A **THESIS**
submitted for a Master's Degree
**Predictive and integrative animal biology (PRIAM)**

# Genetic Determinism of Circadian Activity in Dairy Cattle.

by

**Lydia GHOZLANE**

Supervisors: Herve Acloque

Marie-Pierre Sanchez

Study carried out:

GABI unit – EFISA and GBOS Teams
INRAE Jouy-en-Josas
4 Rue De La Manufacture
78350 Jouy-en-Josas
France

**- 2025 -**

## Acknowledgments

I would like to express my sincere gratitude to my supervisors, Hervé Acloque and Marie-Pierre Sanchez, for their guidance and support throughout this study. I am truly grateful for the trust you placed in me to work on this project --thank you for this opportunity. It was such a pleasure to work on such a vast and fascinating subject.

I am also thankful to Arthur Leroy for his technical support.

I extend my warmest thanks to the members of both EFISA and GBOS teams for their kind welcome, support, and the enriching discussions I had the pleasure of sharing during my time here. In particular, I would like to thank Valentin Sorin, whose help was invaluable throughout this internship, as well as Khadiia El Moumen, Angelo Fierro, Yves Plantard and Armelle Venot.

I am also thankful for the professors of the PRIAM master's program for the quality of their teaching and their guidance.

I gratefully acknowledge the financial support of **GIS Avenir Elevage**, which made this internship possible.

Finally, I also thank my family and loved ones for their unwavering support and encouragement.

# Table of contents

## List of figures:



## List of tables:

## List of abbreviations

**ALCAM**: Activated Leukocyte Cell Adhesion Molecule

**ARNTL**: Aryl Hydrocarbon Receptor Nuclear Translocator-Like Protein 1

**BMAL1**: Brain and Muscle ARNT-Like 1

**BMAL2**: Brain and Muscle ARNT-Like 2

**C2C12**: Mouse Myoblast Cell Line

**CACHD1**: Cache Domain Containing 1

**CANNTG**: Consensus E-box DNA sequence recognized by clock proteins

**CD166**: Cluster of Differentiation 166

**CLOCK**: Circadian Locomotor Output Cycles Kaput (transcription factor)

**CRY**: Cryptochrome Circadian Regulator

**CRY1**: Cryptochrome Circadian Regulator 1

**CRY2**: Cryptochrome Circadian Regulator 2

**CTNNA2**: Catenin Alpha 2

**Dex**: Dexamethasone

**DNA**: Deoxyribonucleic Acid

**E-box**: Enhancer Box

**EM algorithm**: Expectation-Maximization Algorithm

**FERG**: Fertilité-Génisse (Heifer fertility)

**FERV**: Fertilité-Vache (Cow fertility)

**FOXO1**: Forkhead Box Protein O1

**FSH**: Follicle Stimulating Hormone

**FSHR**: Follicle Stimulating Hormone Receptor

**Fsk**: Forskolin

**GRM**: Genomic Relationship Matrix

**GWAS**: Genome-Wide Association Study

**HD**: High Density

**Indels**: Insertions and Deletions

**IVIA** : Intervalle-vêlage-1ère-insémination-artificielle

**LH**: Luteinizing Hormone

**LIMCH1**: LIM and Calponin Homology Domains 1

**LRFN2**: Leucine Rich Repeat and Fibronectin Type III Domain Containing 2

**LRNF5**: Leucine Rich Repeat and Fibronectin Type III Domain Containing 5

**MAF**: Minor Allele Frequency

**mlma**: Mixed Linear Model Analysis

**NPAS2** : Neuronal PAS Domain Protein 2

**PER**: Period Circadian Protein

**PER1**: Period Circadian Protein 1

**PER2**: Period Circadian Protein 2

**PER3**: Period Circadian Protein 3

**QTL**: Quantitative Trait Locus

**RNA**: Ribonucleic Acid

**U7, 7SK**: Small nuclear RNAs; U7 processes histone pre-mRNAs, 7SK regulates transcription elongation.

**RUNX1**: Runt-Related Transcription Factor 1

**SCN**: Suprachiasmatic Nucleus

**SNP**: Single Nucleotide Polymorphism

**TNN**: Tenascin N

**UEP**: Unité Expérimentale du Pin

**VEP**: Variant Effect Predictor

**ZBTB20**: Zinc Finger and BTB Domain Containing 20

## I. Introduction

Distinct daily behavioral and/or physiological rhythms have been observed across animals, plants, fungi and bacteria. known as circadian rhythms, derived from the Latin ‘circa diem’ or about a day. It is driven by an autonomous, intrinsic timekeeping system called the circadian clock, running within cycles that operate in 24 hours approximately, regulating various biochemical, physiological or behavioral processes of organisms (Pittendrigh,1993; Sehgal, 2017).

This rhythmicity is generated at the cellular level by molecular clocks present in all cells of the body. In vertebrates, the master clock is located at the suprachiasmatic nucleus of the hypothalamus. It functions as pacemaker that receives environmental signals captured by the retina, known as zeitgebers (German for ‘time giver’), as well as physiological cues (e.g., nutritional status). As indicated by the extant literature, the master clock is set at the molecular level, thus coordinating peripheral clocks in most cells in every organ of the body (Hastings et al., 2007; Koch et al., 2016; Plaut & Casey, 2011; Xie et al., 2019). These peripheral clocks, in turn, regulate the circadian expression of local genes, thereby orchestrating metabolism and physiological functions across all the organism.

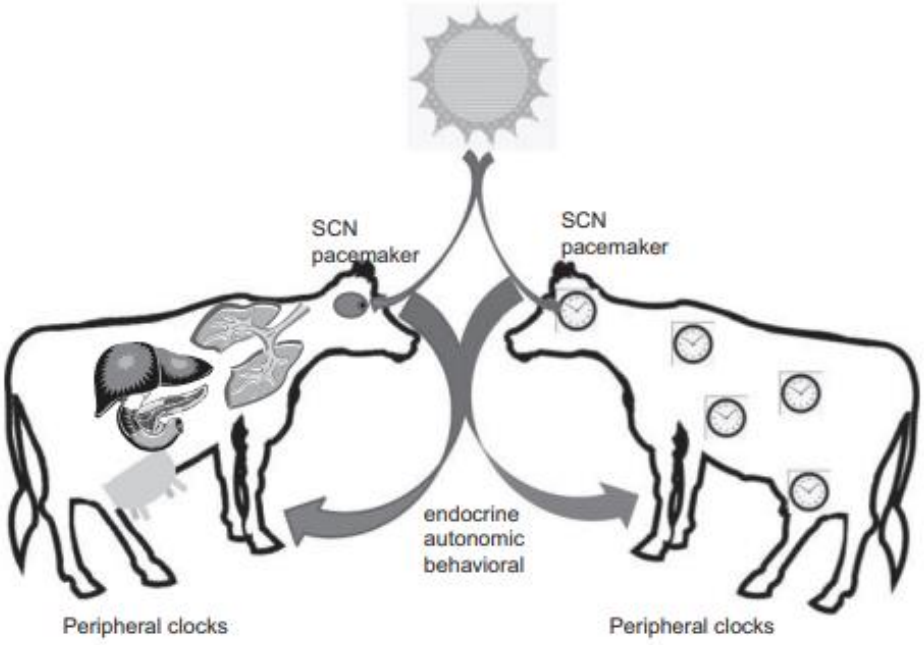


**Figure 01: Schematic representation of the circadian rhythm: coordination between internal physiology and environmental cues.** Adapted from Hastings et al 2007.

This illustration depicts the circadian organization in cows, highlighting the coordination between the suprachiasmatic nucleus (SCN) and peripheral clocks. On the left, the SCN of the cow, located in the brain, acts as the central pacemaker, receiving light cues and synchronizing peripheral clocks present in most tissues through endocrine, autonomic, and behavioral signals.

## 1. Chronotypes

Chronotype refers to an individual's preference or proxy for circadian rhythm expression, reflected in various biological and behavioral processes, such as, body temperature, cortisol secretion, eating and sleeping habits . It is shaped by a combination of internal factors such as genetics and external influences such as light exposure and photoperiod, and plays an important role in health and well-being (Chauhan et al., 2023).

In humans, three chronotypes are defined. They range from morning types (M-types) to evening types (E-types), with intermediate neither types (N-types) (Montaruli et al., 2021). Like humans, animals may exhibit variations in their activities and rhythms. These differences, driven by genetic and environmental factors, play a significant role in determining the timing of key behaviors such as feeding and resting. In livestock, understanding these chronotype-like variations can help optimize management practices, including feeding, milking schedules, and breeding strategies to better align with the animal natural rhythms and enhance their overall health and welfare, as well as their productivity.

## 2. The relationship between the circadian system and animal physiology and metabolism

Circadian rhythms regulate essential physiological functions and metabolic processes across organs. In the kidney, they control blood flow, filtration, ion excretion through rhythmic expression of transport proteins (Solocinski & Gumz, 2015). In the pancreas, they regulate insulin and glucagon secretion (Petrenko et al., 2017). In skeletal muscles, they synchronize processes like respiration and autophagy for energy production (Woldt et al., 2013). The liver coordinates synthesis of key metabolic compounds (Reinke & Asher, 2015), whereas the gastrointestinal tract, in coordination with the microbiome, supports diurnal metabolism (Lynch & Pedersen, 2016), as the response of the circadian clock to metabolic challenges is importantly affected by the microbiome, which in turn is regulated by the clock through the timing of food intake (Reinke & Asher, 2019).

All in all, the circadian system regulates various levels of biological functions, from cellular metabolism to organ coordination. However, ample evidence indicates that metabolic regulation isn't merely just an output of the circadian clock, rather, it also provides an input to the circadian clock. This output-input feedback helps the circadian system stay flexible to adjust physiology to the metabolic needs of cells, tissues and the entire body (Reinke & Asher, 2019).

## 3. The molecular basis of the circadian rhythm

Having established the fundamental role of circadian rhythms in regulating biological functions, we now delve into the genetic mechanism underlying this clock. At the core of this regulation lies a set of genes that drive the circadian rhythm at the molecular level functioning within a feedback loop of positive and negative elements. In mammals, the **positive loop** involves the products of *CLOCK* and *BMAL1* (Brain and Muscle Aren't Like, also known as ARNTL, for Aryl hydrocarbon Receptor Nuclear Translocator-Like) genes, which act as transcription factors (King et al., 1997; Hogenesch et al., 1998; Gekakis et al., 1998). The **negative loop** contains *CRY* (cryptochrome) and *PER* (period) gene families (Sun et al., 1997; Tei et al., 1997). A molecular redundancy is observed, with three PER genes (*PER1*, *PER2*, *PER3*) and two *CRY* genes (*CRY1*, *CRY2*) genes. In addition, paralogues of *BMAL1* like *BMAL2* and *NPAS2*, have been identified (Kume et al., 1999).

In the daytime, the BMAL1-CLOCK complex binds to specific DNA sequences called enhancer box sequences (E-boxes, nucleotides, CANNTG) in the promoter region of clock-controlled genes. This binding activates the transcription of *PER* and *CRY* genes, leading to the accumulation in the cytoplasm of PER and CRY proteins in the afternoon or evening, gradually slowing their own production as they inhibit the transcription of Per and Cry, thereby establishing a negative feedback loop.

Finally, the core circadian oscillator can generate transcriptional cycles of with various phases of expression depending on the presence and combination of cis-elements such as E-boxes in the promoters and enhancers of specific target genes (Ueda et al., 2005). But, they don't operate in isolation. These endogenous oscillations are fine-turned by environmental and physiological cues -known as inputs- to ensure alignment with the light-dark cycle, thus, contributing to the temporal organisation of gene expression across tissues.

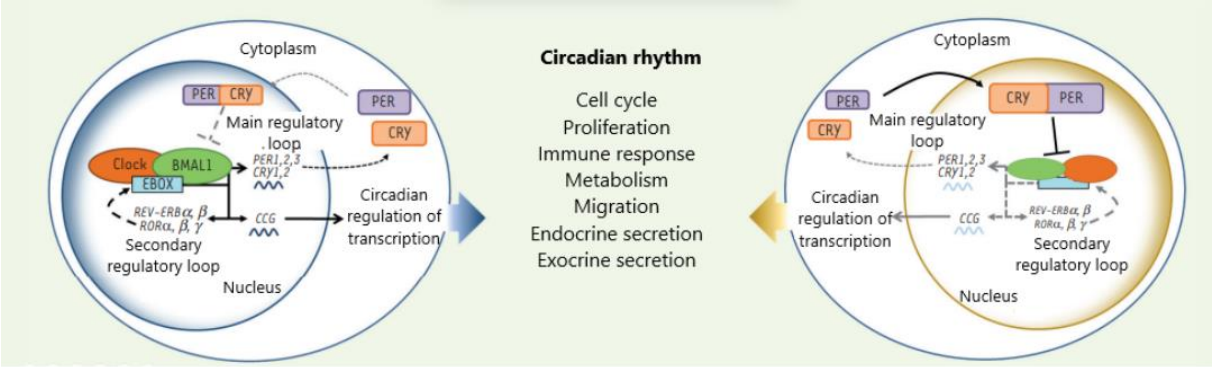


**Figure 02: Molecular mechanism of the circadian rhythm in mammals.** Adapted from (Hadadi & Acloque 2021. The core mechanism relies on transcription-translation feedback loop involving CLOCK and BMAL1, which activate PER and CRY gene expression. PER and CRY accumulate in the cytoplasm, then inhibit CLOCK-BMAL1 complex activity. A secondary loop involves REV-ERBα/β and RORα/β/γ that regulate BMAL1 transcription. These loops control the circadian expression of downstream genes (Clock-Controlled-Genes -- CCG) influencing physiological functions such as immune response and metabolism. The left panel shows activation of circadian transcription, the right panel represents the repression phase, completing the daily cycle of rhythmic gene expression.

