Technische Universität München
Fakultät für Sport- und Gesundheitswissenschaften
Lehrstuhl für Sportbiologie
Prof. Henning Wackerhage

Wissenschaftliche Arbeit zur Erlangung des akademischen Grades
Master of Science (M.Sc.)

Circadian Rhythm Genes & Exercise: A Bioinformatic Analysis

vorgelegt von: Christian Soetebier
Matrikelnummer: 03633614
Studiengang: Diagnostik & Training
Sommersemester 2018

Christian Soetebier
Josef-Wirth-Weg 21, B037
80939 München
Tel.: 0160/94620457
E-Mail: CSoetebier@hotmail.com

Tag der Einreichung:
25. September 2018

Most biological processes follow an intrinsic day and night rhythm, orchestrated by an internal circadian clock. This genetic machinery consists of a core oscillator and multiple secondary loops that interact with each other, driving gene expression up and down throughout the day or in response to outside stimuli.
The expression of the core genes *ARNTL*, *CLOCK*, *PER1*, *PER2*, *CRY1* and *CRY2* as well as a host of functionally related genes of metabolism, myogenesis and cell cycle regulation such as *NR1D1*, *RORA*, *MYOD1* and *WEE1* is relevant to multiple exercise and health related areas of human physiology. This includes muscle (re-)growth, degree of hypertrophy, distribution of muscle fiber types, muscle ageing as well as diseases, phenotype changes and cancer conditions.
However, the ways these genes might interact with each other throughout these different contexts have not yet been linked or contrasted with each other. This study analyzes the behavior of important circadian genes across multiple exercise- and health related conditions.
As shown in previous literature, core circadian oscillator gene activity was often a main driving force in the response of circadian genes to various conditions. However, the present bioinformatic analysis also revealed potentially new overlapping and contrasting patterns of expression of these oscillator genes as well as other relevant genes involved in myogenesis, metabolism and the cell cycle.
These results demonstrate how diversely circadian genes influence the physiological response of mammals in the context of health and exercise. I expect this study to instigate more sophisticated research toward the application of our knowledge of circadian rhythm genes to improve the timing of exercise and drug interventions in chrono-exercise and chronotherapy.

Keywords: circadian rhythm, chrono-exercise, bioinformatics

Contents

List of Tables

List of Figures

0 Preface

0.1 Motivation

This thesis is concerned with the genetics of the human circadian rhythm in the context of health and exercise. This rather specialized topic stems from my own vested interests not only as a student of exercise science but also a coach and an athlete looking to improve performance in himself and others. I am therefore highly interested in the underlying mechanisms of adaptations to exercise in different contexts such as the circadian rhythm. It is also a product of the unique environment of the exercise biology department, which strongly enables and encourages students to investigate research topics they are personally interested in. The means to investigate this topic in the form of a bioinformatic analysis were given to me by my supervisor, Prof. Henning Wackerhage, who is the head of the exercise biology department.

0.2 Thesis Structure

Due to the nature of my research, where I performed a bioinformatic analysis consisting of 20 steps, I decided in accord with my supervisor, that I would

1. present my results and discuss them within the same section
2. introduce the studies I consulted and my method of reanalyzing them later

in order to give the thesis a better and clearer structure.
I will now continue with the introduction to this thesis before presenting the results of my analysis and discussing these in relation to the current literature.

1 Introduction

1.1 Background

The story of life on earth was always bound to the rising and setting of the sun, long before the first mammals evolved. The blueprint for life, the genetic code, of every living being reflects this in the form of mechanisms that establish rhythmic gene expression patterns to enable the organism to adapt to the changes in temperature and light that come with day and night.
Of course, even in our modern times, the human species is no exception to this. Even though we now have the ability to change temperature and light according to our will with the help of radiators, air conditioning and artificial light, we still experience the force of our natural circadian rhythms when it forces us to sleep. This most forcefully comes to bear when we act in contrast to our normal rhythm by leading lifestyles unsuitable to our own genetically coded preferences (chronotype), break our usual (entrained) rhythm or travel across time zones and experience jetlag. This is especially true for individuals who exercise or even work as an athlete: Exercise at unusual times or international competitions tax the individual's body if they are not performed in accordance to the natural rhythm of the relevant factors for performance. Additionally,

the efficacy of treatments involving exercise and medication for patients in hospital settings is also bound to circadian rhythms.

