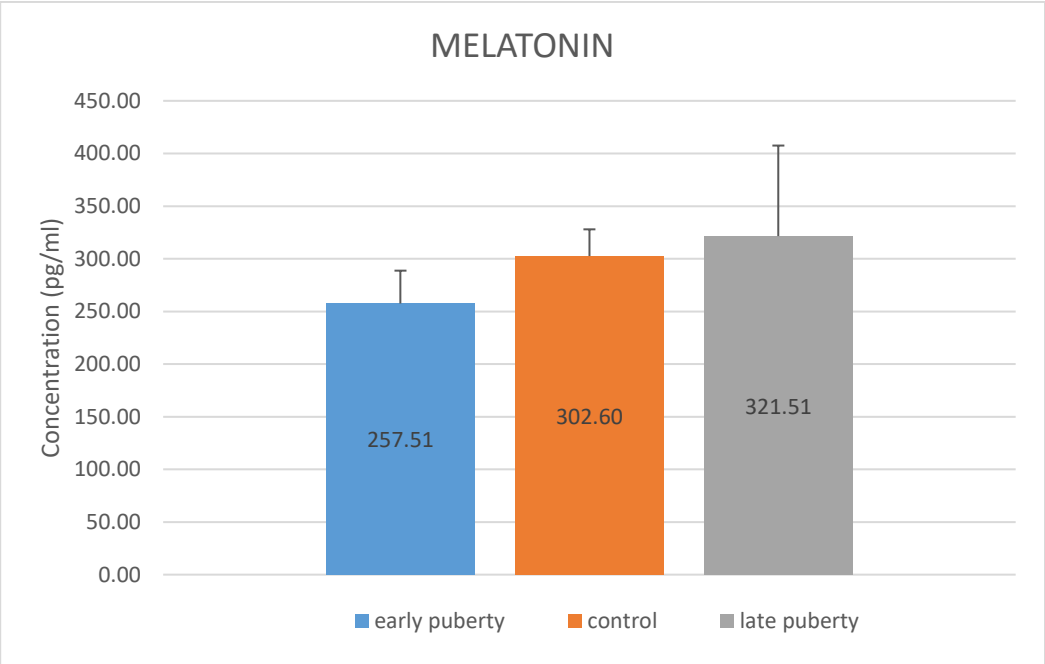


**Figure 24 -** Mean values (±SD) of melatonin concentration among three groups of goat selected according to their timing of puberty. Hormonal values were measured through ELISA tests.

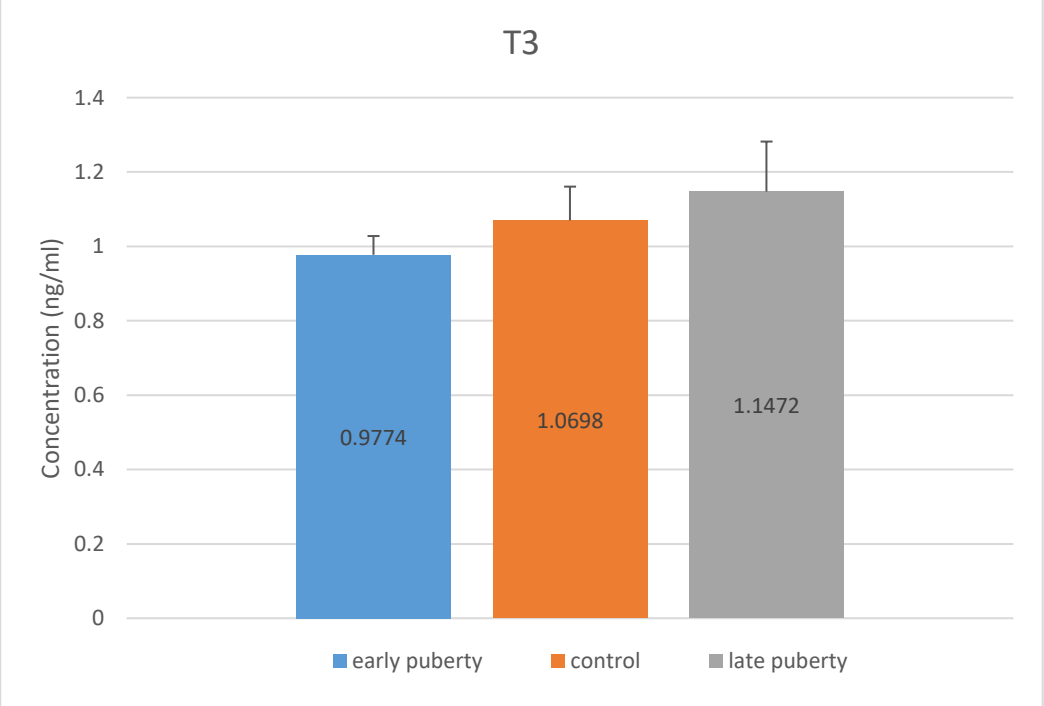


**Figure 25 -** Mean values (±SD) of T3 concentration among three groups of goat selected according to their timing of puberty. Hormonal values were measured through ELISA tests.

## 5.6 SNP discovery on candidate genes for reproduction and photoperiod response

In parallel with epigenetic analysis we conducted a SNP discovery to investigate if a peculiar genetic pattern is linked to specific epigenetic marker. The analysis was performed on three candidate genes for the regulation of the reproductive activity linked to photoperiod for the 20 animals in 4 checkpoints. The selected genes were:

- *MTNR1A* (*Melatonin Receptor 1a)*
- *TSHB* (*Thyroid Stimulating Hormone subunit β)*
- *DIO2* (*Deiodinase 2*)

Sequence alignments and chromatograms analysis (examples are provided on Figure 26 and 27) allow us to detect the presence of 1 SNP in the *DIO2* gene and 9 in the *MTNR1A* gene. No polymorphisms were detected in the *TSHB* gene. We confirmed the high variability of the *MTNR1A* gene, discovering a new SNP (G543T) bringing to a silent mutation. It would be of particular interest to verify if the "late-puberty" goats confirm a lower genetic variability for the loci studied. The *DIO2* SNP (C26T) identified was the first known SNP identified in the gene up to now. The non-synonymous mutation brings to a consequent change of the codon and the corresponding aminoacid (Ala<Val). Details on the SNPs identified are shown in Table 19.

Comparing the genotypes of the animals, we observed a high genetic variability of the *MTNR1A*. This confirms what previously observed in literature by several authors. In particular, we observed a higher variability in the genotype of the animals precocious to puberty, compared to their control and to those who showed a late puberty.

In particular, the group of "late-puberty" did not present variation in the genotype for the genetic locus. A peculiar goat in the group of "early-puberty" was characterized by the highest level of polymorphisms. These difference between genotypes deserves to be analyzed on a higher number of animals to evaluate if exist a positive relation between "early-puberty" and genetic variability with melatonin receptor. We then compared the methylation pattern of these genes among animals.