### 4. Inputs to the master clock in the SCN and Circadian rhythm synchronizers

Light is the primary input to the SCN with the duration of light acting as the most prominent environmental cue for synchronizing circadian clocks (Reppert & Weaver, 2002). Seasonal variations in light influence the phase relationships among cellular circadian oscillators in the SCN, allowing it to gradually adjust its rhythm to changing day lengths (Porcu et al., 2018; Coomans et al., 2015). In response, the SCN transmits photoperiod information to peripheral clocks via rhythmic outputs, notably melatonin, for which circulating levels vary with the length of the dark phase and act as a neuroendocrine mediator of photoperiod (Cipolla-Neto et al., 2014). While light is the dominating entraining cue, other factors such as activity, feeding, stress, temperature, and arousal also modulate circadian timing. These cues, referred to as "synchronizers", "zeitgeber," or "entraining agents", help align both central and peripheral clocks with environmental time (Dardente, 2007; Ebling & Barrett, 2008).

Almost all eukaryotic cells possess a cell-autonomous circadian clock that needs to be influenced by external cues, to align cellular rhythms, in a process called circadian rhythm synchronization, ensuring proper coordination between central and peripheral clocks. Zeitgebers include light, feeding, temperature and hormonal signals as summarised in figure 03.

## 5. Other Non-photic Cues

Feeding acts as a potent non-photic zeitgeber, time and composition of food intake can modulate the amplitude and phase of circadian rhythms independently from the light-dark cycle (Oike et al., 2014; Zerón-Rugerio et al., 2019; Panda, 2016). Furthermore, feeding shape the composition and activity of the gut microbiota, which in turn feeds back into host circadian regulation (Leone et al., 2015; Thaiss et al., 2016). In addition to feeding, stress and arousal influence circadian timing through serotonergic pathways and melatonin signalling affect neural activity in the SCN (Moore et al., 1978; Bosler & Beaudet, 1985; Meyer-Bernstein & Morin, 1996; Jacobs et al., 2002; Moore & Speh, 2004; Wirz-Justice, 2006).

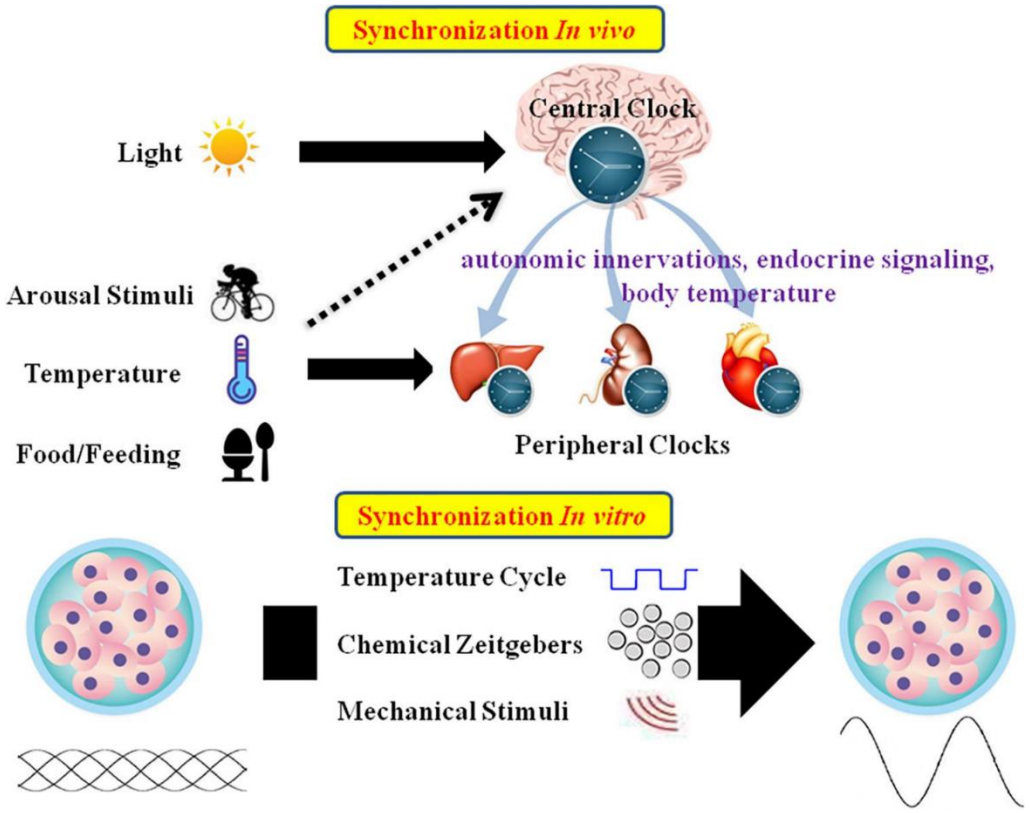


**Figure 03: Schematic summary of *in vivo* and *in vitro* circadian synchronization.** *In vivo*, the photic zeitgeber mainly entrains the central clock, which regulates the peripheral clocks through the internal timing cues including autonomic innervations, endocrine signalling and body temperature; the non-photic zeitgebers including arousal stimuli, temperature and food mainly entrain the peripheral clocks. *In vitro*, the circadian oscillations of cells or explants can be synchronized by temperature cycles, chemical factors (such as Dex, Fsk, or horse serum) and mechanical stimuli (X. Yanling,. Et al 2019).

## 6. Health and welfare

Disruption of circadian synchronizers such as light exposure, feeding schedules, and other environmental cues can significantly impair circadian rhythms. In humans, such disruptions are associated with serious health risks including cancer progression (Momma et al., 2017; Hadadi & Acloque, 2021), depression (Smolensky et al., 2016) and metabolic diseases, as the asynchrony between the inner clock (SCN) and peripheral clocks is considered a risk factor for metabolic dysregulation (Cheng et al., 2021).

In dairy cows, circadian disruption has been linked to heightened vulnerability to both metabolic and infectious diseases. Misalignment of light-dark cycles or feeding routines can impair glucose metabolism and insulin sensitivity, increasing the risk of metabolic disorders, such as ketosis and fatty liver during the transition period from late pregnancy to early lactation (Casey & Plaut, 2022). Furthermore, proteomic analyses show that continuous circadian disruption during late gestational period alters muscle protein expression and increases oxidative stress compromising cow's overall health and performance (McCabe et al., 2021).

These findings highlight the importance of maintaining circadian rhythms, especially in metabolically challenged cows, to lower disease risk and enhance reproductive efficiency and animal welfare (Casey & Plaut, 2022).

## 7. Dairy cattle farming in France

France is Europe's second largest bovine milk producer after Germany. Within December 2024, over 3,3 million dairy cows, mostly of Holstein, Montbéliarde and Normande Breeds, respectively, with around 46000 exploitations in 2023, spread throughout the country with key regions being Brittany, Normandy, and Pays de la Loire (*Institut De L'Élevage. (2023). Bilan Génétique Des Races Bovines Laitières En 2023. Institut De L'Élevage.*, n.d.).

Lately, dairy cattle farming is undergoing significant structural and technological shift as the sector sees a politic of a consolidation of farms, fewer in number of exploitations but larger in herd size alongside an adoption of precision livestock farming aiming at improving efficiency, sustainability and animal health. One innovation is the use of MEDRIA collars, which allow real-time monitoring of activity and physiological events like heat and calving. These tools provide high-resolution, continuous data, essential for studying variability in behavior, metabolism and health traits in cows.

When combined with genomic data, they open new possibilities for genetic evaluation and selection. Traits recorded via these collars, such as activity rhythms, may showcase significant heritability, making them a good target for genome-wide association studies (GWAS) and genomic selection.

## 8. Genomic selection of cattle

The development of genome-wide DNA (deoxyribonucleic acid) genotyping arrays and chips, particularly made possible by the whole-genome sequencing of the first reference- bovine genome and the 1000 Bull Genomes Project (Hayes & Daetwyler, 2018), enabled the genotyping of cattle and thus the implementation of genomic evaluations. These DNA chips contain SNPs (Single Nucleotide Polymorphisms) that are variations in single nucleotides occurring at specific positions in the genome. They are highly abundant (dozens of millions) and are well distributed across the genome, making them particularly useful for genetic studies.

These bi-allelic variations contribute to the genetic diversity between individuals and are the forefront of association studies with phenotypes of interest. In bovine, there exists different chips with different SNP densities, among them, the EuroGMD Beadchip (*Illumina Inc | Microarray Kits for Genotyping & Epigenetics*, n.d.) is currently used in France for genomic selection, considering a set of 54,609 SNP.

For a long time, genetic evaluation for cattle selection was based on pedigree data, but since 2009, with the major development of DNA chips, genomic information has been incorporated, allowing more accurate predictions of genetic values (Boichard et al., 2012).

While genomic selection significantly enhances production (e.g., milk yield) and functional (e.g., udder health and fertility) traits in dairy cows, it can also result in the unintended co-selection of deleterious alleles linked to those traits. This may disrupt the balance between key biological functions, leading to trade-offs that compromise animal health and welfare. Such unintended consequences of selection have been documented in other species—for instance, in cultivated tomato, where the selection process has been associated with a deceleration of the circadian clock (Müller et al., 2015). To date, no studies evaluate the impact of genomic selection on the circadian rhythm of highly selected breeds for dairy production, like Holstein cows.

Given the potential impact of selection on biological rhythms, this study aims to investigate individual variations in circadian rhythmicity among dairy cows. We hypothesize that cows exhibit distinct circadian chronotypes, showcasing differences in behavioral timing, and that these chronotypes may be explained by genetic variation. Our objectives aim to identify chronotypes, asses their genetic variability, and explore possible associations with health and performance traits.

## II. Materials and methods

### 1. Ethical approval and animal experimentation

For this study, no animal experimentation was conducted by the author. Data was obtained from the Pin experimental unit of INRAE, as all animals were handled with care in accordance with the French ministry of Agriculture guidelines for animal research and the applicable European Union guidelines and regulations on animal experiments (UEP, INRAE, 2019., https://doi.org/10.15454/1.5483257052131956E12).

### 2. Animal population and data collection

Activity data was collected from two dairy cattle breeds, Holstein and Normande, 547 and 227 cows, respectively, from 2020 to 2024. Activity was recorded through automated sensor MEDRIA collars developed by MEDRIA solutions, a company specialized in solutions for monitoring and controlling bovine health. These collars were developed primarily for heat detection but also provide valuable indicators related to feeding behavior, health status, reproduction, general well-being, and animal surveillance. The activity data were exported in CSV files. Each file included the following colums: farm_id, animal_id, date, hour, ingestion_trough_pasture, rumination, rest, other_activity, over_activity, and standing_up. For each 5-minute interval, a single predominant activity was recorded, resulting in 288 rows per cow per day (corresponding to the 288 five-minute periods in a 24-hour cycle). The codes used for each activity are described in Table X. All cows were genotyped using the Illumina EuroGMD SNP chip, which includes 54,609 SNPs distributed across the bovine genome.

**Table 01: Activity data from MEDRIA collars lexica.** This table summarize animal behavioral activities using a categorical coding system. The variable Ingestion_through_pasture indicates feeding behavior, where 0 denotes no ingestion, 1 represents ingestion at the trough, and 2 refers to grazing activity at pasture. Rumination is coded as 0 if no rumination was observed and 1 if rumination occurred. Rest is similarly coded, with 0 for no rest and 1 indicating a resting period. Other_activity refers to miscellaneous behaviors not classified under ingestion, rumination, rest, over-activity, or standing; it is coded as 0 for no such activity and from 1 to 10 depending on the intensity or duration of the observed activity. Over_activity captures elevated levels of activity potentially associated with heat stress, ranging from 0 (no signs of over-activity) to 10 (intense over-activity or stress-related behavior). Finally, Standing_up reflects posture, where 0 means the animal is lying down and 1 indicates it is standing.

| Activity | Lexica |
|---|---|
| **Ingestion_trough_pasture** | 0: no ingestion,<br>1: detected ingestion at the trough,<br>2: detected ingestion grazing at pasture. |
| **Rumination** | 0: no rumination,<br>1: detected rumination. |
| **Rest** | 0: no rest,<br>1: detected rest. |
| **Other_activity** | 0: no other activity,<br>1 to 10: detected other activity as it corresponds to periods of activity remaining excluding the other 4. |
| **Over_activity** | 0: no over activity,<br>1 to 10: detected over activity as it corresponds to periods with specific heat stress signs or expression. |
| **Standing_up** | 0: No, lying down,<br>1: standing up. |

### 3. Circadian activity Analysis

With the help of another dataset that provides information on which cows were outdoors, we only selected cows that we confirmed to be outdoors during the period from May to July, to ensure consistency in environmental conditions influencing behavioral rhythms with stable photoperiod. Data was reprocessed, each 5 minutes interval was considered a total of activity and we summed it into an hourly resolution to have a smooth 24h profile. As ingestion location (trough vs pasture) was distinguished, we retained only global ingestion (ingestion > 0) for further analysis, as all cows had access to pasture during the period. Data of collars showing evidence of errors or biologically implausible values were flagged and removed (e.g., cows for which a full 24-hour period showed: ingestion = 0, rumination = 0 or 1440 minutes, or rest = 0 or 1440 minutes, no abnormal data with ingestion = 1440 was detected). For each cow, the average number of minutes spent per hour on each activity was computed over the entire observation period. To reveal circadian individuality and compare these cows, we calculated a deviation score from the global population mean. This score quantified how much a given cow's activity at each hour deviated from the population average, thereby capturing individual chronotype patterns, i.e., tendencies to be more or less active than average at specific times of day.