To be able to act, perform and be treated at the right time, i.e. in line with the circadian rhythm, however, one must first understand what the circadian rhythm is and how it works. For the last decades, scientific research has been trying to find out more about it. So far, we found out the basic structure of these patterns and uncovered multiple interlocked molecular loops that establish the circadian rhythm. However, knowledge as to how these mechanisms can be put into the context of exercise in a useful way, is still in its infancy.

Therefore, I will attempt to broaden the understanding of the genetics of the circadian rhythms in the context of exercise and related health topics with this thesis.

1.2 Objectives of the study

I will now further elaborate on the purpose of this study. This thesis involves a bioinformatic analysis concerning a set of known circadian genes in different contexts relevant to health and exercise (the former being the foundation of the latter). Its fundamental aims are therefore:

1. To establish the current scientific knowledge about the functions and interactions of these genes
2. To reanalyze pre-existing data involving these genes and bring it into the context of how these circadian genes work together to answer health- and exercise-related research questions
3. To find out where our knowledge is still lacking and give recommendations for further studies involving circadian genetics

More specifically, this study will attempt to answer the following research questions:

1. What is the general structure and localization of the circadian genes in question?
2. What is the effect of exercise on the researched circadian genes?
3. What is the role of the circadian genes in myogenesis after mechanical overload or injury?
4. Which phenotypes and diseases are associated with the circadian genes of the corpus?
5. What is the role of the circadian genes of the corpus regarding cancer risk, expression and survival?
6. Which practical recommendations can be given in regard to circadian genes and exercise?

In order to answer these research questions, I will first lay out the theoretical foundations of circadian rhythm genes, including their general structure, interaction and entrainment factors (chapter 2). In the following chapter (chapter 3), I will present the results of my bioinformatic analysis of data material stemming from different data bases and publications. After that I will summarize my results and put them into context to answer the research questions I outlined above (chapter 4). Within the same chapter I will also give practical recommendations. Methods will be shown in chapter 5. Finally, I will give a summary of my research questions, methods and results (chapter 6).

2 Theory of Circadian Rhythm and Genetics

2.1 Gene Transcription and Translation

To understand how circadian genes work, one must first have a basic understanding about gene expression and regulation in general. Therefore, I will first explain some underlying mechanisms of how genes work before discussing the circadian rhythm and its genetic machinery.
Genes are unable to influence the rest of the organism without direct interaction. Since they exist as a string of DNA within the nucleus, they need to be read by and converted into a protein, which can influence the rest of the system. For example, enzymes, a form of protein, metabolize nutrients or copy DNA when acting as a DNA polymerase during cell division (Clancy and Brown 2008). Proteins enact what is written in our genetic code. Differences between cells are determined by which sets of their genes are expressed (Hoopes 2008).
There are two important steps for a gene to be expressed and ultimately be converted into a protein (Clancy and Brown 2008):

1. Transcription, where the information of our DNA is transferred to a messenger RNA (mRNA)
2. Translation, where the single-stranded mRNA is read to create a protein

For the first step, transcription, to occur, the part of the DNA to be transcribed has first to be unwound from its original form. This is done by multiple histone modifications, transcription factor binding and other chromatin remodeling processes. Only then, specific genes can be read by proteins, who bind to the specific DNA sequences to either activate or repress it (Hoopes 2008). All of this happens within the cell's nucleus.
If the gene is activated, single-stranded mRNA will be formed to be translated. During translation, mRNA is used as a template to build amino acid chains that form a protein. To make this process happen, the mature mRNA molecules must leave the nucleus and travel to the ribosomes within the cytoplasm (Clancy and Brown 2008).
In summary, for a gene to have an effect, its code must be read during transcription within the nucleus. Here, it will either be activated or repressed. To form a protein, the mRNA resulting from transcription must move from the nucleus to the cytoplasm's ribosomes, where it will be translated into a protein.