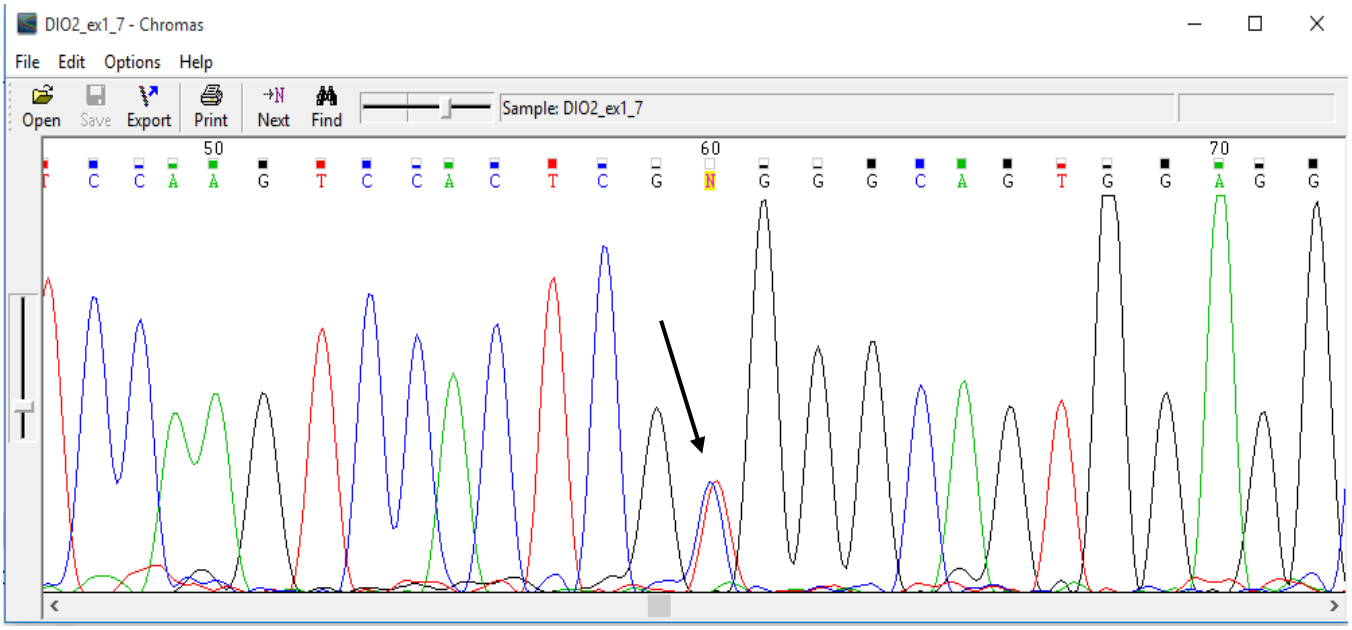


**Figure 26** - Example for chromatogram of the exon1 of the gene DIO2. Arrow indicates the heterozygous SNP

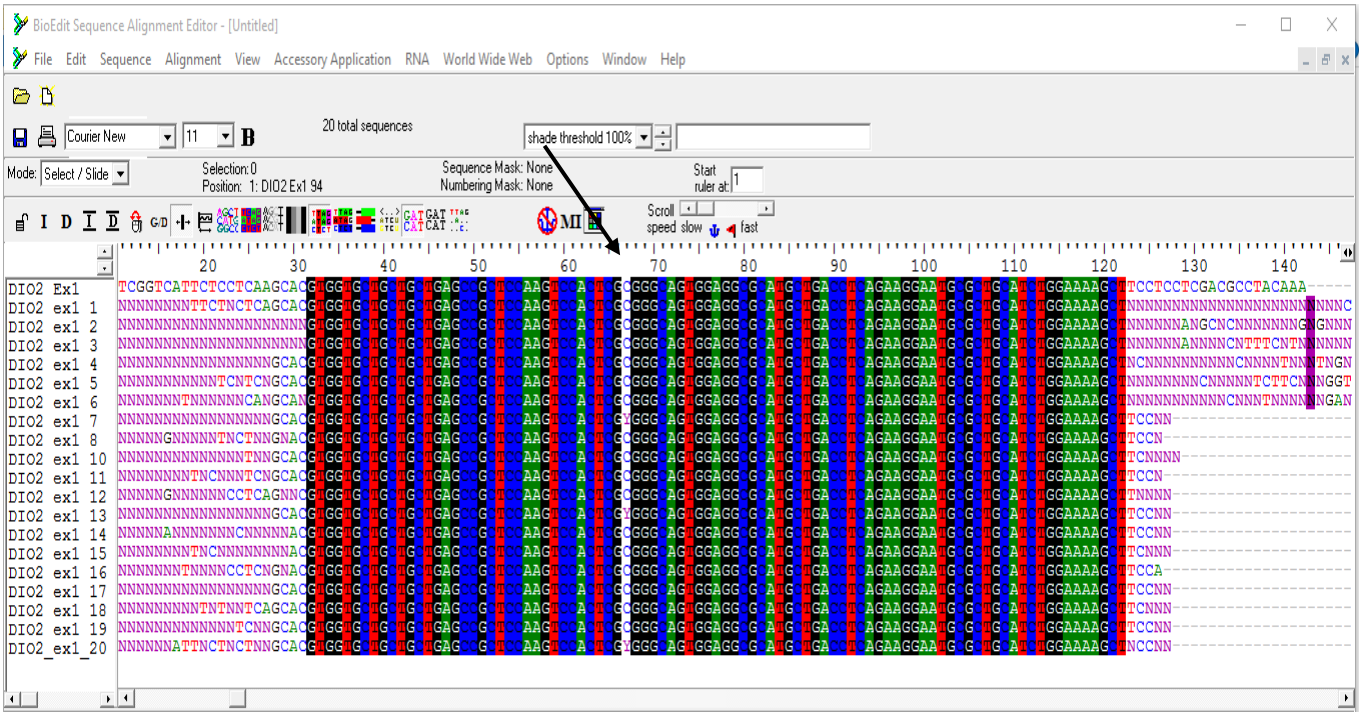


**Figure 27** - Example of alignment of the exon1 of the gene DIO2. Arrow show the column with the detected SNP

| Gene | SNP name | Location | SNP type | aa. change | References | Allele | MAF |
|---|---|---|---|---|---|---|---|
| DIO2 | C62T | exon 1 | C/T | Ala>Val | Present work | T | 0.08 |
| MTNR1A | C426T | exon 2 | C/T | no | Migaud et al. (2002) | T | 0.28 |
| MTNR1A | G469C | exon 2 | G/C | Gly>Arg | Migaud et al. (2002) | C | 0.11 |
| MTNR1A | G474A | exon 2 | G/A | no | Sagawa et al. (2012) | A | 0.06 |
| MTNR1A | G513A | exon 2 | G/A | no | Sagawa et al. (2012) | A | 0.06 |
| MTNR1A | G543T | exon 2 | G/T | no | Present work | T | 0.11 |
| MTNR1A | T657C | exon 2 | T/C | no | Migaud et al. (2002) | C | 0.22 |
| MTNR1A | C660T | exon 2 | C/T | no | Migaud et al. (2002) | T | 0.17 |
| MTNR1A | C813T | exon 2 | C/T | no | Migaud et al. (2002) | T | 0.03 |
| MTNR1A | C825A | exon 2 | C/A | no | Migaud et al. (2002) | A | 0.14 |

**Table 19 -** Characteristics of 10 SNPs found in *Capra hircus DIO2* and *MTNR1A* genes, in a sample of 20 Alpine goats.

# 6 - Discussion

A methyl binding domain sequencing approach was taken to study genome-wide DNA methylation patterns in two organs from the HPG axis in Saanen goats. A large number of differentially methylated regions were identified: 382,850 and 413,010 methylation peaks were found in hypothalamus and ovary, respectively.

Considering the methylation density across gene regions, hypothalamus and ovary showed a similar trend: exons and promoters had the highest methylation frequency. This high methylation density in exons has already been observed in horse, rat and human (Laurent et al. 2010; Sati et al. 2012; Lee et al. 2014). However, the high density of methylation in promoter regions seen in the present study was unexpected. It is well documented that most promoters have a low level of methylation and subsequently facilitate gene expression (Klose & Bird 2006). The high level of methylation could reflect the type of tissues studied.