### 4. Clustering - chronotype detection

To identify individual chronotypes, we analysed deviations in hourly ingestion and rest from the population mean over a 24h cycle. We used the MagmaClustR package that implements two main algorithms, called Magma (Leroy et al., 2022) and MagmaClust (Leroy et al., 2023), using a multi-task gaussian processes (GP) model to perform predictions for supervised learning problems. The method clusters animals into groups, each associated to specific mean process, based on temporal activity profiles. Clustering was performed separately for ingestion and rest behavior. Cows identifiers were structured in the dataset as **ID**, the hour of the day was used as the temporal **input**, and the ingestion/rest deviation was used as the **output** or response variable.

The model was trained using the train_magmaclust() function with default settings. After convergence of the variational EM algorithm, cows were allocated to their most probable cluster using the data_allocate_cluster() function, each cluster representing a chronotype.

To evaluate the robustness and generalizability of the clustering model, we applied an 80/20 train-test split strategy. Specifically, 80% of the cows were randomly selected and used to train the Magma clustering model, while the remaining 20% were used to test the model's predictive ability. This approach allowed us to evaluate the model's robustness and its potential for generalization across cows not included in the clustering phase. Correlation was calculated between true values and predicted values in the testing population.

To further test robustness, we later applied a leave-one-out-like strategy, training the model on all cows except one and assessing cluster assignment for the excluded individual.

In preparation for the genome wide association studies, we increased the reliability on chronotype definitions focussing only on cows that overlapped between ingestion and rest-based clusters. This intersection was visualised using Venn diagrams. Given the biological assumption that a cow cannot simultaneously be engaged in ingestion and rest, we retained only animals whose chronotype assignments overlapped in both ingestion and rest-based clustering.

## 5. Genome wide association studies

a. Animals, phenotypes and genotypes:

A total of 432 Holstein and 172 Normande cows intersected between the clusters of rest and ingestion clusters. Chronotypes derived from the clustering were then considered as categorical traits for the GWAS, 0,1 for clusters 1, 2 of ingestion, respectively.

Thereafter, these phenotyped animals were cross-referenced with the database that contains EuroGMD chip genotyping data processed and used for genomic selection.

Based on the chip genotyping data, sequence level genotypes were obtained through a two-step imputation process. The first step imputation was carried out on the EUROGMD density (54,609 SNP) to the HD density (777,000 SNP) using FImpute software (Sargolzaei et al., 2014). The second step imputed from the HD density to the full sequence using Minimac software (Howie et al., 2012).

b. GWAS:

Genome wide association studies were performed using sequence-level genotype density. Association analyses were conducted using the mlma option (Mixed Linear Model Analysis) of the GCTA software (Yang et al., 2010), which applies a mixed linear model:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{x}b + \mathbf{u} + \mathbf{e}$$

Where **y** is the phenotype vector (coded as 0/1); μ is the overall mean; b is the additive fixed effect of the variant to be tested; **x** is the vector of imputed genotypes coded as dosages varying from 0 to 2 (number of copies of the second allele); $\mathbf{u} \sim N(0, \mathbf{G}\sigma^2_u)$ is the vector of random polygenic effects, with **G** the Genomic Relationship Matrix (GRM) that is calculated using the 50K SNP genotypes, $\sigma^2_u$ the polygenic variance that is estimated based on the null model $\mathbf{y} = \mathbf{1}\mu + \mathbf{u} + \mathbf{e}$ and then fixed while testing for the association between each variant and the trait of interest; and $\mathbf{e} \sim N(0, \mathbf{I}\sigma^2_e)$ is the vector of random residual effects, with **I** the identity matrix and $\sigma^2_e$ the residual variance. Heritability ($h^2$) was estimated as the ratio of genetic variance to total phenotypic variance.

$$\text{heritability} = \sigma^2_u / (\sigma^2_u + \sigma^2_e) = variance\ génétique/variance\ totale$$

As heritability ($h^2$) is the proportion of phenotypic variance explained by genetic additive effects and indicates the potential effectiveness of selection. It is specific to a population and ranges from 0 to 1, increasing with the similarity of performance among related animals. Traits are considered lowly heritable if $h^2 < 0.20$, moderately heritable if $0.20 < h^2 < 0.40$, and highly heritable if $h^2 > 0.40$.

For sequence-level GWAS, only variants with MAF>0.01 and imputation accuracy of $R^2 > 0.2$ were retained. To account for multiple testing, we applied the Bonferroni correction, assuming 50,000 independent genomic regions. Therefore, the genome-wide threshold at the 5% level corresponded to a nominal P value of $10^{-6}$ ($-\log10(P)=6$).

c. QTL identification:

Quantitative Trait Loci (QTL) are genomic regions that influence phenotypic variation of a complex trait, commonly identified through a statistical analysis called QTL mapping (Powder, 2020). When performed at the sequence level, GWAS increases detection power and resolution

as we have access to a more exhaustive set of genomic variants (Daetwlyer et al., 2014). However, identifying one single candidate variant remains challenging because of strong LD present in a QTL region. To refine candidate identification, GWAS results can be cross-referenced with functional cattle genome annotations from resources such as Ensembl (McLaren et al., 2016). This approach helps identifying candidate variants based on their position in functional regions of the genome, within genes or their coding regions (Sanchez, 2019).

In this study, significant SNPs with p-values $< 1 \times 10^{-5}$ were extracted from the GWAS results and visually inspected across chromosomes. Due to the limited number of significant signals, QTL regions were defined manually by grouping nearly SNPs based on their chromosomal proximity in bp (base pairs). Variant annotation was done using Ensembl genome browser (Dyer et al., 2024). SNPs were classified in genic or intergenic regions. In the case where the variants with significant results were located in intergenic regions, we retrieved their gene proximity.

### 6. Performance indicators

In this part of study, we explored potential associations between chronotype and reproductive performance in both breeds separately, Holstein and Normande. Three fertility-related traits were analysed: Conception rate in heifers (FERG), conception rates in cows (FERV), and calving-to-first-insemination interval (IVIA). These traits were expressed as yield deviations, i.e., phenotypes adjusted for non-genetic effects estimated on the whole Holstein and Normande population for the routine genetic evaluation. For each trait, a linear model was fitted using the lm() function in R, with clusters as the fixed effect. Model assumptions were assessed through diagnostic plots (residuals vs fitted, Q-Q plots, etc.). No major violations were observed that would compromise interpretation. Results were summarized using F-statistics (Type II ANOVA), estimated effect sizes, and coefficients of determination ($R^2$). Boxplots were used to illustrate the distribution of each trait by chronotype group.

## III. Results

### 1. Circadian activity analysis results

After filtering and processing, circadian activity data were analysed for 457 Holstein and 186 Normande cows, that were confirmed to be on pasture between May and July (2020-2024). The aggregated average hourly activity profiles showed clear 24h rhythmic pattern in the two analysed activities: ingestion and rest.

Deviation profiles, computed from the population mean level at the measuring time, highlighted individual variability both in the timing and intensity of activity across the day. While some cows showed high ingestion activity peaks in early mornings, others peaked in the afternoon or in the evening, illustrating the presence of circadian variability and individuality within the population, as illustrated in figure (04).

Ingestion patterns (a, c) show clear rhythms with individual variability in peak timing and amplitude. Holstein cows (a) tend to show higher peaks, in early morning and evening, peaks around thrice a day, with more pronounced variability between individuals. Normande cows (c) also display rhythmic ingestion patterns, though slightly more synchronized midday peaks.

Rest patterns (b, d), are more synchronized across individuals in both breeds, with most cows resting predominantly at night (00:00 – 05:00) and showing reduced rest during daylight. However, Normande cows (d) tend to have a less varied rest activity compared to Holsteins (b), although one cow in the Holstein group seem to show unusual high deviations that may reflect a measurement noise that wasn't filtered.

These profiles highlight a 24 hours rhythmicity in both breeds and reveal a variability of activity between individuals. This reinforces the idea of individual chronotypes, motivating the clustering step in the next part.

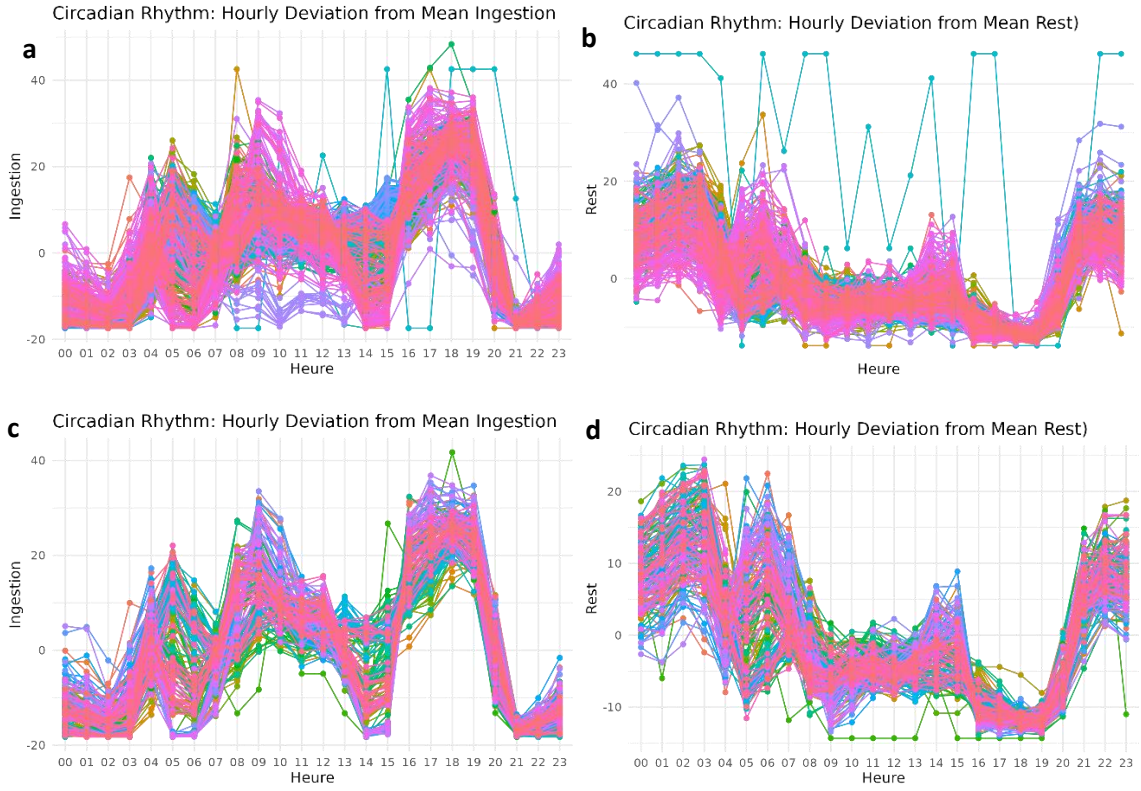


**Figure 04: Circadian hourly deviations for ingestion and rest profiles across the 24h cycle in both Holstein (a and b) and Normande (c and d) cows**, as in x axis: time and y axis: deviation of the activity. Each line represents one individual cow's deviation from the population mean across 24 hours.

## 2. Clustering results

Before going to the actual clustering of cows, we first ensured the robustness and accuracy of the clustering model, by applying an 80/20 train-test split strategy. 80% of the cows were randomly selected and used to train the Magma clustering model, while the remaining 20% validation population was used to test the model's predictive ability. We then calculated correlation between predicted and true values in the validation populations. Correlation was equal to 0.953, and the plot showed model's robustness.

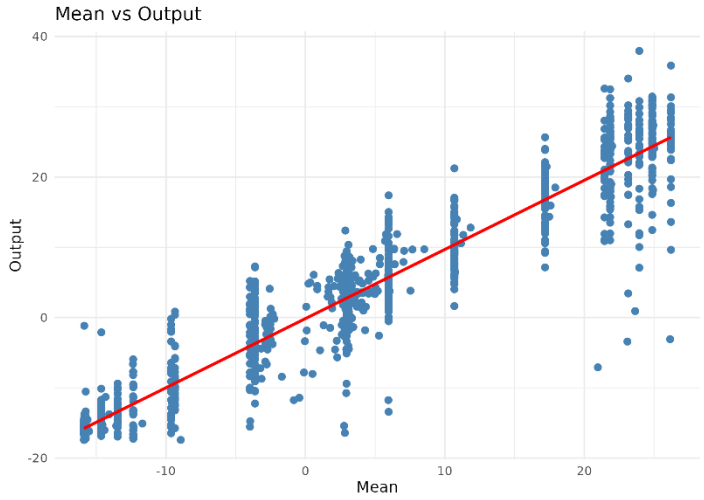


**Figure 05: Scatterplot between predicted values (Mean) and true values (Output).**

After performing clustering on both ingestion and rest profiles separately using the MagmaClustR package, cows were grouped based on their temporal activity pattern. Hourly time was used as the input variable, and deviations from the population mean served as the output. Unlike the default setting of three clusters, we manually specified two clusters to better align with biological interpretability in the GWAS analyses. Cows were allocated to their most probable cluster, each cluster representing a chronotype as showed in figures (06, 07). In Holstein cows, clustering based on ingestion deviation profiles resulted in two clusters where 168 and 288 individuals allocated to clusters K1 and K2, respectively. When clustering was based on rest activity, 304 Holsteins were allocated to K1 and 152 to K2, with K1 and K2 being arbitrarily named. For Normande cows, 106 and 109 were assigned to K1 and K2, for ingestion, while 111 and 74 were clustered in K1, K2, respectively, for rest.

In Holstein (figure 06), for ingestion panel (a), cows in cluster K1 show sharper peaks in the early morning and in the evening as they also present an earlier ingestion activity than those on the second cluster. In contrast, cows in K2 exhibit a more gradual activity as the ingestion deviation adds up throughout the day. For rest panel (b), K1 cows show clear consistent peaks during the night, while K2 cows display more variable rest patterns, particularly between morning and midday afternoon. K1 and K2 for ingestion and K2 and K1 for rest represent respectively complementary activity profiles, as expected.