2.2 The Circadian Rhythm and Genetic Loops

Circadian Rhythms have four major properties (Andreani et al. 2015):

1. They mirror the 24-hour solar day
2. Periodical patterns are shown even without exogenous clues
3. They can adjust to exogenous signals
4. The periodicity of rhythms is stable across a wide range of temperatures

Anatomically, these properties are realized within multiple tissues: In mammals, the suprachiasmatic nuclei (SCN) governs the daily rhythms of sleep and wake, functioning as master pacemaker. To perceive time of day, photic information is collected by the retina and transmitted to the SCN via the retinohypothalamic tract (RHT) (Froy 2011). The SCN also entrains various parts of the body through its output to certain regions of the hypothalamus. Those parts – in turn – signal the rest of the body and brain.

The master pacemaker consists of interconnected pacemaker neurons within each brain hemisphere. Independent groups within this network communicate with one another to integrate external cues and determine phase and period.

The networks of the master pacemaker drive behavioral rhythmicity utilizing the rhythms of neuronal activity within the nuclei, which is in turn regulated in part by the clock control of neuronal resting membrane potential. This membrane potential increases and decreases diurnally. More positive membrane potential increases excitability and firing rates, more negative potentials decrease excitability and firing rates.

Apart from the SCN, all tissues and organs are equipped with their own internal clock, whose genes cycle in most nucleated cells, except for the thymus and testis, and provide cycling gene expression throughout the body (Andreani et al. 2015).

The clocks of pancreas and liver are especially well-studied because of their importance for glucose homeostasis and metabolism. Studies on *ARNTL* knockout mice suggest a link between peripheral clock disfunction and diabetes and highlight the importance of these peripheral clocks (Rudic et al. 2004).

To make major circadian properties possible on the genetic level, our circadian system needs multiple functional genetic loops. The most important one is called the *Core Circadian Oscillator* (CCO).

The Core Circadian Oscillator, a cell intrinsic transcriptional/translational feedback loop (TTFL) drives the rhythmic expression of our circadian genes. The forward portion of the loop consists of transcription factors who upregulate the expression of genes that are the “negative” elements of the loop, which in turn downregulate the forward arm. The genes *ARNTL* and *CLOCK* form this forward portion of the loop (Andreani et al. 2015).

Over the course of the day the activity of the forward arm continues, the levels of the negative arm increase high enough that they can translocate into the cell nucleus and repress the positive component. When the negative elements are degraded, the cycle restarts and the forward arm is active again. The negative arm of the CCO consists of the genes *PER1*, *PER2*, *CRY1* and *CRY2*. These negative elements also serve distinct roles outside of the core mechanism, such as the regulation of internal clocks in different tissues, depending on which gene is looked at (Andreani et al. 2015). Another *PER* gene, *PER3*, was formerly regarded as part of the CCO. However, research by Bae et al. (2001) placed it outside the core oscillator, stating that “mPer3 is not essential for the core clock loops” (Bae et al. 2001).