Compared with other livestock, goats show similar distribution of DNA methylation across different gene regions (Hu et al. 2013; Lee et al. 2014; Su et al. 2014), e.g. comparing the DNA methylation distribution pattern in goat to that found in bovine placenta and horse (Su et al. 2014; Lee et al. 2014). In bovine placenta, the DNA methylation level decreased before the TSS, noticeably increases in intragenic regions, and is constantly low downstream of genes. Methylation density in several tissues in horse, also showed a decrease before the TSS, followed by a gradual increase of methylation in the gene body with an acute decline at the TTS and subsequent plateau. Measures of methylation density in hypothalamus and ovary in goat indicates a slight increase before the CDS start, followed by a gradual increase in the intragenic region which contains scattered areas of increased methylation. After the CDS end the DNA methylation level falls sharply and remains constantly low for all the downstream region in both organs.

Analysis of peaks of methylation per chromosome showed that CHI X has the lowest number of methylation peaks. A possible explanation is that the X chromosome has one of the lowest CpG content within the genome. On the other hand, the X chromosome has one of the highest percentages of methylated CpG islands in both hypothalamus and ovary (59.36% and 66.52% respectively). Despite the lower degree of methylation, strong methylation in CpG islands could suggest a gene regulation of the CHI X. This high level of methylation observed in CGIs may be associated with X-chromosome inactivation (Sharp et al. 2011; Barakat & Gribnau 2012).

The associations between CGIs DNA methylation and gene expression found here were only partially in accordance with previous findings in mammals (Hu et al. 2013; Couldrey et al. 2014; Su et al. 2014; Lee et al. 2014). CpG islands have usually been reported to be regions of gene regulation via DNA methylation, most likely through the mechanism of transcriptional repression. These regions in vertebrate genomes are known to be generally unmethylated, in spite of having a high GC content (Deaton & Bird 2011). However, recent findings suggest that a relatively high level of DNA methylation can occur in CGIs (Straussman et al. 2009; Panchin et al. 2016). In the present study, CpG islands showed

a higher methylation density compared with previous studies. Recent works on horse, cattle and chicken epigenomes found 10 to 20% of methylated CGIs across the genomes (Su et al. 2014; Lee et al. 2014; Hu et al. 2013). The results in the present work indicate that almost 50% of the CGIs in goat are methylated in hypothalamus and ovary. This may be a result of the enrichment method used (Nair et al. 2011). Recent evidence suggests a more complicated role of DNA methylation than simply inhibitory expression (Jones 2012). In the present paper, the correlation between RNA-seq and methyl-seq profiles in the two tissues was more consistent for genes with greater differences in expression.

The comparison of hypo/hyper-methylated with the under/over-expressed genes showed that the methylation of a significant fraction of DMGs was negatively correlated with the expression level, in agreement with the idea that DNA methylation represses gene expression (Moore et al. 2013). However, the two organs considered here may exhibit different effects of methylation. A strong negative correlation between hypo-methylation and over-expression was seen in the hypothalamus (P=1.26E-46) while only a slight increase in expression was seen in the ovary (P=0.0167). In contrast with a previous reported study (Wan et al. 2015), there was no positive correlation between gene methylation and expression.

Pathway analysis revealed that the highly expressed/low methylated genes in the hypothalamus are involved in brain specific signaling. No correlation with specific pathways and tissue function was observed in the ovary. Thus, an interesting observation that emerge from this study is that a low level of gene methylation in CpGs is linked to a high level of expression in the hypothalamus. This is in contrast to what has been observed elsewhere, with a high level of methylation in gene body being associated with a high expression level (Jones 2012; Aran et al. 2011). Epigenetic regulation in the hypothalamus may be in line with the observation that for slowly dividing and non-dividing cells, such as those in the brain, gene body methylation is not associated with increased gene expression (Moore et al. 2013). The comparison of the transcriptomes between hypothalamus and ovary showed that a higher level of methylation is not necessarily accompanied by the suppression of expression. This, confirms that methylation is not constantly associated with gene silencing, and its effect on gene expression regulation could be both positive and negative (Lou et al. 2014; Wan et al. 2015). There are also methods for single-cell genome and transcriptome sequencing that have contributed to our understanding of cellular heterogeneity, whereas methods for single-cell epigenomics are much less established. Farlik and colleagues recently validated a protocol in both mouse and human cells for single-cell methylome sequencing. This will be probably the future of the new era of epigenetic analysis (Farlik et al. 2015).

One of the main goals of our project was to compare the epigenetic differences on the genome of the samples collected at each checkpoint. Unfortunately, our choice to not influence goats destined to the “out of season breeding” with external factor, as the “buck effect” and melatonin injections, did not pay off and goats did not show an out of season heat. The risk was taken because we wanted to observe a clear epigenetic response, without any external perturbation that could hide the epi-effect we were looking for. Said that, we decided to focus our efforts on the characterization of the factors that drives to puberty in the species.

In a first comparison we detected 1,232 DMGs between pre-pubertal and pubertal goats. Chromosome X showed one of the highest number of DMGs. This chromosome was mentioned as the one with the highest frequency of methylation, considering CpG islands, in the comparison of hypothalamus and ovary in adult goats. In the aforementioned comparison among young goats, differences observed could be due to the different age and size of the animals analyzed. To avoid this possible bias, we also chose to sequence and analyze 5 goats as controls of the 5 at their pubertal stage (same age and weight). DMGs analysis revealed 664 genes that differentiated the methylome of pubertal goats and their control. The percentage of C methylated for goats with similar age and weight to the pubertal ones, but still prepubescent, were slightly lower than their counterpart (85.00% for control vs 85.66% for Early puberty goats) (Table 12). This suggests that the almost 5% of difference in methylation observed between pre-pubertal and early puberty goats could be due to the variable age of the animals. In fact, even though the methylation profile of the mammalian genome is largely stable, the methylation of specific genomic regions changes dynamically across development, and different cellular lineages have unique methylation profiles (Ziller et al. 2013). The dynamic change in DNA methylation profile was confirmed by the analysis of the 3 candidate genes (*MTNR1A*, *DIO2* and *TSHB*), in which the methylation level was the highest in the group of pre-pubertal goats, it decreased in the control, and increased again in pubertal goats.

Matching DMGs found between the comparisons of pre-pubertal and pubertal goats, and the group of pubescent goats and their sexually immature control revealed 232 genes in common. Pathway analysis highlighted, as most significant biological processes (FDR<0.01) embryonic skeletal system development and morphogenesis and regulation of nitrogen compound. If pathways involving development and morphogenesis were expected players in the puberty onset, the involvement of nitrogen compound was not that obvious. Several publications showed how a correct balance of nitrogen compounds is fundamental in reaching and maintaining puberty (Soliman et al. 2014; Gattas et al. 1990). The crossing of DMGs and DEGs among pubescent goats and their control, revealed 8 genes in common. Pathway analysis did not show any significant interaction, probably because of the reduced number of interactions. Despite this fact, it seemed that the list of genes could fit as puberty related. As an example, *ATG16L1* (autophagy related 16 like 1) mainly involved in Crohn's disease, is indirectly implicated in the delay of puberty (Naser et al. 2012), *LRP5* (low-density lipoprotein receptor-related protein 5) is associated with calcium homeostasis and bone mineral density, fundamental factors in reaching puberty (Ashouri et al. 2015). *MERTK* (proto-oncogene tyrosine-protein kinase MER) whose expression is low during puberty, maturity, and pregnancy, but is rapidly induced during postpartum involution (Stanford et al. 2014), is for sure another suitable example of puberty-related genes.