In Normande, in the ingestion panel (a) of the figure (07), K1 shows clear three sharp peaks a day, early morning (04:00–05:00), morning to midday (07:00 - 12:00), and evening (18:00-19:00). Cows in cluster K2 displays a flatter and evenly distributed ingestion pattern across the day. In rest panel (b), cows in cluster K2 (blue dashed line) rest mainly during the night as the illustration shows clear rest activity between 20:00 and 05:00, with minimal rest during the day, while cluster K1 cows show a less structured pattern, and they rest almost during all day, with no clear peak time for resting.

Circadian rhythm profiles differ noticeably between Holstein and Normande cows, with Normande cows exhibiting a more pronounced alternation between diurnal activity and nocturnal rest.

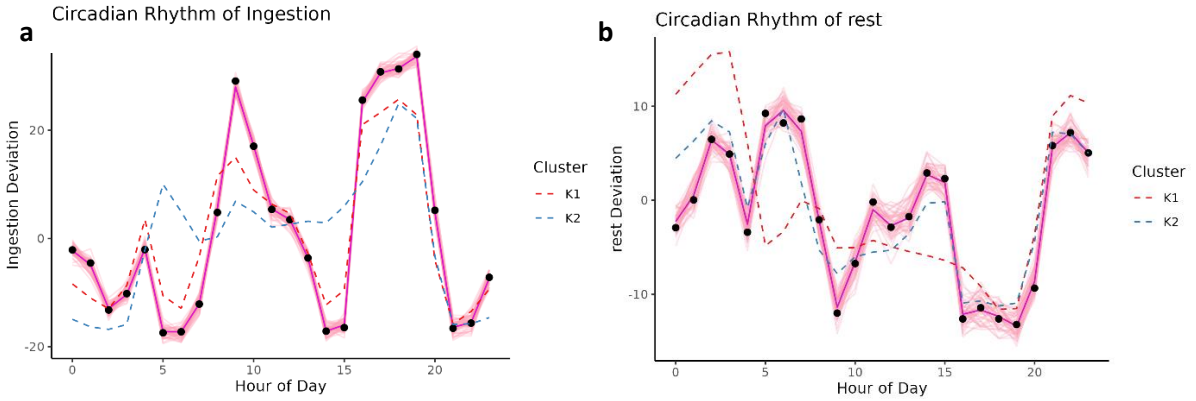


**Figure 06: Circadian rhythm profiles of ingestion (a) and rest (b) in Holstein cows identified by the MagmaClustR algorithm.** Two clusters, K1 (red dashed line) and K2 (blue dashed line), represent each a distinct group-level mean learned from the training data. The solid pink line represents the model's predicted mean deviation profile for the test cow, while the pink shaded band around it shows the 95% credible interval around this prediction, reflecting model's uncertainty. The black dots represent the actual observed deviation values for the test cow at different hours. x axis in the hour of day and y axis is the deviation of the activity in minutes.

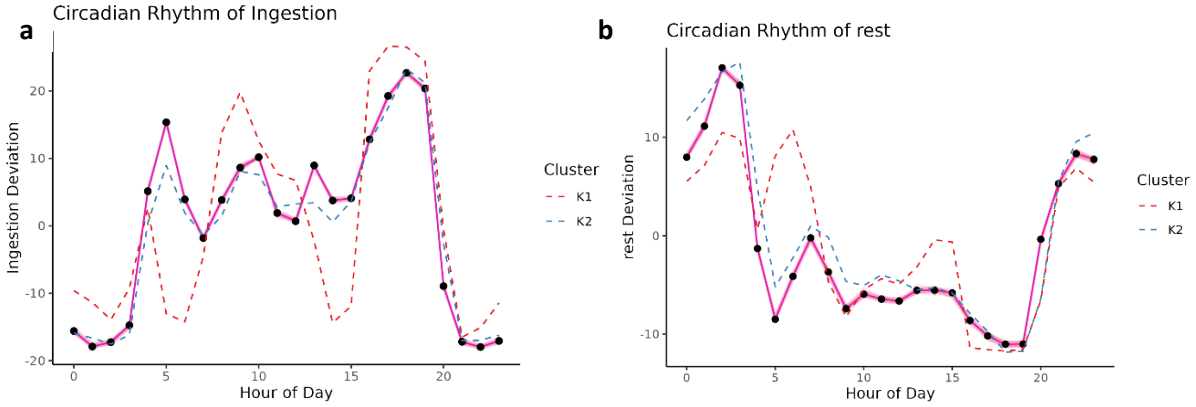


**Figure 07: Circadian rhythm profiles by ingestion/rest profiles in Normande cows**, x axis in the hour of day and y axis is the deviation of the activity in minutes.

To increase our chronotype definition reliability, we did an intersection between clusters of ingestion and rest. We retained only cows who belong to the shared subset, as those outside the overlap were excluded to reduce phenotypic noise, as illustrated in Figure (08 and 09).

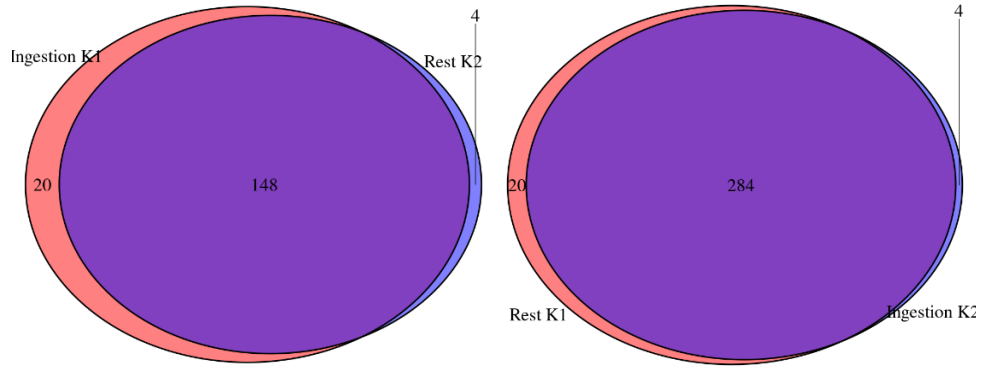


**Figure 08: Number of overlapping Holstein cows** a Cows in cluster 1 of ingestion and cluster 2 of rest b cows in cluster 2 of ingestion and cluster 1 of rest.

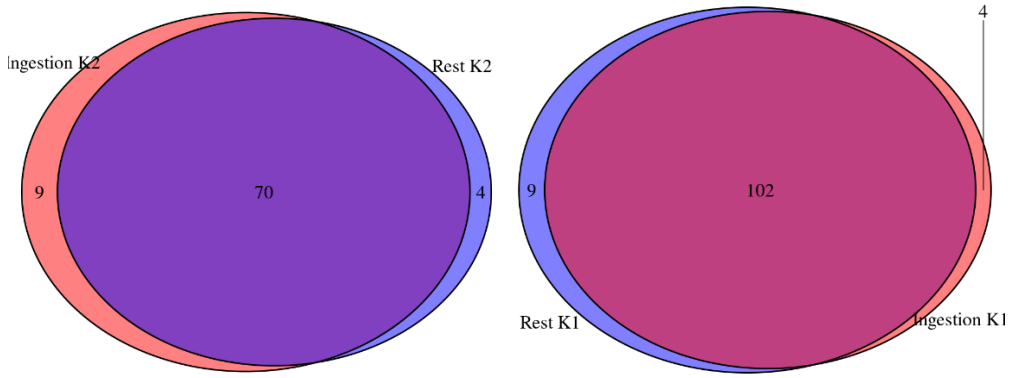


**Figure 09: Number of overlapping Normande cows** a Cows in cluster 1 of ingestion and cluster 1 of rest b cows in cluster 2 of ingestion and cluster 2 of rest.

### 3. GWAS results

Genome wide association studies were conducted on cows presenting an overlap between ingestion and rest clustering. Therefore, GWAS were applied on 432 Holstein and 172 Normande cows to identify QTL associated with the chronotype (0 or 1) using imputed whole genome sequence data.

Genetic and residual variances were estimated using the genomic relationship matrices from 50K genotypes. Heritability estimates were very different in Holstein and Normande cows; 0.61 and 0.08, respectively, as shown in Table 02.

Table 02: Variance and heritability values estimated from the genomic relationship matrix.

| Breed | Genetic variance | Residual variance | $h^2$ |
|---|---|---|---|
| Holstein | 0.14 | 0.09 | 0.61 |
| Normande | 0.02 | 0.23 | 0.08 |

Using a genome-wide p-value significance threshold of 1*10e-06, a total of 182 significant variants (SNPs or Indels) was detected: 65 variants located on six chromosomes in Holstein and 117 variants located on four chromosomes in Normande. In Holstein cows, 8 regions were identified on chromosome 1, 3, 6, 7, 11, and 21 (Figure 10, table 03), the strongest association being located on chromosome 6 at 60,232,663 bp. In Normande cows, 4 regions were identified on chromosome 1, 12, 16 and 23 (Figure 11, table 03), with strongest associations located on chromosome 16 at 56,365,799 bp and on chromosome 23 at 14,433,101 bp.

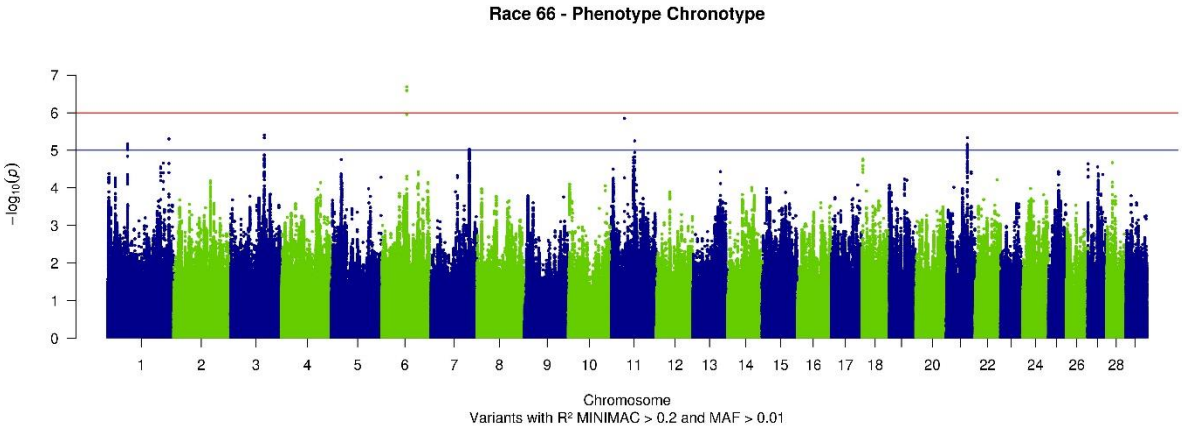


**Figure 10: Manhattan plot for GWAS results of Holstein cows on chronotype 0/1 trait on sequence level density.**

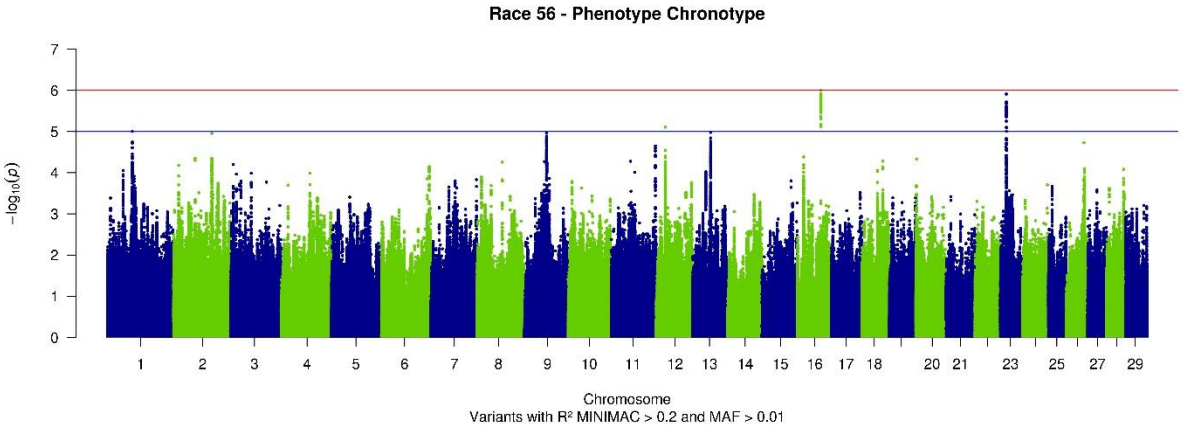


**Figure 11: Manhattan plot for GWAS results of Normande cows on chronotype 0/1 trait on sequence level density.**

Functional annotation:

After using Ensembl VEP v114 tool, functional annotations revealed that 57% of variants among the 65 variants detected on the six chromosomes of the Holstein breed were located in intronic regions, 34% were in intergenic regions, 9% were in upstream regions and 1% left were in regulatory regions (figure 12). In the Normande breed, the majority among the 117 detected variants on the four chromosomes were in intronic regions (79%), 9% were intergenic, 7% were synonymous variants, and 4% were classified missense (figure 13).