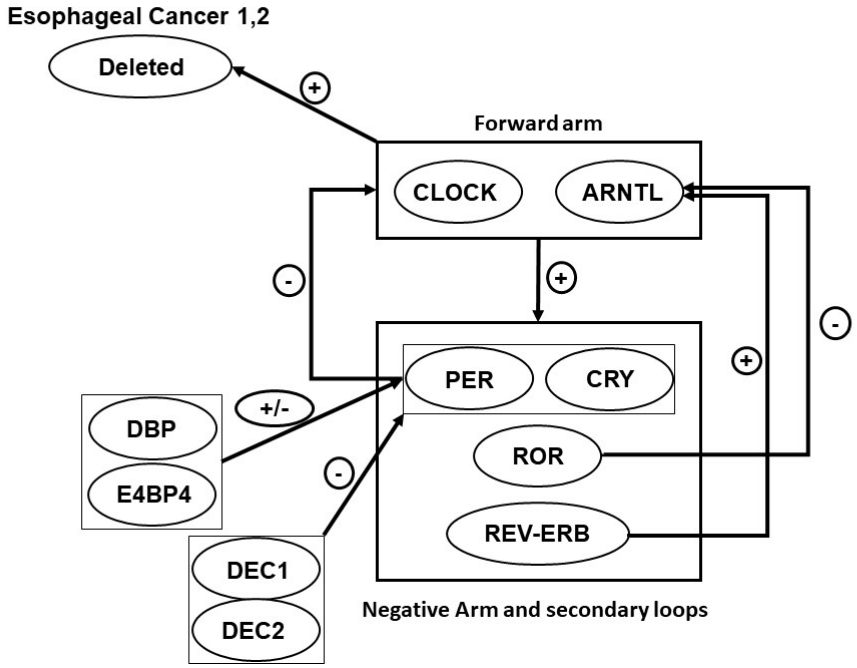


Figure 1: Schematic of the circadian gene system

REV-ERB and *ROR* regulate *ARNTL* instead of *CLOCK* itself. *DBP* and *E4BP4* bind to the promoter sequence (D-Box) of *mPer* and alter its expression pattern. *CLK-BMAL1* (= *CLOCK-ARNTL*) facilitates expression of Deleted in esophageal cancer1,2. *DEC1* and *DEC2* repress E-Box targets including *mPER* and their own site forming an interlocked feedback loop.

There are secondary loops in mammals that add complexity to the system (Figure 1) (Andreani et al. 2015):

- *REV-ERB* and *ROR* regulate *ARNTL* instead of *CLOCK* itself
- *DBP* and *E4BP4* bind to the promoter sequence (D-Box) of *mPer* and alter its expression pattern
- *CLK-BMAL1* (= *CLOCK-ARNTL*) facilitates expression of Deleted in esophageal cancer1,2
- *DEC1* and *DEC2* repress E-Box targets including *mPER* and their own site forming an interlocked feedback loop

2.3 Posttranslational Modifications

Phosphorylation of the clock proteins influences the components of the TTFL after translation to time their activity, subcellular localization and stability. A key protein kinase is *DOUBLETIME (DBT)/CASEIN KINASE I delta* and *epsilon* (*CK1*δ/ε). In mammals, *CK1*δ/ε mediate phosphorylation and degradation of *mPER2* and facilitate its entry into the nucleus. Mutations in *CK1*δ cause Familial Advanced Sleep Phase Disorder (FASPD). *CASEIN KINASE 2 (CK2)* promotes *PER* nuclear localization. Phosphatates also modulate the clock after translation: *PROTEIN PHOSPHATASE 1 (PP1)* dephosphorylates *PER2* and thus stabilizes it for reentry into the nucleus (Andreani et al. 2015).

2.4 Entrainment

Organisms use sensory cues to synchronize their clock to the 24h cycle. These cues are called zeitgebers, the german word for “time giver”. Light and food are well studied zeitgebers in current research (Andreani et al. 2015). However, according to Roenneberg, Kumar, and Merrow (2007), “there is no doubt that light is the most potent zeitgeber for the biological clock, but other environmental signals also entrain”.
To understand entrainment, one must differentiate between the central circadian clock (i.e. the SCN) and peripheral clocks within different tissues. For example, Tahara et al. (2015) found that sub-acute physical and psychological stress changed clock gene expression in the peripheral tissues in mice, but not in the suprachiasmatic nucleus. Therefore, we must differentiate the effects of zeitgebers on central and peripheral circadian clocks through photic and non-photic entrainment (Tahara, Aoyama, and Shibata 2017). I will now give a short overview about three zeitgebers currently being discussed in scientific literature that are related to the topic of this thesis: Light, food and exercise.