In a recent study, published at the end of October 2016, a group of researchers compared the DNA methylation pattern in hypothalamus of three Anhuai goat, a rural Chinese breed, across puberty (Yang et al. 2016). Unfortunately, the goats analyzed in the paper were not well characterized from a phenotypic point of view. It seems that the differences observed among pubertal a pre-pubertal goats

were only partially due to the onset of puberty, and they could be related to aging (Jones et al. 2015), as in our comparison among the early puberty group and the pre-pubertal one.
Another important epigenetic feature analyzed in the present study were microRNAs. Extensive high-throughput sequencing studies of miRNAs have been performed in several animal models (see paragraph 2.1.2). However, little is known about the diversities of these regulatory RNAs in goat species. So, we decided to investigate the microRNAs expression of reproductive-related organs in dairy goat. The expression of sequencing miRNAs was confirmed by qRT-PCR analysis on 5 conserved miRNAs (miR-141, miR-7-5p, miR- 9-5p, miR 124a and miR-10a-5p). The results of qRT-PCR analysis indicated that the high-throughput sequencing data were available and reliable. It is well known that conserved miRNAs were found in many animal species and had important functions in mammalian development and physiological processes (Ji et al. 2012). In this study, we found that of the 760 sequencing miRNAs identified, 402 miRNAs were already known in the species, 210 miRNAs were conserved to sheep and cows, and 148 miRNAs were predicted on the basis of their peculiar sequences. These analyses on miRNAs reveals a comprehensive survey of large and small RNA repertoire in the HPG axis in goat. Differentially expressed miRNA will be further investigated in order to evaluate their role in organ specific mRNA expression and their possible function in the regulation of reproductive physiology.

Recent research has indicated that specific members of the let-7 family can affect mammalian reproduction, development, cell proliferation and apoptosis (Roush & Slack 2008; Miles et al. 2012; Sangiao-Alvarellos et al. 2013). In the present study, three members of the let-7 family (let-7b, let-7c, and let-7i) were expressed at high levels in ovary and pituitary, and chi-let-7g and chi-let-71 were the two most highly expressed known miRNAs in the hypothalamus of pubertal goats (if compared to pre-pubertal goats). In 2003, Zhang and colleagues identified chi-let-7 family in goat ovary up-regulated in non-pregnant goats, if compared to pregnant goats. Our data added an interesting role for miRNA let-7 family (in our case i and g), not only as markers of non-pregnancy, but they also seemed to distinguish pubertal goats from those younger which did not get the first heat yet.

Increased GnRH expression during the infantile-to-juvenile transition and that impairing microRNA synthesis in GnRH neurons leads to hypogonadotropic hypogonadism and infertility in mice. Two essential components of this switch, miR-200 family and miR-155, respectively regulate Zeb1, a repressor of GnRH transcriptional activators and GnRH itself, and Cebpb, a nitric oxide–mediated repressor of GnRH that acts both directly and through Zeb1, in GnRH neurons (Messina et al. 2016). In our comparison of miRNA populations among organs, miR-200a and c resulted upregulated in the pituitary. Furthermore, considering the significance of p-values ($P<0.001$) in the comparison of early puberty goats and their controls, one of the twelve miRNAs that overcame the threshold was miR-200 itself. In particular, mir-200 was upregulated in the controls, so in goat that did not reached puberty yet. A member of the mir-200 family was also differentially expressed in the previous comparison, among pre-pubertal goat and pubertal. Our data indicated not only that miRNAs are active in the hypothalamic GnRH network, but confirmed that miRNA-200 is key components of a complex developmental switch that controls GnRH promoter activity, whose correct functioning is essential for

the normal initiation of puberty and adult fertility. It is worth mentioning that, considering a more robust statistical correction (in our case FDR), according to our data, miRNA population in hypothalamus could not play a crucial role in the reaching of puberty in goat species in the organ analyzed.

The potential contribution of epigenetics to the regulation of puberty has been only recently addressed (Lomniczi et al. 2013). According with our findings, it seems that the onset of puberty is something very finely regulated from and epigenetic point of view in the hypothalamus. Lomniczi and colleagues (2013) provided evidence that an epigenetic mechanism of transcriptional repression, operating in the neuroendocrine brain, influences the timing of female puberty. They identify the PcG system of transcriptional silencing as a central element of this repressive mechanism. Hypothalamic expression of *CBX7* and *EED*, two PcG genes required for PcG action decreased before the onset of puberty, and this change was associated with increased DNA methylation of their 5' flanking regions. They suggested that the initiation of puberty was accompanied by changing methylation of the promoters of several members of the PcG repressive complex and genes encoding proteins that interact with the PcG system. In our samples, gene body DNA methylation in *EED* gene showed a lower methylation level in controls, while in the gene body of *CBX7* early pubertal goats had a DNA profile almost 6% less methylated than controls. On the other hand, analyzing the methylation pattern in the 5' flanking regions of both genes we confirmed a lower percentage of CpG methylated in the group of early pubertal goats if compared with the other two groups. If this observation was only a slight increase in CBX7, the 5' flanking region of EED in goats belonging to control group was about 33% more methylated than their counterpart that reached puberty. Our comparison among the group of early puberty and their control, suggested 8 new candidate genes that could be directly or indirectly involved in the onset of puberty. The next effort must be focused on disentangling the eventual bond between the target of miRNAs and DMGs.

Concerning the SNP discovery on three candidate genes and ELISA tests of photoperiod-related hormones, interesting considerations emerged. In fact, the level of the T3 was significantly lower at the arrest of the cycle, corresponding to the artificially increasing of day-length (checkpoint 3 described in materials and methods paragraph 4.1). This seems to be in contrast with previous studies, showing that the plasma T3 concentration increases during long days and decreases during short days in goat (Todini et al. 2006). It is now well known that thyroid hormones do not influence transition into the breeding season but their presence is required for its termination. In ewes, thyroid hormones must be produced for the seasonal increase in hypothalamic responsiveness to estradiol negative feedback and the seasonal inhibition of pulsatile GnRH secretion that causes the transition from breeding season to anoestrus (Thrun et al. 1997).

The T3 concentration was also significantly higher (P=0.03) in the "late-puberty" group compared with the "early-puberty" (figure 24). Similarly, we observed a significantly lower melatonin concentration (P=0.04) for the "early-puberty" goats compared with their control (figure 23). The role of melatonin is well established for what concerns the perception of photoperiod, and it can be considered as 'pro-gonadotrophic' stimulus (Abecia et al. 2012). Melatonin stimulates GnRH and LH secretion during anoestrus, by reducing tyrosine hydroxylase activity and, therefore, the secretion of dopamine in the

median eminence (Viguié et al. 1997). A low dopamine secretion during anoestrus is associated with an improvement of the reproductive activity. Melatonin implants cause a Short Day-like response by lengthening the duration of the melatonin signal. However, a specific role of melatonin and T3 in the onset of puberty without considering the photoperiodic changes is less clear. These mechanisms are strictly connected as in mammals, initiation of puberty requires an increase in pulsatile release of GnRH from the hypothalamus and this is positively influence by short photoperiod in goat. Further analysis on a larger sample could enhance the knowledge about these mechanisms.

We found a novel non-synonymous polymorphism in the caprine *DIO2* gene and we confirmed the high variability of the *MTNR1A* gene, discovering a new SNP bringing to a silent mutation. It would be of particular interest to verify if goats showing "late-puberty" confirm a lower genetic variability for the loci studied. We observed significant differences in respect to the timing of puberty for melatonin and T3 plasma levels. Further analyses on a large sample could enhance our understanding of the complex mechanisms behind reproductive seasonality and the onset of puberty in goat, also allowing the detection of more SNPs useful for marker assisted selection.