Table 03: Description of QTL regions identified in GWAS analysis of Holstein and Normande cows

| Breed | Chromosome | QTL | Region (bp) start-end | Number of variants with significant effects | Location of the most significant variant | MAF | Effect size | Standard Error | P-value | Annotation | Gene |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Holstein** | 1 | QTL1.1 | 147299552 - 147299558 | 2 | 147299558 | 0.181 | -0.307 | 0.067 | 4.92e-06 | Intergenic region | RUNX1 |
| **Holstein** | 1 | QTL1.2 | 48621526 - 48633585 | 35 | 48623408 | 0.155 | -0.560 | 0.126 | 6.98e-06 | Intergenic region | U7 ALCAM |
| **Holstein** | 3 | 1 QTL | 80815984 - 80816007 | 3 | 80816007 | 0.236 | -0.379 | 0.082 | 3.90e-06 | Intron variant | CACHD1 |
| **Holstein** | 6 | 1 QTL | 60211257 - 60232663 | 7 | 60232663 | 0.012 | -3.207 | 0.608 | 1.10e-06 | Intron variant | LIMCH1 |
| **Holstein** | 7 | 1 QTL | 92265331 - 92267970 | 2 | 92267970 | 0.230 | -0.380 | 0.085 | 9.32e-06 | Intergenic region | 7SK ENSBTAG00000048629 |
| **Holstein** | 11 | QTL11.1 | 31357791 | 1 | 31357791 | 0.074 | -0.396 | 0.082 | 1.39e-06 | Intron variant | FSHR |
| **Holstein** | 11 | QTL11.2 | 53646427 | 1 | 53646427 | 0.108 | -0.458 | 0.101 | 5.59e-06 | Intron variant | CTNNA2 |
| **Holstein** | 21 | 1 QTL | 55646427 - 51415799 | 14 | 51415799 | 0.051 | -0.724 | 0.158 | 4.62e-06 | Intron variant | LRFN5 |
| **Normande** | 1 | 1 QTL | 59889548 | 1 | 59889548 | 0.728 | -0.565 | 0.127 | 9.90e-06 | Intergenic region | ZBTB20 ENSBTAG00000047946 |
| **Normande** | 12 | 1 QTL | 22102121 | 1 | 22102121 | 0.371 | -0.345 | 0.077 | 7.83e-06 | Non-coding transcript | FOXO1 ENSBTAG00000053790 |
| **Normande** | 16 | 1 QTL | 56360968 - 56372052 | 68 | 56365799 | 0.312 | -0.433 | 0.089 | 1.22e-06 | Intron variant | TNN |
| **Normande** | 23 | 1 QTL | 14381537 - 14444918 | 47 | 14433101 | 0.311 | -0.586 | 0.120 | 1.22e-06 | Intron variant | LRFN2 |

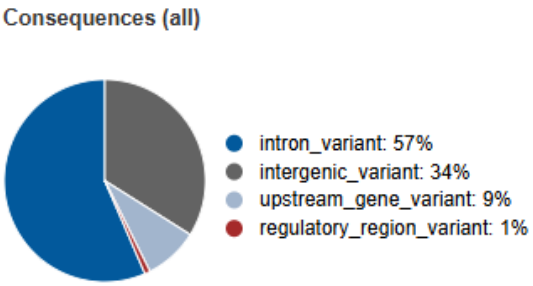


**Figure 12:** Variant repartition based on their location in Holstein breed chromosomes (Ensembl)

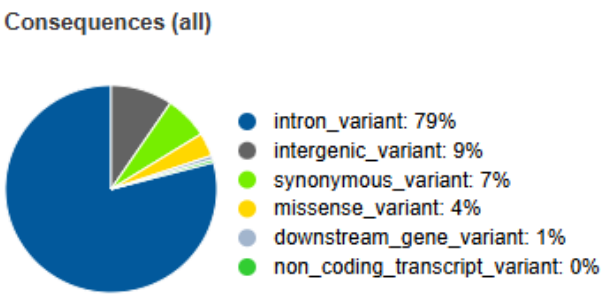


**Figure 13:** Variant repartition based on their location in Normande breed chromosomes (Ensembl)

From these results, 16 genes were identified as potentially impacted by the most significant variants. Notably, several variants were located within or near candidate genes as ALCAM (Activated Leucocyte Cell Adhesion Molecule), FSHR (Follicle Stimulating Hormone Receptor), CTNNA2 (Catenin Alpha 2), TNN (Tenascin N). A summary of QTL regions, lead variants, MAF, effects sizes, annotations and candidate genes, is provided in Table 03.

### 4. Performance indicators results

To evaluate the association between chronotype and reproductive performance. We applied linear models on both breeds. In the Holstein breed, the linear models revealed no statistically significant association between the chronotype and any of the three fertility traits analysed. However, IVIA showed a weak non-significant trend with cows in K2 exhibiting a longer interval (+5,8 days; p = 0.10). A summary of the results is presented in table 04.

For the Normande heifers, a statistically significant association between the chronotype and the heifer conception rate FERG (p-value =7.109e-05). heifers in cluster K2 had an average reduction of 0.245 units in conception rate compared to heifers in K1. This indicates a meaningful difference in fertility between chronotypes. For FERV, there was a trend toward lower performance in K2 cows, though it did not reach statistical significance. For IVIA, cows in K2 tend to have a shorter interval from calving to first insemination (-12,1 days) though it's not significant. A summary of the results is presented in table 05.

Boxplots presenting the distribution of each fertility trait by chronotype group in the Holstein cows in figure (14). For FERG and FERV, no clear differences was observed between chronotype clusters K1 and K2. For IVIA, a slight upward shift in the distribution for K2 was observed, consistent with the linear model trend (+5.81 days), but this difference wasn't statistically significant. High variability and several extreme values were present in all traits, especially in IVIA. Whears , Normande cows, showed signficant heifer conception rate as animals in K2 group showed a significantly lower conception rate compared to K1 ($p < 0.001$). Conception rate in cows and calving to first insemination was not statistically significant, but a tendency toward lower fertility was observed in K2 cows although this same cluster had shorter calving to first insemination interval.

**Table 04: A summary of linear models results for the Holstein breed.** The table reports the estimated effect of chronotype cluster K2 compared to K1, the p-value assessing the statistical significance of this effect, the coefficient of determination ($R^2$) indicating the proportion of the variance explained by the model. While none of the associations were statistically significant ($p > 0.05$), cows in cluster K2 show slightly a low conception rate (FERG and FERV) and a longer calving to insemination interval (IVIA), suggesting a possible trend of delayed reproduction recovery postpartum in this group.

| Trait | Effect (K2 vs K1) | p-value | $R^2$ |
|---|---|---|---|
| **FERG (heifer conception rate)** | -0.03 | 0.44 | 0.0015 |
| **FERV (cow conception rate)** | -0.015 | 0.72 | 0.0005 |
| **IVIA (calving to AI interval)** | +5.81 days | 0.1 | 0.009 |

**Table 05: A summary of linear models results for the Normande breed.** The table reports the estimated effect of chronotype cluster K2 compared to K1, the p-value assessing the statistical significance of this effect, the coefficient of determination ($R^2$) indicating the proportion of the variance in the trait explained by the model. A significant difference was observed in FERG (p= 00007), with cows in K2 showing substantially lower heifer conception rates compared to K1. For FERV and IVIA, p-values were near the 0.05 threshold, supporting potential trends where K2 have lower conception rates and return to oestrus and insemination quickly than K1 cows.

| Trait | Effect (K2 vs K1) | p-value | $R^2$ |
|---|---|---|---|
| **FERG (heifer conception rate)** | -0.245 | **0.00007*** | 0.095 |
| **FERV (cow conception rate)** | -0.189 | 0.076 | 0.033 |
| **IVIA (calving to AI interval)** | -12,1 days | 0.085 | 0.028 |

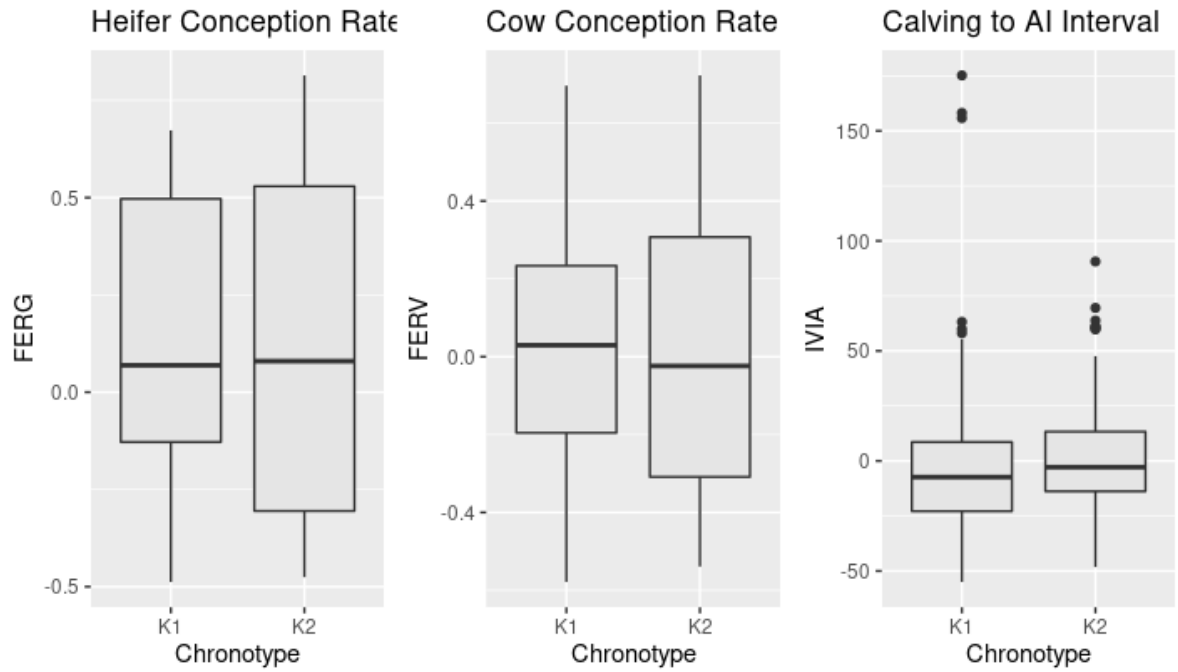


**Figure 14: Distribution of the three fertility traits by chronotype group (K1 vs K2) in the Holstein breed.** From left to right: heifer conception rate (FERG), cow conception rate (FERV), and calving-to-first-insemination interval (IVIA). Although, no statistic significant differences were found between the groups, a trend suggest that cows in chronotype K1 may have slightly higher conception rate (FERV), and shorter insemination calving to insemination interval (IVIA).

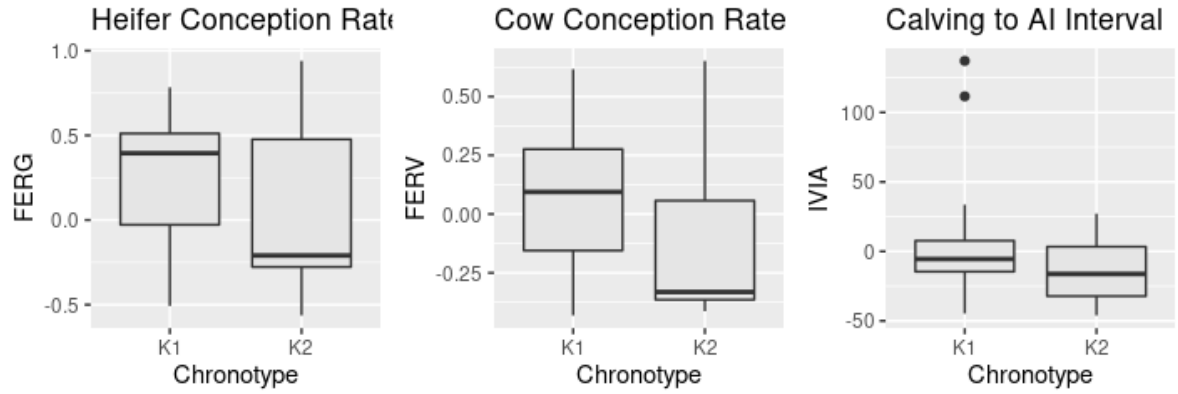


**Figure 15: Distribution of the three fertility traits by chronotype group (K1 vs K2) in the Normande breed.** From left to right: FERG (heifer conception rate): animals in K2 group showed a significantly lower conception rate compared to K1 ($p < 0.001$), FERV (cow conception rate): a tendency toward lower fertility was observed in K2 cows, although the difference is not statistically significant (0.076). IVIA (calving-to-first-insemination interval): K2 had shorter calving to first insemination interval, but this trend did not reach significance ($p = 0.085$).

## IV. Discussion

We aimed to study the genetic determinism of circadian activity, i.e., chronotypes, which is, to our knowledge, a subject that has never been studied before in dairy cows. By analysing their behavioral rhythms during a 24-hour cycle, we eventually succeeded to identify two different chronotypes in Holstein and Normande cows based on ingestion activity. A GWAS was then conducted using these chronotypes as traits, revealing 182 significant genetic variants across both breeds. These variants are associated with several promising candidate genes (*ALCAM*, *CACHD1*, *LIMCH1*, *FSHR*, *LRNF5*, *ZBTB20*, *FOXO1*, *TNN*, *LRFN2*), (See table x for positions and annotations).

Several types of activity data were available, including rest, rumination, standing up, overactivity, and ingestion. While each of these activities reflects certain aspects of a cow's daily routine, ingestion was selected as a proxy for circadian activity rhythm, given that it involves coordinated muscle activity, alertness, and voluntary movement. Other activities, such as rumination, tend to be more passive, intermingled and therefore less suitable for defining individual chronotypes. We reinforced this approach by incorporating rest activity, as the combination of ingestion and rest captures key aspects of the cow's daily rhythm and provides a more robust assessment of their underlying circadian phenotype.

One limitation of this study is the relatively small size of cows (172 Normande and 432 Holstein), particularly for genetic analyses, which may reduce statistical power and increase the risk of false positives. However, the use of well detailed activity data helps to strengthen the biological relevance of the findings despite the limited size. Despite this limitation, we were able to identify chronotypes and to perform GWAS with significant hits and QTLs detection.