2.4.1 Photic entrainment: Light

As I already mentioned, light is the most potent zeitgeber according to Roenneberg, Kumar, and Merrow (2007). It also has the important ability of directly influencing the central circadian clock, i.e. the SCN, which was investigated over the course of multiple studies (Baehr, Fogg, and Eastman 1999; Roenneberg, Kumar, and Merrow 2007; Colwell 2011) and has become an important topic in the context of modern life filled with electrical sources of light.
Baehr, Fogg, and Eastman (1999) found that “bright light can phase shift human circadian rhythms”, while Colwell (2011) explains that “exposing an animal to light during the night also drives a robust increase in neural activity in the SCN”. The consequences of light exposure depend on the phase of the daily cycle, i.e. “a light stimulus during the night can induce a phase shift of the molecular oscillator, whereas it has no such effect when applied during the day” (Colwell 2011). However, in the absence of light stimuli, there are other non-photic stimuli that can entrain the circadian clocks.

2.4.2 Non-photic entrainment: Food

There have long been studies which observed that animals not only change their circadian behaviour in response to light, but to food as well. Richter (1922) observed that rats changed their behavior in anticipation of food. This led to the discovery of the pattern now known as food anticipatory activity (FAA). More recent studies found that feeding behavior can entrain organs such as the liver independently of the SCN and light exposure (Stokkan et al. 2001). By now, such entrainment processes could also be found in other organs such as the kidney, heart or pancreas. The food-induced phase resetting proved fastest in the liver (Damiola et al. 2000). However, food does not entrain the SCN itself (Stokkan et al. 2001).
This led to the still ongoing search for possible food-entrainable oscillators (FEOs) (Stephan 1984), who can entrain peripheral clocks to the external zeitgeber, i.e. food. So far, we know that food is an important zeitgeber for our peripheral clocks. However, the loci of potential oscillators as well as their hierarchy in relation to the SCN are still unknown to us.

2.4.3 Non-photic entrainment: Exercise

Another potentially important zeitgeber, especially in the context of this thesis, is exercise. Buxton et al. (2003) found that exercise can elicit phase shifts and changes to the levels of melatonin in human subjects, with longer workout durations having more consistent effects. Miyazaki et al. (2001) reported similar findings when they reported that "Phase-advance shifts of human circadian pacemaker are accelerated by daytime physical exercise" (Miyazaki et al. 2001). In their study, physical exercise in the afternoon or at midnight significantly phase delayed melatonin rhythms.
In summary, multiple studies report that phase shifts to melatonin levels seem to be the major mechanism of circadian entrainment by exercise. Richardson et al. (2017) proposes that there could be multiple mechanisms at play. Exercise may

- suppress and either delay or advance excretion of melatonin (depending on time of day)
- alter melatonin degradation and thus plasma melatonin concentrations
- reduce melatonin production at the exercise time

2.5 Chronotypes

Apart from the mechanisms that entrain our clocks, we can also look at the general behavioural manifestations of the circadian rhythm. These preferences in the timing of sleep and wake are called chronotypes and are broadly differentiated in early types ("larks") and late types ("owls"). They depend on genetic disposition as well as entrainment, i.e. light exposure and sleeping patterns (Roenneberg, Wirz-Justice, and Merrow 2003). Lack et al. (2009) found that the two types generally have different sleep patterns, with late types having their sleep patterns initiated 2-3 hours later. The same was true for arising times, circadian phase markers for temperature and melatonin as well as objective sleepiness.
The fact that circadian signals and the kinetics of sleep pressure buildup differ between the two chronotypes (Taillard et al. 2003) shows that it is important to differentiate between early and late chronotypes. Yet, this is only a broad differentiation and there will be many individuals in between those extremes. To find out where one fits in the spectrum, it is possible to use questionnaires such as the "Horne and Östberg Questionnaire" (Horne and Östberg 1976) or the "Munich Chronotype Questionnaire" (Juda, Vetter, and Roenneberg 2013).

3 Results and Discussion

Now, having established the theoretical basics of circadian genetics and other relevant factors, I will commence with presenting the results of my analyses and promptly discuss them.

3.1 String Database

To find out whether the circadian genes I study are functionally linked and whether they form distinct groups, I employed STRING database. This database includes known and predicted protein-protein inter-

actions, including direct physical as well as indirect functional associations stemming from computational prediction. It also employs knowledge transfers from other organisms and searches for interactions aggregated from other databases (string-db.org 2018).