# 7 - Conclusion

This work presents a global methylation pattern for hypothalamus and ovary in dairy goat. The differences observed at the DNA methylation level in hypothalamus and ovary suggest the organ-specificity of methylation patterns. The DNA methylation landscape of the *Capra hircus* ovary and hypothalamus methylomes associated with the analysis of the transcriptome highlighted the complexity of epigenetic regulation. The correlation between RNA-seq and methyl-seq profiles in the two tissues was more consistent for genes with greater differences in expression. The comparison of hypo/hyper-methylated with the under/over-expressed genes showed that the methylation of a significant fraction of DMGs was negatively correlated with the expression level, in agreement with the idea that DNA methylation represses gene expression (Moore et al. 2013). However, the two organs considered here may exhibit different effects of methylation. A strong negative correlation between hypo-methylation and over-expression was seen in the hypothalamus (P=1.26E-46) while only a slight increase in expression was seen in the ovary (P=0.0167). In contrast with a previous reported study (Wan et al. 2015), there was no positive correlation between gene methylation and expression. Thus, an interesting observation that emerge from this study is that a low level of gene methylation in CpGs is linked to a high level of expression in the hypothalamus. Our results offer a basis for a better understanding of the influence of epigenetics on livestock development, response to complex diseases and production traits under different conditions. This will be likely an aid to increase animal productivity and sustainability. There is no doubt that a more extensive analysis is needed in the next years to confirm and generalize these observations. A promising topic for future research involves the examination of the influence of epigenetic marks relative to livestock development, complex diseases and production traits under different conditions and the inclusion of this information to increase animal productivity and sustainability (Ibeagha-Awemu & Zhao 2015).

Our study also revealed the first miRNA profile comparison related to the biology of ovary, hypothalamus and pituitary in goat. We identified subsets of miRNAs that might be involved in dairy goat organ-specific development. The characterization of these conserved miRNAs could contribute to a better understanding of the molecular mechanisms of reproductive physiology and development in the dairy goat. Our data indicated not only that miRNAs are active in the hypothalamic GnRH network, but confirmed that miRNA-200 is key components of a complex developmental switch that controls GnRH promoter activity, whose correct functioning is essential for the normal initiation of puberty and adult fertility. We also added an interesting role for miRNA let-7 family (in our case i and g), not only as markers of non-pregnancy, but they also seemed to distinguish pubertal goats from those younger which did not get the first heat yet. Future work to characterize the expression of ovarian miRNAs at different stages of reproduction and in different breeds of goat, or in specific cell lines derived from ovarian tissues, is necessary to fully elucidate the functions of miRNAs in the puberty onset of goat, which will help to understand the relationships between miRNAs and mammalian reproduction.

It would be of particular interest also associate the individual variability of SNPs found in goats in candidate genes, as *MTNR1A*, to peculiar epigenetic pattern (at microRNA and DNA methylation level) to enhance our understanding of the complexity behind reproductive seasonality and the onset of puberty in goat. With our genome-wide analysis of DNA methylation profiles of the goat hypothalamus during puberty we started to move the first steps towards the epigenetic machinery behind reproduction. DMGs analysis revealed 664 genes that differentiated the methylome of pubertal goats and their control. The percentage of C methylated for goats with similar age and weight to the pubertal ones, but still prepubescent, were slightly lower than their counterpart (85.00% for control vs 85.66% for Early puberty goats). This suggests that the almost 5% of difference in methylation observed between pre-pubertal and early puberty goats could be due to the variable age of the animals. In fact, even though the methylation profile of the mammalian genome is largely stable, the methylation of specific genomic regions changes dynamically across development, and different cellular lineages have unique methylation profiles (Ziller et al. 2013). By confirming that the neuroendocrine control of female puberty entail mechanisms subjected to epigenetic regulation, our results provide insight into genes and pathways involved in the integrative mechanisms used by the hypothalamus to control the initiation of mammalian puberty. As such, they are consistent with the concept that the pubertal process depends not only on genetic determinants, but also on developmentally regulated changes in epigenetic information.

In conclusion, it is clear that phenotypic complexity is not just a simple matter of Mendelian genetics: a fundamental role of epigenetics is assessed for adulthood reproductive seasonality. Understanding the (epi)genetic mechanisms behind out-of-season breeding in dairy goat will allow the pick of genetic (epi)mutations in animals more prone to deseasonalization. In this way breeders could easily spread milk production throughout the year, satisfying the increasing costumers' demand.

## Acknowledgments

The entire work was part of the Genhome project, funded by MIUR.

Many professors and doctors helped me with suggestions, tips, and, more practically, in sample collection, molecular analysis and data analysis. I would like to thank: my tutor Professor Giulio Pagnacco, Doctor Emanuele Capra, Professor Paola Crepaldi, Doctor Beatrice Coizet, Doctor Barbara Lazzari, Professor Stephanie McKay, Doctor Andrea Talenti, Doctor Alessandra Stella, Doctor Bianca Castiglioni, Doctor Marcello DelCorvo, Professor Silvana Arrighi, Professor Alessandro Pecile, Doctor Debora Groppetti, Professor Angelo Belloli, Professor Davide Pravettoni, Doctor Pietro Riccaboni, Doctor Stefania Chessa, Doctor Mario Villa, Professor John Williams.

# Bibliography


Abecia, J.A., Forcada, F. & González-Bulnes, A., 2012. Hormonal control of reproduction in small ruminants. *Animal Reproduction Science*, 130(3), pp.173–179.

Abecia, J.A., Forcada, F. & González-Bulnes, A., 2011. Pharmaceutical control of reproduction in sheep and goats. *The Veterinary clinics of North America. Food animal practice*, 27(1), pp.67–79.

Anders, S., Pyl, P.T. & Huber, W., 2015. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), pp.166–169.

Andrews, S., 2010. A quality control tool for high throughput sequence data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Anway, M.D. & Skinner, M.K., 2006. Epigenetic Transgenerational Actions of Endocrine Disruptors. *Endocrinology*, 147(6), pp.s43–s49.

Aran, D. et al., 2011. Replication timing-related and gene body-specific methylation of active human genes. *Human Molecular Genetics*, 20(4), pp.670–680.

Ashouri, E. et al., 2015. The impact of LRP5 polymorphism (rs556442) on calcium homeostasis, bone mineral density, and body composition in Iranian children. *Journal of Bone and Mineral Metabolism*, 33(6), pp.651–657.

Austin, C.R. & Short, R.V., 1982. *Reproduction in mammals* Second Edi. C. U. Press, ed., Cambridge.

Barakat, T.S. & Gribnau, J., 2012. X chromosome inactivation in the cycle of life. *Development*, 139, pp.2085–2089.

Baril, G., Leboeuf, B. & Saumande, J., 1993. Synchronization of estrus in goats: the relationship between time of occurrence of estrus and fertility following artificial insemination. *Theriogenology*, 40(3), pp.621–8.

Barrett, P. et al., 2007. Hypothalamic thyroid hormone catabolism acts as a gatekeeper for the seasonal control of body weight and reproduction. *Endocrinology*, 148(8), pp.3608–3617.

Bartolomei, M.S., 2009. Genomic imprinting: Employing and avoiding epigenetic processes. *Genes and Development*, 23(18), pp.2124–2133.

Beltramo, M. et al., 2014. Cellular mechanisms and integrative timing of neuroendocrine control of GnRH secretion by kisspeptin. *Molecular and Cellular Endocrinology*, 382(1), pp.387–399.