Through heritability estimations, we found that the studied trait (chronotypes) was highly heritable in Holstein cows ($h^2$=0.61) but lowly in Normande cows ($h^2$=0.08). This result shows a remarkable difference that can be explained partly by the population size, as we had access to a limited number of Normande cows compared to Holstein cows (172 Normande cows vs. 432 Holstein cows). Since both breeds were raised in the same environmental conditions, macro-environmental factors are unlikely to explain this discrepancy. However, individual cows may still experience unique micro-environmental influences that cannot be entirely ruled out. Additionally, we observed that Normande cows displayed a clear circadian rhythm, characterized by diurnal feeding and nocturnal resting behavior—a pattern that was much less pronounced in Holstein cows. This less pronounced phenotype in Normande cows may also explain the low heritability observed.

The genes we found associated to the chronotypes, are likely related to different physiological, neural, metabolic, and pathological processes. As previously mentioned, nearly all physiological processes are regulated by the circadian clock, from hormone release to metabolism and feeding behavior.

A notable observation is the identification of genes from the same family, *LRFN*, in both breeds. In the Holstein breed, a variant was found on chromosome 21 at 51,415,799 bp within *LRFN5* (Leucine Rich Repeat And Fibronectin Type III Domain Containing 5). This gene plays a role in neural connectivity and synapse formation, as well as immune functions, including the restriction of T cell responses and neuroinflammation, as it decreased expression has been reported to promote neuroinflammation (Zhu et al., 2016). This variation likely affects the

timing of activities in dairy cows by modulating neural pathways involved in activity preference and responsiveness to environmental stimuli. In the Normande breed, *LRFN2*, an important paralog of *LRFN5*, was identified as a candidate gene associated with an intronic variant located on chromosome 23 at position 14,433,101 bp. This gene is predicted to be involved in the modulation of chemical synaptic transmission and regulation of post synapse organization as well as neural network formation. Its function overlaps with *LRFN5*, as this family gene contributes to the development and function of excitatory synapses in the brain and behavior. The identification of *LRFN2* in Normande cows, alongside *LRFN5* in Holsteins, suggests that these genes' different family members may influence behavioral timing through shared or complementary neuroregulatory pathways.

In Holstein breed, a variant at 48,623,408 bp on chromosome 1 was identified near *ALCAM*, a candidate gene known as CD166 antigen, a transmembrane receptor which has been thoroughly studied in human cancer research, has been implicated in leukocyte adhesion/migration and T cell activation (Hassan et al., 2004; Zimmerman et al., 2006; Cayrol et al., 2007). In addition, *ALCAM* was shown to be over-expressed in the milk somatic cells of a mastitis-resistant line of sheep (Bonnefont et al., 2011), and seemed as a plausible candidate gene to be involved in mastitis resistance in Holstein-Friesian cattle (Meredith et al 2013).

A significant variant was found in the Holstein breed on chromosome 3 at 80,816,007 bp within the *CACHD1* (Cache Domain Containing 1), a gene that has a role in regulating calcium channel activity, crucial for various physiological processes, including neural excitability, synaptic transmission, and muscle contraction (Dahimen S et al., 2018; Stephens, G. J., & Cottrell, G. S. 2019). Although the exact function of *CACHD1* in cattle remains to be fully defined, its involvement in calcium signalling makes it biologically, a plausible candidate gene for traits related to circadian regulation and feeding behavior. Calcium signalling plays a central role in the molecular clock, where its intracellular levels help regulate the expression of core clock genes such as *BMAL1*, *PER* and *CRY*. Therefore, fluctuations with these intracellular levels influence the transcriptional activity of these genes (Said et al., 2020; Noguchi et al., 2017; Ikeda et al.,2003; Cavieres-Lepe & Ewer, 2021). Hypocalcaemia -both clinical and subclinical- is a common metabolic disorder in high producing dairy cows, like Holstein in our study, particularly during the transition period. It can disrupt muscle activity leading to weakness, collapse, and altered feeding patterns, while also affecting overall metabolic balance. These disruptions may potentially interfere with regular circadian rhythms. Given that *CACHD1* regulates calcium channels function, genetic variation within this gene may affect how circadian and behavioral rhythms respond to calcium fluctuations. As chronotypes are based on activity rhythms, variation in *CACHD1* may influence behavior through calcium-dependent pathways, potentially affecting the timing of activity. Thus, this gene may represent a molecular link between mineral metabolism, circadian rhythm, and overall health of cows.

In the same breed, another significant variant was found at position 60,232,663 bp on chromosome 6 within *LIMCH1* (LIM and Calponin Homology Domains 1), a protein coding gene involved in cytoplasmic actin-based contraction. *LIMCH1* is associated with cell motility, myosin II binding, positive regulation of stress fiber assembly, and focal adhesion dynamics (Alliance of Genome Resources, 2025). These functions are central to regulating cytoskeletal tension and mechanical stress, especially during periods of activity. Given that the cow chronotypes in this study are defined by ingestion rhythms, these movements create localized mechanical stress in skeletal muscle and epithelial tissues (tongue, jaw, neck...) that may trigger intracellular signalling cascades affecting circadian gene expression contributing to the regulation of activity rhythms and peripheral clock entrainment. This is supported by findings

in C2C12 myoblasts where mechanical stress load reduces *Per* and *Cry* and enhances *Clock*/*Bmal1* gene expression (Wang M et al., 2021), suggesting a mechanotransduction-clock interaction that may underline chronotype variation in cows. Furthermore, studies have shown that *LIMCH1* knockout in mice leads to muscle weakness, highlighting its functional importance in muscle physiology (Penna et al., 2023). Additionally, *LIMCH1* has been implicated in cattle body size variation, particularly forehead size in Brahman cattle (Chen Q et al., 2020). While this association is primarily morphological, it's involvement in muscle regulative pathways suggests that selection for structural traits could have indirect effects on circadian rhythms and behavioral traits such as feeding.

Another interesting variant was found in the Holstein breed on chromosome 11 at 31,357,791 bp within the *FSHR* gene that encodes the follicle stimulating hormone receptor. This gene directly affects the female reproduction cycle as it is highly expressed in ovarian tissue and mediates follicular development and estrogen production. *FSHR* is considered a candidate gene for traits such as multiple birth and twinning in Holstein, as it encodes receptors for three essential hormones for female reproduction: luteinizing hormone (LH), choriogonadotropin, and follicle stimulating hormone (FSH) (Widmer et al., 2021; Lett & Kirkpatrick, 2022). Hormonal cycles, including FSH, are regulated by the circadian clock. The hypothalamic-pituitary-gonadal axis, which controls *FSHR* activity, is tightly integrated with the circadian system and closely linked to metabolic signals (Sellix, 2013). Therefore, variation near *FSHR* could be related to the timing of activity rhythms. In our study, we explored the association between these chronotypes and reproductive performances in the Holstein cows analysed. Regarding the calving to first insemination interval, a key reproductive performance indicator, cows in cluster K2 had an average interval 5.8 days longer than those in cluster K1, although this difference was not statistically significant (p-value = 0.10). There seems to be a plausible influence of circadian rhythm on postpartum reproductive recovery, as cows in cluster K1 may return to oestrus and insemination sooner.

In contrast, although FSHR was not identified through Normande GWAS, an interesting pattern emerged. Normande cows in cluster K2 also exhibited lower conception rates (FERG and FERV) like the Holstein cows, with a particular strong and significant decrease observed for conception rate in heifers. However, these cows showed tendency toward a shorter IVIA ($p = 0.085$), indicating they were inseminated earlier after calving. This apparent contradiction – earlier insemination despite lower fertility – may reflect a mismatch between behavioral oestrus and actual physiological readiness. It raises the possibility that K2 cows may resume to heat quickly but not under optimal endocrine conditions for conception, how this remains a small speculation, or would also put the management of reproduction that may be different between the two breeds. This contrasting findings between the two breeds may reflect differences in the genetic background, physiological sensitivity to circadian rhythm misalignment, or interactions with management conditions. These observations suggest that the relationship between chronotype and fertility may be breed-dependant warranting further investigation whether these rhythms reflect different adaptive strategies or vulnerabilities to circadian misalignment.

The *CTNNA2* gene associated to variant found in the Holstein breed on chromosome 11 at 53,646,427 bp, is involved in the development of the nervous system and behavior regulation. It has previously been associated with the regulation of L-alanine in bovine blood (Li et al., 2020) and identified as a candidate gene under positive selection for tolerance to trypanosomiasis in Boran and N'dama cattle (Kim SJ et al., 2017; Noyes H et al., 2019), as well as for climate adaptation in Mediterranean cattle (Flori L et al., 2019). L-alanine is a known metabolic marker that reflects energy balance, especially in dairy cows, and high levels may

indicate extended intervals of feeding, as observed in cluster 1 of both Holstein and Normande cows, which showed distinct peaks of ingestion. However, its direct connection to a circadian chronotype remains speculative in this context. Nevertheless, the identification of *CTNNA2* may point underlying neurobehavioral mechanisms contributing to differences in activity. Further research is needed to investigate its potential role in circadian rhythm regulation and its broader impact on cow's health.

In the Normande breed, a significant variant located on chromosome 1 at 59,889,548 bp near *ZBTB20* (Zinc Finger and BTB Domain Containing 20), a candidate gene that encodes a transcription factor involved in brain development and metabolic regulation. In mice, it was shown that loss of *ZBTB20* impairs circadian output and leads to unimodal behavioral rhythms (Qu et al., 2016). This suggest a potential role in modulating behavioral timing through circadian regulatory pathways.

As well in Normande cows, we identified a variant located on chromosome 16 at 56,365,799 bp within *TNN* (Tenascin N). This candidate gene is involved in cell adhesion, migration, neural development, and integrin binding. It plays a role in neurite outgrowth and behavioral regulation through its effect on cell migration and tissue remodelling. Given its role in neuro-muscular signalling, *TNN* may influence timing and coordination of circadian activity through pathways related to neuronal connectivity and motor control. Interestingly, *TNN* has also been linked to mammary tumor progression by promoting the migratory behavior of breast cancer cells (Degen et al., 2007). Moreover, circadian rhythm disruptions have been shown to accelerate mammary tumor progression (Hadadi & Acloque, 2021). Further research into *TNN* in cattle is needed, particularly regarding its potential role in circadian regulation and mammary gland health.

Within the same breed, we have also identified a variant with significant effects on chronotypes located within a non-coding region of *FOXO1*. While the direct functional impact of this SNP is still unknown in cattle, *FOXO1* encodes a transcription factor with key roles in metabolism, circadian rhythm regulation, and mammary gland function. It is involved in tumor suppression by inducing apoptosis (Duffy et al., 2020; Zhang et al., 2025) and has been associated with body weight and growth in chicken ((Xie et al., 2012; Abdalhag et al., 2015; Yuan et al., 2017; Wang et al., 2022). These findings suggest a potential relevance of *FOXO1* to dairy cow health and performance, possibly influenced by circadian rhythms.

Other QTLs we identified in the Holstein breed, included a variant near *RUNX1*, a transcription factor involved in haematopoiesis and immune function, as well as variants near non-coding RNAs U7 and 7SK, which are known for their roles in histone RNA processing and transcriptional regulation, respectively. There is no research establishing their functions in circadian regulation, therefore their identification needs further investigation.

It is worth to note that *ALCAM* and *ZBTB20* have been both mentioned in a genome-wide association study (GWAS) to identify genetic loci associated with somatic cell score (SCS), an indicator trait of mammary gland inflammation in Holstein-Friesian cattle (Meredith et al 2013). It is tempting to hypothesize that chronotype clusters (e.g., K1 vs K2) may differ in their susceptibility to mastitis or inflammatory responses. Although not tested in the present study, such associations could be explored in future work just like the performance indicators tests we applied, using SCS records to assess potential links between behavioral rhythms and udder health.

## V. Conclusion

This study successfully allowed to define different chronotypes in dairy cows and through comprehensive genome-wide association studies analysis, several genomic regions associated with those chronotypes were identified. The findings revealed promising candidate genes involved with neural regulation, calcium signalling, energy metabolism and circadian control. These results support the hypothesis of the existence of genetic variation in the daily rhythm of cows, potentially linked to broader physiological processes such as milk production and fertility.

Despite the small size of cows, the use of high-resolution data strengthens substantial biological relevance. All in all, it remains an exploratory work that is paving the way for other future research, opening promising perspectives for understanding the connections between circadian rhythms, cow's performance, and health. Ultimately, this work offers a potential foundation for future breeding strategies that consider both circadian rhythm variation and physiological traits to improve productivity and welfare.