Bernal, J., 2002. Action of thyroid hormone in brain. *Journal of Endocrinological Investigation*, 25(3), pp.268–288.

Bird, A., 2002. DNA methylation patterns and epigenetic memory DNA methylation patterns and epigenetic memory. *Genes & Development*, pp.6–21.

Busk, P.K., 2014. A tool for design of primers for microRNA-specific quantitative RT-qPCR. *BMC Bioinformatics*, 15(1), p.29.

Carcangiu, V. et al., 2011. Characterization of the Mediterranean Italian buffaloes melatonin receptor 1A (MTNR1A) gene and its association with reproductive seasonality. *Theriogenology*, 76(3), pp.419–426.

Carcangiu, V., Mura, M.C., et al., 2009. Polymorphism of the melatonin receptor MT1 gene and its relationship with seasonal reproductive activity in the Sarda sheep breed. *Animal Reproduction Science*, 116(1–2), pp.65–72.

Carcangiu, V., Vacca, G.M., et al., 2009. Relationship between MTNR1A melatonin receptor gene polymorphism and seasonal reproduction in different goat breeds. *Animal Reproduction Science*, 110(1–2), pp.71–78.

Cech, T.R. & Steitz, J.A., 2014. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell*, 157(1), pp.77–94.

Chemineau, P. et al., 2006. Male-induced short oestrous and ovarian cycles in sheep and goats: a working hypothesis. *Reproduction, nutrition, development*, 46(4), pp.417–429.

Chemineau, P. et al., 2004. Seasonal ovulatory activity exists in tropical Creole female goats and Black Belly ewes subjected to a temperate photoperiod. *BMC physiology*, 4, p.12.

Corteel, J.M., Leboeuf, B. & Baril, G., 1988. Artificial breeding of adult goats and kids induced with hormones to ovulate outside the breeding season. *Small Ruminant Research*, 1(1), pp.19–35.

Cotton, A.M. et al., 2015. Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Human molecular genetics*, 24(6), pp.1528–39.

Couldrey, C. et al., 2014. Genome-wide DNA methylation patterns and transcription analysis in sheep muscle. *PLoS ONE*, 9(7), pp.1–7.

Couldrey, C. & Cave, V., 2014. Assessing DNA methylation levels in animals: Choosing the right tool for the job. *Animal Genetics*, 45(SUPPL.1), pp.15–24.

Deaton, A. & Bird, A., 2011. CpG islands and the regulation of transcription. *Genes & development*, 25(10), pp.1010–1022.

Dobin, A. et al., 2013. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), pp.15–21.

Dupont, C. et al., 2012. Maternal environment and the reproductive function of the offspring. *Theriogenology*, 78(7), pp.1405–1414.

Dupré, S.M. et al., 2008. Identification of melatonin-regulated genes in the ovine pituitary pars tuberalis, a target site for seasonal hormone control. *Endocrinology*, 149(11), pp.5527–39.

Ebling, F.J.P. & Luckman, S.M., 2008. RFAmide-related peptide: another sexy peptide? *Endocrinology*, 149(3), pp.899–901.

FAO, 2015. *The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture* Second Rep. S. B. D. & D. P. F. C. on G. R. for F. and Agriculture3, ed., Rome.

Farlik, M. et al., 2015. Single-Cell DNA Methylome Sequencing and Bioinformatic Inference of Epigenomic Cell-State Dynamics. *Cell Reports*, 10(8), pp.1386–1397.

Fatet, A., Pellicer-Rubio, M.-T. & Leboeuf, B., 2011. Reproductive cycle of goats. *Animal reproduction science*, 124(3–4), pp.211–9.

Franchini, D.-M., Schmitz, K.-M. & Petersen-Mahrt, S.K., 2012. 5-Methylcytosine DNA Demethylation: More Than Losing a Methyl Group. *Annu Rev Genet*, 46(August), pp.419–441.

Gabory, A., Attig, L. & Junien, C., 2011. Developmental programming and epigenetics 1 – 4. *Am J Clin Nutr*, 94, pp.1943–1952.

Galm, O. et al., 2002. Enzymatic regional methylation assay: a novel method to quantify regional CpG methylation density. *Genome research*, 12(1), pp.153–7.

Gattas, V. et al., 1990. Protein-energy requirements of prepubertal school-age boys determined by using the nitrogen-balance response to a mixed-protein diet. *The American journal of clinical nutrition*, 52(6), pp.1037–42.

Georges, A. et al., 2014. FOXL2: a central transcription factor of the ovary. *Journal of molecular endocrinology*, 52(1), pp.R17-33.

Giannopoulou, E.G. & Elemento, O., 2011. An integrated ChIP-seq analysis platform with customizable workflows. *BMC bioinformatics*, 12, p.277.

Gil, J. et al., 2004. Polycomb CBX7 has a unifying role in cellular lifespan. *Nature cell biology*, 6(1), pp.67–72.

Grieco, A. et al., 2013. Investigation of peripubertal expression of Lin28a and Lin28b in C57BL/6 female mice. *Molecular and cellular endocrinology*, 365(2), pp.241–8.

Guo, H. et al., 2010. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, 466(7308), pp.835–840.

Hackenberg, M. et al., 2006. CpGcluster: a distance-based algorithm for CpG-island detection. *BMC bioinformatics*, 7, p.446.

Han, L. & Zhao, Z., 2009. CpG islands or CpG clusters: how to identify functional GC-rich regions in a genome? *BMC bioinformatics*, 10, p.65.

Helm, B. et al., 2013. Annual rhythms that underlie phenology: biological time-keeping meets environmental change. *Proceedings of the Royal Society B.Biological sciences*, 280, p.20130016.

Hiller-Sturmhöfel, S. & Bartke, A., 1998. The endocrine system: an overview. *Alcohol health and research world*, 22(3), pp.153–64.

Hu, W. et al., 2014. MicroRNA mediates DNA methylation of target genes. *Biochemical and biophysical research communications*, 444(4), pp.676–81.

Hu, Y. et al., 2013. Comparison of the Genome-Wide DNA Methylation Profiles between Fast-Growing and Slow-Growing Broilers. *PLoS ONE*, 8(2).

Huang, D.W. et al., 2013. Analysis on DNA sequence of TSHB gene and its association with reproductive seasonality in goats. *Molecular Biology Reports*, 40(2), pp.1893–1904.

Huang, J. et al., 2011. Solexa sequencing of novel and differentially expressed microRNAs in testicular and ovarian tissues in Holstein cattle. *International journal of biological sciences*, 7(7), pp.1016–26.

Ibeagha-Awemu, E.M. & Zhao, X., 2015. Epigenetic marks: Regulators of livestock phenotypes and conceivable sources of missing variation in livestock improvement programs. *Frontiers in Genetics*, 6(SEP), pp.1–17.

Ikegami, K. & Yoshimura, T., 2012. Circadian clocks and the measurement of daylength in seasonal

reproduction. *Molecular and Cellular Endocrinology*, 349(1), pp.76–81.

Jaenisch, R. & Bird, A., 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature genetics*, 33 Suppl(march), pp.245–254.

Ji, Z. et al., 2012. Identification and characterization of microRNA in the dairy goat (Capra hircus) mammary gland by Solexa deep-sequencing technology. *Molecular Biology Reports*, 39(10), pp.9361–9371.

Jin, B., Li, Y. & Robertson, K.D., 2011. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes & cancer*, 2(6), pp.607–17.

Jones, M.J., Goodman, S.J. & Kobor, M.S., 2015. DNA methylation and healthy human aging. *Aging Cell*, 14(6), pp.924–932.

Jones, P.A., 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature reviews. Genetics*, 13(7), pp.484–92.

Kaminen-Ahola, N. et al., 2010. Maternal ethanol consumption alters the epigenotype and the phenotype of offspring in a mouse model. *PLoS Genetics*, 6(1).

Kaneko, H. et al., 1991a. Ovarian follicular dynamics and concentrations of oestradiol-17 beta, progesterone, luteinizing hormone and follicle stimulating hormone during the periovulatory phase of the oestrous cycle in the cow. *Reproduction, fertility, and development*, 3(5), pp.529–35.

Kaneko, H. et al., 1991b. Ovarian follicular dynamics and concentrations of oestradiol-17 beta, progesterone, luteinizing hormone and follicle stimulating hormone during the periovulatory phase of the oestrous cycle in the cow. *Reproduction, fertility, and development*, 3(5), pp.529–35.

Klose, R.J. & Bird, A.P., 2006. Genomic DNA methylation: The mark and its mediators. *Trends in Biochemical Sciences*, 31(2), pp.89–97.

Kuhn, D.E. et al., 2008. Experimental validation of miRNA targets. *Methods*, 44(1), pp.47–54.

Kurian, J.R. & Terasawa, E., 2013. Epigenetic control of gonadotropin releasing hormone neurons. *Frontiers in Endocrinology*, 4(MAY), pp.1–7.

Lai, P. et al., 2013. Polymorphism of the melatonin receptor genes and its relationship with seasonal reproduction in the gulin Ma goat breed. *Reproduction in Domestic Animals*, 48(5), pp.732–737.

Lan, X. et al., 2011. High resolution detection and analysis of CpG dinucleotides methylation using MBD-seq technology. *PLoS ONE*, 6(7).

Laurent, L. et al., 2010. Dynamic changes in the human methylome during differentiation. *Genome Research*, 20(3), pp.320–331.

Lee, J.-R. et al., 2014. Genome-wide analysis of DNA methylation patterns in horse. *BMC genomics*, 15(1), p.598.

Li, H. & Durbin, R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), pp.589–595.

Li, M. et al., 2011. Repertoire of porcine microRNAs in adult ovary and testis by deep sequencing. *International journal of biological sciences*, 7(7), pp.1045–55.

Lincoln, G.A., 2006. Melatonin entrainment of circannual rhythms. *Chronobiology international*, 23(1–2), pp.301–6.

Ling, Y.-H. et al., 2014. Identification and characterization of microRNAs in the ovaries of multiple and uniparous goats (Capra hircus) during follicular phase. *BMC genomics*, 15(1), p.339.

Livak, K.J. & Schmittgen, T.D., 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, 25(4), pp.402–408.

Lokk, K. et al., 2014. DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biology*, 15(4), p.r54.

Lomniczi, A. et al., 2013. Epigenetic control of female puberty. *Nature neuroscience*, 16(3), pp.281–9.

Lou, S. et al., 2014. Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. *Genome biology*, 15(7), p.408.

Maffucci, J.A. & Gore, A.C., 2009. Chapter 2 Hypothalamic Neural Systems Controlling the Female Reproductive Life Cycle. Gonadotropin-Releasing Hormone, Glutamate, and GABA. *International Review of Cell and Molecular Biology*, 274(C), pp.69–127.

Malpaux, B. et al., 1997. Control of the circannual rhythm of reproduction by melatonin in the ewe. *Brain Research Bulletin*, 44(4), pp.431–438.

Malpaux, B. et al., 1989. Regulation of the onset of the breeding season of the ewe: importance of long days and of an endogenous reproductive rhythm. *The Journal of endocrinology*, 122(1), pp.269–78.

Manfredi, E. et al., 2001. Genetic parameters of type appraisal in Saanen and Alpine goats. *Livestock Production Science*, 70(3), pp.183–189.

Martin, G.B. et al., 2004. Natural methods for increasing reproductive efficiency in small ruminants. *Animal reproduction science*, 82–83, pp.231–45.

Matzuk, M.M. & Lamb, D.J., 2002. Genetic dissection of mammalian fertility pathways. *Nature cell biology*, 4 Suppl, pp.s41-9.

McCarthy, M.M., 2013a. A piece in the puzzle of puberty. *Nature Publishing Group*, 16(3), pp.251–253.

McCarthy, M.M., 2013b. A piece in the puzzle of puberty. *Nature neuroscience*, 16(3), pp.251–3.

McFarlane, L. & Wilhelm, D., 2009. Non-coding RNAs in mammalian sexual development. *Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation*, 3(6), pp.302–16.

Medvedeva, Y.A. et al., 2014. Effects of cytosine methylation on transcription factor binding sites. *BMC genomics*, 15(1), p.119.

Mercer, T.R., Dinger, M.E. & Mattick, J.S., 2009. Long non-coding RNAs: insights into functions. *Nature reviews. Genetics*, 10(3), pp.155–9.

Messina, A. et al., 2016. A microRNA switch regulates the rise in hypothalamic GnRH production before puberty. *Nature Neuroscience*, 19(6), pp.835–844.

Migaud, M., Daveau, A. & Malpaux, B., 2005. MTNR1A melatonin receptors in the ovine premammillary hypothalamus: day-night variation in the expression of the transcripts. *Biology of reproduction*, 72(2), pp.393–8.

Miles, J.R. et al., 2012. MicroRNA expression profile in bovine cumulus–oocyte complexes: Possible role of let-7 and miR-106a in the development of bovine oocytes. *Animal Reproduction Science*, 130(1–2), pp.16–26.

Mishima, T. et al., 2008. MicroRNA (miRNA) cloning analysis reveals sex differences in miRNA expression profiles between adult mouse testis and ovary. *Reproduction (Cambridge, England)*, 136(6), pp.811–22.

Moore, L.D., Le, T. & Fan, G., 2013. DNA methylation and its basic function. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 38(1), pp.23–38.

Nair, S.S. et al., 2011. Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias. *Epigenetics*, 6(1), pp.34–44.

Naser, S.A. et al., 2012. Role of ATG16L, NOD2 and IL23R in Crohn's disease pathogenesis. *World journal of gastroenterology*, 18(5), pp.412–24.

Panchin, A.Y., Makeev, V.J. & Medvedeva, Y.A., 2016. Preservation of methylated CpG dinucleotides in human CpG islands. *Biology Direct*, 11(1), p.11.

Pellicer-Rubio, M.T. et al., 2016. Evaluation of hormone-free protocols based on the "male effect" for artificial insemination in lactating goats during seasonal anestrus. *Theriogenology*, 85(5), pp.960–969.

Perera, B.M.A.O., 2011. Reproductive cycles of buffalo. *Animal Reproduction Science*, 124(3), pp.194–199.

Plant, T.M., Ramaswamy, S. & Dipietro, M.J., 2006. Repetitive activation of hypothalamic G protein-coupled receptor 54 with intravenous pulses of kisspeptin in the juvenile monkey (Macaca mulatta) elicits a sustained train of gonadotropin-releasing hormone discharges. *Endocrinology*, 147(2), pp.1007–1013.

Rakyan, V.K. et al., 2012. Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics*, 12(8), pp.529–541.

Rivera, C.M. & Ren, B., 2013. Mapping human epigenomes. *Cell*, 155(1), pp.39–55.

Robinson, M.D., McCarthy, D.J. & Smyth, G.K., 2009. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), pp.139–140.

Rodríguez-Paredes, M. & Esteller, M., 2011. Cancer epigenetics reaches mainstream oncology. *Nature medicine*, 17(3), pp.330–339.

Roush, S. & Slack, F.J., 2008. The let-7 family of microRNAs. *Trends in Cell Biology*, 18(10), pp.505–516.