**References:**


Abdalhag, M., Zhang, T., Fan, Q., Zhang, X., Zhang, G., Wang, J., Wei, Y., & Wang, Y. (2015). Single nucleotide polymorphisms associated with growth traits in Jinghai yellow chickens. *Genetics and Molecular Research*, *14*(4), 16169–16177. https://doi.org/10.4238/2015.december.8.6

*Alliance of Genome Resources*. (2025, May). Gene Function and Annotation Overview. https://www.alliancegenome.org/gene-function

Boichard, D., Guillaume, F., Baur, A., Croiseau, P., Rossignol, M. N., Boscher, M. Y., Druet, T., Genestout, L., Colleau, J. J., Journaux, L., Ducrocq, V., & Fritz, S. (2012). Genomic selection in French dairy cattle. *Animal Production Science*, *52*(3), 115. https://doi.org/10.1071/an11119

Bonnefont, C. M. D., Toufeer, M., Caubet, C., Foulon, E., Tasca, C., Aurel, M., Bergonier, D., Boullier, S., Robert-Granié, C., Foucras, G., & Rupp, R. (2011). Transcriptomic analysis of milk somatic cells in mastitis resistant and susceptible sheep upon challenge with Staphylococcus epidermidis and Staphylococcus aureus. *BMC Genomics*, *12*(1). https://doi.org/10.1186/1471-2164-12-208

Bosler, O., & Beaudet, A. (1985). VIP neurons as prime synaptic targets for serotonin afferents in rat suprachiasmatic nucleus: a combined radioautographic and immunocytochemical study. *Journal of Neurocytology*, *14*(5), 749–763. https://doi.org/10.1007/bf01170826

Casey, T. M., & Plaut, K. (2022). Circadian clocks and their integration with metabolic and reproductive systems: our current understanding and its application to the management of dairy cows. *Journal of Animal Science*, *100*(10). https://doi.org/10.1093/jas/skac233

Cavieres-Lepe, J., & Ewer, J. (2021). Reciprocal relationship between calcium signaling and circadian clocks: implications for calcium homeostasis, clock function, and therapeutics. *Frontiers in Molecular Neuroscience*, *14*. https://doi.org/10.3389/fnmol.2021.666673

Cayrol, R., Wosik, K., Berard, J. L., Dodelet-Devillers, A., Ifergan, I., Kebir, H., Haqqani, A. S., Kreymborg, K., Krug, S., Moumdjian, R., Bouthillier, A., Becher, B., Arbour, N., David, S., Stanimirovic, D., & Prat, A. (2007). Activated leukocyte cell adhesion molecule promotes leukocyte trafficking into the central nervous system. *Nature Immunology*, *9*(2), 137–145. https://doi.org/10.1038/ni1551

Chauhan, S., Norbury, R., Faßbender, K. C., Ettinger, U., & Kumari, V. (2023). Beyond sleep: A multidimensional model of chronotype. *Neuroscience & Biobehavioral Reviews*, *148*, 105114. https://doi.org/10.1016/j.neubiorev.2023.105114

Chen, Q., Huang, B., Zhan, J., Wang, J., Qu, K., Zhang, F., Shen, J., Jia, P., Ning, Q., Zhang, J., Chen, N., Chen, H., & Lei, C. (2020). Whole-genome analyses identify loci and selective signals associated with body size in cattle. *Journal of Animal Science*, *98*(3). https://doi.org/10.1093/jas/skaa068

Cheng, H., Liu, Z., Wu, G., Ho, C., Li, D., & Xie, Z. (2021). Dietary compounds regulating the mammal peripheral circadian rhythms and modulating metabolic outcomes. *Journal of Functional Foods*, *78*, 104370. https://doi.org/10.1016/j.jff.2021.104370

Cipolla-Neto, J., Amaral, F. G., Afeche, S. C., Tan, D. X., & Reiter, R. J. (2014). Melatonin, energy metabolism, and obesity: a review. *Journal of Pineal Research*, *56*(4), 371–381. https://doi.org/10.1111/jpi.12137

Coomans, C. P., Lucassen, E. A., Kooijman, S., Fifel, K., Deboer, T., Rensen, P. C. N., Michel, S., & Meijer, J. H. (2015). Plasticity of circadian clocks and consequences for metabolism. *Diabetes Obesity and Metabolism*, *17*(S1), 65–75. https://doi.org/10.1111/dom.12513

Daetwyler, H. D., Capitan, A., Pausch, H., Stothard, P., Van Binsbergen, R., Brøndum, R. F., Liao, X., Djari, A., Rodriguez, S. C., Grohs, C., Esquerré, D., Bouchez, O., Rossignol, M., Klopp, C., Rocha, D., Fritz, S., Eggen, A., Bowman, P. J., Coote, D., . . . Hayes, B. J. (2014). Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nature Genetics*, *46*(8), 858–865. https://doi.org/10.1038/ng.3034

Dahimene, S., Page, K. M., Kadurin, I., Ferron, L., Ho, D. Y., Powell, G. T., Pratt, W. S., Wilson, S. W., & Dolphin, A. C. (2018). The α2δ-like Protein Cachd1 Increases N-type Calcium Currents and Cell Surface Expression and Competes with α2δ-1. *Cell Reports*, *25*(6), 1610-1621.e5. https://doi.org/10.1016/j.celrep.2018.10.033

Dardente, H. (2007). Does a Melatonin-Dependent circadian oscillator in the pars tuberalis drive prolactin seasonal rhythmicity? *Journal of Neuroendocrinology*, *19*(8), 657–666. https://doi.org/10.1111/j.1365-2826.2007.01564.x

Degen, M., Brellier, F., Kain, R., Ruiz, C., Terracciano, L., Orend, G., & Chiquet-Ehrismann, R. (2007). Tenascin-W is a novel marker for activated tumor stroma in low-grade human breast cancer and influences cell behavior. *Cancer Research*, *67*(19), 9169–9179. https://doi.org/10.1158/0008-5472.can-07-0666

Duffy, T., Bekki, H., & Lotz, M. K. (2020). Genome-Wide occupancy profiling reveals critical roles of FOXO1 in regulating extracellular matrix and circadian rhythm genes in human chondrocytes. *Arthritis & Rheumatology*, *72*(9), 1514–1523. https://doi.org/10.1002/art.41284

Dyer, S. C., Austine-Orimoloye, O., Azov, A. G., Barba, M., Barnes, I., Barrera-Enriquez, V. P., Becker, A., Bennett, R., Beracochea, M., Berry, A., Bhai, J., Bhurji, S. K., Boddu, S., Lins, P. R. B., Brooks, L., Ramaraju, S. B., Campbell, L. I., Martinez, M. C., Charkhchi, M., . . . Yates, A. D. (2024). Ensembl 2025. *Nucleic Acids Research*, *53*(D1), D948–D957. https://doi.org/10.1093/nar/gkae1071

Ebling, F. J. P., & Barrett, P. (2008). The regulation of seasonal changes in food intake and body weight. *Journal of Neuroendocrinology*, *20*(6), 827–833. https://doi.org/10.1111/j.1365-2826.2008.01721.x

Flori, L., Moazami-Goudarzi, K., Alary, V., Araba, A., Boujenane, I., Boushaba, N., Casabianca, F., Casu, S., Ciampolini, R., D'Acier, A. C., Coquelle, C., Delgado, J., El-Beltagi, A., Hadjipavlou, G., Jousselin, E., Landi, V., Lauvie, A., Lecomte, P., Ligda, C., . . . Gautier, M. (2018). A genomic map of climate adaptation in Mediterranean cattle breeds. *Molecular Ecology*, *28*(5), 1009–1029. https://doi.org/10.1111/mec.15004

Gekakis, N., Staknis, D., Nguyen, H. B., Davis, F. C., Wilsbacher, L. D., King, D. P., Takahashi, J. S., & Weitz, C. J. (1998). Role of the CLOCK protein in the mammalian circadian mechanism. *Science*, *280*(5369), 1564–1569. https://doi.org/10.1126/science.280.5369.1564

Hadadi, E., & Acloque, H. (2021). Effets des altérations du rythme circadien sur l'évolution du cancer du sein. *Médecine/Sciences*, *37*(5), 538–541. https://doi.org/10.1051/medsci/2021046

Hassan, N. J., Barclay, A. N., & Brown, M. H. (2004). Frontline: Optimal T cell activation requires the engagement of CD6 and CD166. *European Journal of Immunology*, *34*(4), 930–940. https://doi.org/10.1002/eji.200424856

Hastings, M., O'Neill, J. S., & Maywood, E. S. (2007). Circadian clocks: regulators of endocrine and metabolic rhythms. *Journal of Endocrinology*, *195*(2), 187–198. https://doi.org/10.1677/joe-07-0378

Hayes, B. J., & Daetwyler, H. D. (2018). 1000 Bull Genomes Project to map simple and complex genetic traits in cattle: applications and outcomes. *Annual Review of Animal Biosciences*, *7*(1), 89–102. https://doi.org/10.1146/annurev-animal-020518-115024

Hogenesch, J. B., Gu, Y., Jain, S., & Bradfield, C. A. (1998). The basic-helix–loop–helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors. *Proceedings of the National Academy of Sciences*, *95*(10), 5474–5479. https://doi.org/10.1073/pnas.95.10.5474

Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., & Abecasis, G. R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics*, *44*(8), 955–959. https://doi.org/10.1038/ng.2354

Ikeda, M., Sugiyama, T., Wallace, C. S., Gompf, H. S., Yoshioka, T., Miyawaki, A., & Allen, C. N. (2003). Circadian dynamics of cytosolic and nuclear CA2+ in single suprachiasmatic nucleus neurons. *Neuron*, *38*(2), 253–263. https://doi.org/10.1016/s0896-6273(03)00164-8

*Illumina Inc | Microarray Kits for Genotyping & Epigenetics*. (n.d.). https://www.illumina.com/products/by-type/microarray-kits.html?tags=&pageSize=20&pageNumber=1&sortBy=title

*Institut de l'Élevage. (2023). Bilan génétique des races bovines laitières en 2023. Institut de l'Élevage.* (n.d.). Retrieved June 9, 2025, from https://idele.fr/?eID=cmis_download&oID=workspace%3A%2F%2FSpacesStore%2F5173106d-ab95-464e-8ce4-073ec3c92720&cHash=35c67b92156dc7c82a91d5b5d235276a

Jacobs, B. L., Martín-Cora, F. J., & Fornal, C. A. (2002). Activity of medullary serotonergic neurons in freely moving animals. *Brain Research Reviews*, *40*(1–3), 45–52. https://doi.org/10.1016/s0165-0173(02)00187-x

Kim, S., Ka, S., Ha, J., Kim, J., Yoo, D., Kim, K., Lee, H., Lim, D., Cho, S., Hanotte, O., Mwai, O. A., Dessie, T., Kemp, S., Oh, S. J., & Kim, H. (2017). Cattle genome-wide analysis reveals genetic signatures in trypanotolerant N'Dama. *BMC Genomics*, *18*(1). https://doi.org/10.1186/s12864-017-3742-2

King, D. P., Zhao, Y., Sangoram, A. M., Wilsbacher, L. D., Tanaka, M., Antoch, M. P., Steeves, T. D., Vitaterna, M. H., Kornhauser, J. M., Lowrey, P. L., Turek, F. W., & Takahashi, J. S. (1997). Positional cloning of the mouse circadian gene. *Cell*, *89*(4), 641–653. https://doi.org/10.1016/s0092-8674(00)80245-7

Koch, C., Leinweber, B., Drengberg, B., Blaum, C., & Oster, H. (2016). Interaction between circadian rhythms and stress. *Neurobiology of Stress*, *6*, 57–67. https://doi.org/10.1016/j.ynstr.2016.09.001

Kume, K., Zylka, M. J., Sriram, S., Shearman, L. P., Weaver, D. R., Jin, X., Maywood, E. S., Hastings, M. H., & Reppert, S. M. (1999). MCRY1 and MCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell*, *98*(2), 193–205. https://doi.org/10.1016/s0092-8674(00)81014-4

Leone, V., Gibbons, S. M., Martinez, K., Hutchison, A. L., Huang, E. Y., Cham, C. M., Pierre, J. F., Heneghan, A. F., Nadimpalli, A., Hubert, N., Zale, E., Wang, Y., Huang, Y., Theriault, B., Dinner, A. R., Musch, M. W., Kudsk, K. A., Prendergast, B. J., Gilbert, J. A., & Chang, E. B. (2015). Effects of diurnal variation of gut microbes and High-Fat feeding on host circadian clock function and metabolism. *Cell Host & Microbe*, *17*(5), 681–689. https://doi.org/10.1016/j.chom.2015.03.006

Leroy, A., Latouche, P., Guedj, B., & Gey, S. (2022). MAGMA: inference and prediction using multi-task Gaussian processes with common mean. Machine Learning, 111(5), 1821–1849. https://doi.org/10.1007/s10994-022-06172-1

Leroy, A., Latouche, P., Guedj, B., & Gey, S. (2023). Cluster-Specific Predictions with Multi-Task Gaussian Processes. http://jmlr.org/papers/v24/20-1321.html

Lett, B. M., & Kirkpatrick, B. W. (2022). Identifying genetic variants and pathways influencing daughter averages for twinning in North American Holstein cattle and evaluating the potential for genomic selection. *Journal of Dairy Science*, *105*(7), 5972–5984. https://doi.org/10.3168/jds.2021-21238

Li, J., Akanno, E. C., Valente, T. S., Abo-Ismail, M., Karisa, B. K., Wang, Z., & Plastow, G. S. (2020). Genomic Heritability and Genome-Wide Association studies of plasma metabolites in crossbred beef cattle. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.538600

Lynch, S. V., & Pedersen, O. (2016). The human intestinal microbiome in health and disease. *New England Journal of Medicine*, *375*(24), 2369–2379. https://doi.org/10.1056/nejmra1600266

McCabe, C. J., Aryal, U. K., Casey, T., & Boerman, J. (2021). Impact of exposure to chronic Light–Dark phase shifting circadian rhythm disruption on muscle proteome in periparturient dairy cows. *Proteomes*, *9*(3), 35. https://doi.org/10.3390/proteomes9030035

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl variant effect Predictor. *Genome Biology*, *17*(1). https://doi.org/10.1186/s13059-016-0974-4

Meredith, B. K., Berry, D. P., Kearney, F., Finlay, E. K., Fahey, A. G., Bradley, D. G., & Lynn, D. J. (2013). A genome-wide association study for somatic cell score using the Illumina high-density bovine beadchip identifies several novel QTL potentially related to mastitis susceptibility. *Frontiers in Genetics*, *4*. https://doi.org/10.3389/fgene.2013.00229

Meyer-Bernstein, E., & Morin, L. (1996). Differential serotonergic innervation of the suprachiasmatic nucleus and the intergeniculate leaflet and its role in circadian rhythm

modulation. *Journal of Neuroscience*, *16*(6), 2097–2111. https://doi.org/10.1523/jneurosci.16-06-02097.1996

Momma, T., Okayama, H., Saitou, M., Sugeno, H., Yoshimoto, N., Takebayashi, Y., Ohki, S., & Takenoshita, S. (2017). Expression of circadian clock genes in human colorectal adenoma and carcinoma. *Oncology Letters*. https://doi.org/10.3892/ol.2017.6876

Montaruli, A., Castelli, L., Mulè, A., Scurati, R., Esposito, F., Galasso, L., & Roveda, E. (2021). Biological Rhythm and Chronotype: New Perspectives in health. *Biomolecules*, *11*(4), 487. https://doi.org/10.3390/biom11040487