Sairam, M.R. & Li, C.H., 1977. Human pituitary thyrotropin. The primary structure of the alpha and beta subunits. *Canadian journal of biochemistry*, 55(7), pp.755–60.

Sangiao-Alvarellos, S. et al., 2013. Changes in hypothalamic expression of the Lin28/let-7 system and related microRNAs during postnatal maturation and after experimental manipulations of puberty. *Endocrinology*, 154(2), pp.942–55.

Sati, S. et al., 2012. High resolution methylome map of rat indicates role of intragenic DNA methylation in identification of coding region. *PLoS ONE*, 7(2), pp.1–12.

Sharp, A. et al., 2011. DNA methylation profiles of human active and inactive X chromosomes. *Genome research*, 21(10), pp.1592–1600.

Shinomiya, A. et al., 2014. Regulation of seasonal reproduction by hypothalamic activation of thyroid hormone. *Frontiers in Endocrinology*, 5(FEB), pp.1–7.

Simon, J.A. & Kingston, R.E., 2009. Mechanisms of Polycomb gene silencing: knowns and unknowns. *Nature Reviews Molecular Cell Biology*, 10, pp.1–12.

Smith, Z.D. & Meissner, A., 2013. DNA methylation: roles in mammalian development. *Nature reviews. Genetics*, 14(3), pp.204–20.

Soliman, A., De Sanctis, V. & Elalaily, R., 2014. Nutrition and pubertal development. *Indian journal of endocrinology and metabolism*, 18(Suppl 1), pp.S39-47.

Srivastava, V.K., Hiney, J.K. & Dees, W. Les, 2011. Hypothalamic glial-to-neuronal signaling during puberty: influence of alcohol. *International journal of environmental research and public health*, 8(7), pp.2876–94.

Stanford, J.C. et al., 2014. Efferocytosis produces a prometastatic landscape during postpartum mammary gland involution. *Journal of Clinical Investigation*, 124(11), pp.4737–4752.

Straussman, R. et al., 2009. Developmental programming of CpG island methylation profiles in the human genome. *Nature structural & molecular biology*, 16(5), pp.564–571.

Sturn, A., Quackenbush, J. & Trajanoski, Z., 2002. Genesis: cluster analysis of microarray data. *Bioinformatics (Oxford, England)*, 18(1), pp.207–8.

Su, J. et al., 2014. Genome-wide analysis of DNA methylation in bovine placentas. *BMC genomics*, 15(1), p.12.

Szklarczyk, D. et al., 2015. STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, 43(D1), pp.D447–D452.

Szyf, M., 2012. DNA methylation signatures for breast cancer classification and prognosis. *Genome medicine*, 4(3), p.26.

Tena-Sempere, M., 2013. Keeping puberty on time: novel signals and mechanisms involved. *Current topics in developmental biology*, 105, pp.299–329.

Terzano, G. & Terzano, 2012. Overview on reproductive endocrine aspects in buffalo. *Journal of Buffalo Science*, pp.126–138.

Thrun, L.A. et al., 1997. A Critical Period for Thyroid Hormone Action on Seasonal Changes in Reproductive Neuroendocrine Function in the Ewe [1]. *Endocrinology*, 138(8), pp.3402–3409.

Todini, L. et al., 2006. Plasma total T3 and T4 concentrations in bucks as affected by photoperiod. *Small Ruminant Research*, 65(1), pp.8–13.

Tomikawa, J. et al., 2012. Epigenetic regulation of Kiss1 gene expression mediating estrogen-positive feedback action in the mouse brain. *Proceedings of the National Academy of Sciences of the United States of America*, 109(20), pp.E1294-301.

Triantaphyllopoulos, K.A., Ikonomopoulos, I. & Bannister, A.J., 2016. Epigenetics and inheritance of phenotype variation in livestock. *Epigenetics & chromatin*, 9, p.31.

Trimmomatic, 2014. a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, p.2114.

Tsang, J., Zhu, J. & van Oudenaarden, A., 2007. MicroRNA-Mediated Feedback and Feedforward Loops Are Recurrent Network Motifs in Mammals. *Molecular Cell*, 26(5), pp.753–767.

Untergasser, A. et al., 2012. Primer3--new capabilities and interfaces. *Nucleic acids research*, 40(15), p.e115.

Vasantha, I., 2015. Physiology of Seasonal Breeding: A Review. *Journal of Veterinary Science & Technology*, 7(3), pp.1–4.

Viguié, C. et al., 1997. Characterization of the Short Day-Induced Decrease in Median Eminence Tyrosine Hydroxylase Activity in the Ewe: Temporal Relationship to the Changes in Luteinizing Hormone and Prolactin Secretion and Short Day-Like Effect of Melatonin [1]. *Endocrinology*, 138(1), pp.499–506.

Waddington, C.H., 1959. Canalization of Development and Genetic Assimilation of Acquired Characters. *Nature*, 183(4676), pp.1654–1655.

Wan, J. et al., 2015. Characterization of tissue-specific differential DNA methylation suggests distinct modes of positive and negative gene expression regulation. *BMC Genomics*, 16(1), p.49.

Windsor-Engnell, B.M. et al., 2007. An increase in in vivo release of LHRH and precocious puberty by posterior hypothalamic lesions in female rhesus monkeys (Macaca mulatta). *American journal of physiology. Endocrinology and metabolism*, 292(4), pp.E1000-9.

Winter, J. et al., 2009. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nature cell biology*, 11(3), pp.228–234.

Woo, C.J. et al., 2010. A Region of the Human HOXD Cluster that Confers Polycomb-Group Responsiveness. *Cell*, 140(1), pp.99–110.

Wood, S. et al., 2006. Seasonal variation in assisted conception cycles and the influence of photoperiodism on outcome in in vitro fertilization cycles. *Human fertility (Cambridge, England)*, 9(4), pp.223–9.

Wu, J. et al., 2014. Identification of conservative microRNAs in Saanen dairy goat testis through deep sequencing. *Reproduction in domestic animals = Zuchthygiene*, 49(1), pp.32–40.

Xu, T. et al., 2013. Differential gene expression analysis between anagen and telogen of Capra hircus skin based on the de novo assembled transcriptome sequence. *Gene*, 520(1), pp.30–38.

Yang, C. et al., 2016. DNA Methylation Patterns in the Hypothalamus of Female Pubertal Goats. *Plos One*, 11(10), p.e0165327.

Yuan, X.-L. et al., 2016. Profiling the genome-wide DNA methylation pattern of porcine ovaries using reduced representation bisulfite sequencing. *Scientific Reports*, 6(April 2015), p.22138.

Zetouni, L. et al., 2014. Polymorphisms in the MTRN1A gene and their effects on the productive and reproductive traits in buffaloes. *Tropical Animal Health and Production*, 46(2), pp.337–340.

Zhang, X.-D. et al., 2013. Characterization and differential expression of microRNAs in the ovaries of pregnant and non-pregnant goats (Capra hircus). *BMC genomics*, 14(1), p.157.

Zicarelli, L. et al., 1997. Effects of using vasectomized bulls in artificial insemination practice on the reproductive efficiency of Italian buffalo cows. *Animal reproduction science*, 47(3), pp.171–80.

Ziller, M.J. et al., 2013. Charting a dynamic DNA methylation landscape of the human genome. *Nature*, 500(7463), pp.477–481.

## Web references

http://www.fao.org/faostat

http://www.goatit.eu

http://www.lifetechnologies.com

http://www.epibeat.com

http://www.mirbase.org

http://www.bmr-genomics.it

http://www.mbio.ncsu.edu/bioedit