Moore, R. Y., Halaris, A. E., & Jones, B. E. (1978). Serotonin neurons of the midbrain raphe: Ascending projections. *The Journal of Comparative Neurology*, *180*(3), 417–438. https://doi.org/10.1002/cne.901800302

Moore, R. Y., & Speh, J. C. (2004). Serotonin innervation of the primate suprachiasmatic nucleus. *Brain Research*, *1010*(1–2), 169–173. https://doi.org/10.1016/j.brainres.2004.02.024

Müller, N. A., Wijnen, C. L., Srinivasan, A., Ryngajllo, M., Ofner, I., Lin, T., Ranjan, A., West, D., Maloof, J. N., Sinha, N. R., Huang, S., Zamir, D., & Jiménez-Gómez, J. M. (2015). Domestication selected for deceleration of the circadian clock in cultivated tomato. *Nature Genetics*, *48*(1), 89–93. https://doi.org/10.1038/ng.3447

Noguchi, T., Leise, T. L., Kingsbury, N. J., Diemer, T., Wang, L. L., Henson, M. A., & Welsh, D. K. (2017). Calcium circadian rhythmicity in the suprachiasmatic nucleus: cell autonomy and network modulation. *eNeuro*, *4*(4), ENEURO.0160-17.2017. https://doi.org/10.1523/eneuro.0160-17.2017

Noyes, H., Brass, A., Obara, I., Anderson, S., Archibald, A. L., Bradley, D. G., Fisher, P., Freeman, A., Gibson, J., Gicheru, M., Hall, L., Hanotte, O., Hulme, H., McKeever, D., Murray, C., Oh, S. J., Tate, C., Smith, K., Tapio, M., . . . Kemp, S. J. (2011). Genetic and expression analysis of cattle identifies candidate genes in pathways responding to Trypanosoma congolense infection. *Proceedings of the National Academy of Sciences*, *108*(22), 9304–9309. https://doi.org/10.1073/pnas.1013486108

Oike, H., Oishi, K., & Kobori, M. (2014). Nutrients, clock genes, and chrononutrition. *Current Nutrition Reports*, *3*(3), 204–212. https://doi.org/10.1007/s13668-014-0082-6

Panda, S. (2016). Circadian physiology of metabolism. *Science*, *354*(6315), 1008–1015. https://doi.org/10.1126/science.aah4967

Penna, M. S., Hu, R., Rodney, G. G., & Cooper, T. A. (2023). The role ofLimch1alternative splicing in skeletal muscle function. *Life Science Alliance*, *6*(6), e202201868. https://doi.org/10.26508/lsa.202201868

Petrenko, V., Saini, C., Giovannoni, L., Gobet, C., Sage, D., Unser, M., Masson, M. H., Gu, G., Bosco, D., Gachon, F., Philippe, J., & Dibner, C. (2017). Pancreatic α- and β-cellular clocks have distinct molecular properties and impact on islet hormone secretion and gene expression. *Genes & Development*, *31*(4), 383–398. https://doi.org/10.1101/gad.290379.116

Pittendrigh, C. (1993). Temporal Organization: Reflections of a Darwinian Clock-Watcher. *Annual Review of Physiology*, *55*(1), 17–54. https://doi.org/10.1146/annurev.physiol.55.1.17

Plaut, K., & Casey, T. (2011). Does the circadian system regulate lactation? *Animal*, *6*(3), 394–402. https://doi.org/10.1017/s1751731111002187

Porcu, A., Riddle, M., Dulcis, D., & Welsh, D. K. (2018). Photoperiod-Induced neuroplasticity in the circadian system. *Neural Plasticity*, *2018*, 1–13. https://doi.org/10.1155/2018/5147585

Qu, Z., Zhang, H., Huang, M., Shi, G., Liu, Z., Xie, P., Li, H., Wang, W., Xu, G., Zhang, Y., Yang, L., Huang, G., Takahashi, J. S., Zhang, W. J., & Xu, Y. (2016). Loss of ZBTB20 impairs circadian output and leads to unimodal behavioral rhythms. *eLife*, *5*. https://doi.org/10.7554/elife.17171

Reinke, H., & Asher, G. (2015). Circadian clock control of liver metabolic functions. *Gastroenterology*, *150*(3), 574–580. https://doi.org/10.1053/j.gastro.2015.11.043

Reinke, H., & Asher, G. (2019). Crosstalk between metabolism and circadian clocks. *Nature Reviews Molecular Cell Biology*, *20*(4), 227–241. https://doi.org/10.1038/s41580-018-0096-9

Reppert, S. M., & Weaver, D. R. (2002). Coordination of circadian timing in mammals. *Nature*, *418*(6901), 935–941. https://doi.org/10.1038/nature00965

Said, R., Lobanova, L., Papagerakis, S., & Papagerakis, P. (2020). Calcium sets the clock in ameloblasts. *Frontiers in Physiology*, *11*. https://doi.org/10.3389/fphys.2020.00920

Sanchez, M.-P. (2019). *Analyse génétique de la composition protéique & des aptitudes fromagères du lait de vache prédites à partir des spectres moyen infrarouge* [Thèse de doctorat]. UniversitÈ Paris-Saclay (ComUE).

Sargolzaei, M., Chesnais, J. P., & Schenkel, F. S. (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*, *15*(1), 478. https://doi.org/10.1186/1471-2164-15-478

Sehgal, A. (2017). Physiology Flies with Time. *Cell*, *171*(6), 1232–1235. https://doi.org/10.1016/j.cell.2017.11.028

Sellix, M. T. (2014). Circadian clock function in the mammalian ovary. *Journal of Biological Rhythms*, *30*(1), 7–19. https://doi.org/10.1177/0748730414554222

Smolensky, M. H., Hermida, R. C., Reinberg, A., Sackett-Lundeen, L., & Portaluppi, F. (2016). Circadian disruption: New clinical perspective of disease pathology and basis for chronotherapeutic intervention. *Chronobiology International*, *33*(8), 1101–1119. https://doi.org/10.1080/07420528.2016.1184678

Solocinski, K., & Gumz, M. L. (2015). The circadian clock in the regulation of renal rhythms. *Journal of Biological Rhythms*, *30*(6), 470–486. https://doi.org/10.1177/0748730415610879

Stephens, G. J., & Cottrell, G. S. (2019). CACHD1: A new activity-modifying protein for voltage-gated calcium channels. *Channels*, *13*(1), 120–123. https://doi.org/10.1080/19336950.2019.1600968

Sun, Z. S., Albrecht, U., Zhuchenko, O., Bailey, J., Eichele, G., & Lee, C. C. (1997). RIGUI, a Putative Mammalian Ortholog of the Drosophila period Gene. *Cell*, *90*(6), 1003–1011. https://doi.org/10.1016/s0092-8674(00)80366-9

Tei, H., Okamura, H., Shigeyoshi, Y., Fukuhara, C., Ozawa, R., Hirose, M., & Sakaki, Y. (1997). Circadian oscillation of a mammalian homologue of the Drosophila period gene. *Nature*, *389*(6650), 512–516. https://doi.org/10.1038/39086

Thaiss, C. A., Levy, M., Korem, T., Dohnalová, L., Shapiro, H., Jaitin, D. A., David, E., Winter, D. R., Gury-BenAri, M., Tatirovsky, E., Tuganbaev, T., Federici, S., Zmora, N., Zeevi, D., Dori-Bachash, M., Pevsner-Fischer, M., Kartvelishvily, E., Brandis, A., Harmelin, A., . . . Elinav, E. (2016). Microbiota diurnal rhythmicity programs host transcriptome oscillations. *Cell*, *167*(6), 1495-1510.e12. https://doi.org/10.1016/j.cell.2016.11.003

Ueda, H. R., Hayashi, S., Chen, W., Sano, M., Machida, M., Shigeyoshi, Y., Iino, M., & Hashimoto, S. (2005). System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nature Genetics*, *37*(2), 187–192. https://doi.org/10.1038/ng1504

UEP, INRAE, 2019. Unité expérimentale du PIN. (2019). *Unité Expérimentale Du Pin*. https://doi.org/10.15454/1.5483257052131956e12

Wang, M., Yu, D., Zheng, L., Hong, B., Li, H., Hu, X., Zhang, K., & Mou, Y. (2021). Mechanical stress affects circadian rhythm in skeletal muscle (C2C12 myoblasts) by reducing PER/CRY gene expression and increasing BMAL1 gene expression. *Medical Science Monitor*, *27*. https://doi.org/10.12659/msm.928359

Wang, S., Wang, Y., Li, Y., Xiao, F., Guo, H., Gao, H., Wang, N., Zhang, H., & Li, H. (2022). Genome-Wide association study and selective sweep analysis reveal the genetic architecture of body weights in a chicken F2 resource population. *Frontiers in Veterinary Science*, *9*. https://doi.org/10.3389/fvets.2022.875454

Widmer, S., Seefried, F. R., Von Rohr, P., Häfliger, I. M., Spengeler, M., & Drögemüller, C. (2021). A major QTL at the LHCGR/FSHR locus for multiple birth in Holstein cattle. *Genetics Selection Evolution*, *53*(1). https://doi.org/10.1186/s12711-021-00650-1

Wirz-Justice, A. (2006). Biological rhythm disturbances in mood disorders. *International Clinical Psychopharmacology*, *21*(Supplement 1), S11–S15. https://doi.org/10.1097/01.yic.0000195660.37267.cf

Woldt, E., Sebti, Y., Solt, L. A., Duhem, C., Lancel, S., Eeckhoute, J., Hesselink, M. K. C., Paquet, C., Delhaye, S., Shin, Y., Kamenecka, T. M., Schaart, G., Lefebvre, P., Nevière, R., Burris, T. P., Schrauwen, P., Staels, B., & Duez, H. (2013). Rev-erb-α modulates skeletal muscle oxidative capacity by regulating mitochondrial biogenesis and autophagy. *Nature Medicine*, *19*(8), 1039–1046. https://doi.org/10.1038/nm.3213

Xie, L., Luo, C., Zhang, C., Zhang, R., Tang, J., Nie, Q., Ma, L., Hu, X., Li, N., Da, Y., & Zhang, X. (2012). Genome-Wide Association Study Identified a Narrow Chromosome 1 Region Associated with Chicken Growth Traits. *PLoS ONE*, *7*(2), e30910. https://doi.org/10.1371/journal.pone.0030910

Xie, Y., Tang, Q., Chen, G., Xie, M., Yu, S., Zhao, J., & Chen, L. (2019). New insights into the circadian rhythm and its related diseases. *Frontiers in Physiology*, *10*. https://doi.org/10.3389/fphys.2019.00682

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2010). GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, *88*(1), 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Yuan, J., Chen, S., Shi, F., Wu, G., Liu, A., Yang, N., & Sun, C. (2017). Genome-wide association study reveals putative role of gga-miR-15a in controlling feed conversion ratio in layer chickens. *BMC Genomics*, *18*(1). https://doi.org/10.1186/s12864-017-4092-9

Zerón-Rugerio, M. F., Cambras, T., & Izquierdo-Pulido, M. (2019). Social Jet Lag Associates Negatively with the Adherence to the Mediterranean Diet and Body Mass Index among Young Adults. *Nutrients*, *11*(8), 1756. https://doi.org/10.3390/nu11081756

Zhang, Y., Wang, H., Wang, Y., & Ma, B. (2025). FOXO1 mediates miR-99a-5p/E2F7 to restrain breast cancer cell proliferation and induce apoptosis. *BMC Cancer*, *25*(1). https://doi.org/10.1186/s12885-025-14111-1

Zhu, Y., Yao, S., Augustine, M. M., Xu, H., Wang, J., Sun, J., Broadwater, M., Ruff, W., Luo, L., Zhu, G., Tamada, K., & Chen, L. (2016). Neuron-specific SALM5 limits inflammation in the CNS via its interaction with HVEM. *Science Advances*, *2*(4). https://doi.org/10.1126/sciadv.1500637

Zimmerman, A. W., Joosten, B., Torensma, R., Parnes, J. R., Van Leeuwen, F. N., & Figdor, C. G. (2005). Long-term engagement of CD6 and ALCAM is essential for T-cell proliferation induced by dendritic cells. *Blood*, *107*(8), 3212–3220. https://doi.org/10.1182/blood-2005-09-3881

# Title: Genetic determinism of circadian activity in dairy cattle.

Keywords: Circadian rhythm, chronotypes, dairy cattle, MAGMA clustering, Genome-wide association studies GWAS.

Abstract: Circadian rhythms play a crucial role in regulating animal physiology processes and behavior. Studies have shown that there is variation in human circadian rhythms creating what is called a chronotype, and disruptions of these rhythms have deleterious consequences. Although circadian rhythms are well studied in humans, it remains less studied in animals and even lesser in cattle. In this exploratory work, we aimed to characterize individual chronotypes in dairy cows using high-frequency activity data and to explore their genetic determinism. Continuous activity records from Holstein and Normande cows were analyzed to quantify deviations from the mean of the population in daily activity patterns. We applied the MAGMA clustering algorithm to identify chronotypes and performed genome-wide-association-studies (GWAS) to detect genetic variants associated with these rhythmic phenotypes. Heritability estimates were moderately high in Holstein cows ($h^2$=0.61) but low in Normande ($h^2$=0.08). Variants with significant effects on chronotypes were identified in both breeds. They were annotated and several candidate genes involved in neuro-regulation and hormonal signaling pathways were identified (*ALCAM*, *CACHD1*, *LIMCH1*, *FSHR*, *LRNF5*, *ZBTB20*, *FOXO1*, *TNN*, *LRFN2*). These findings, for a unique study that is not yet explored in this field of bovine research, highlight the potential of chronotype as a novel trait and open avenues in precise livestock management through integrative chronobiological and genomic approaches.