StringDB Analysis (Figure2) revealed that the gene corpus has more interactions than what would be expected by a random set of genes, which “indicates that the proteins are at least partially biologically connected, as a group” (string-db.org 2018).

number of nodes:	32	expected number of edges:	3
number of edges:	102	PPI enrichment p-value:	< 1.0e-16
average node degree:	6.38	*your network has significantly more interactions*	
avg. local clustering coefficient:	0.636	*than expected (what does that mean?)*	

Figure 2: StringDB Network stats

If we look at the interactions shown in Figure 3, the first noticeable pattern is formed by a cluster consisting of *ARNTL*, *CLOCK*, *PER1*, *PER2*, *CRY1* and *CRY2*. These genes have strong interactions with a high combined score ($>$0.99), occupying six of the first ten highest ranked interactions (for further explanation see Chapter 5.2). These genes are known to form the core circadian oscillator, with *ARNTL* and *CLOCK* forming the forward arm, *PER1*, *PER2*, *CRY1* and *CRY2* forming the negative arm (Andreani et al. 2015; Lowrey and Takahashi 2011; Pittendrigh 1960). The vital role of *CLOCK* has been challenged by DeBruyne et al. (2006), when they found that mouse *CLOCK* is not required for circadian oscillator function. An absence of *CLOCK* is, however, not without consequences like altered response to light or other alterations. This is unsurprising because of the many roles *CLOCK* plays regarding transcription and expression of many genes (Miller et al. 2007).

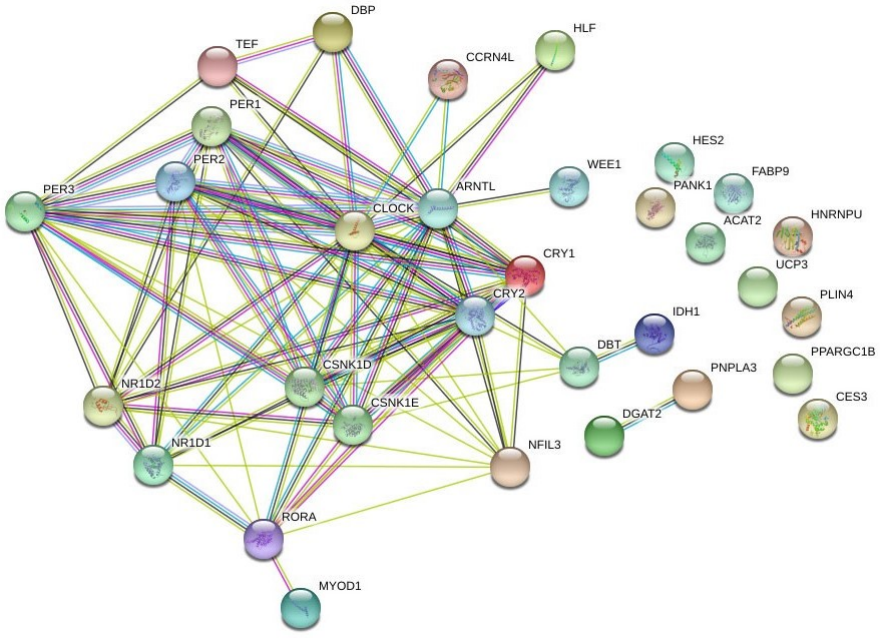


Figure 3: Functional interactions between the clock genes

Another interesting result of string database analysis is the strong influence of *PER3* on many members of the core circadian oscillator: Its interactions with *CRY2*, *CRY1*, *PER2* and *PER1* have high combined scores ($>$0.99). This interaction was also researched by Bae et al. (2001) when they placed *PER3* outside of the core circadian oscillator. Even though it seems to be not essential, it is a relevant modulator of the CCO.

Also connected to the group forming the core circadian oscillator are multiple genes concerned with metabolism: *CCRN4L* interacts with both components of the forward arm (*ARNTL* and *CLOCK*) with high combined scores (0.943 and 0.925 respectively). This is interesting since Kawai et al. (2010) revealed that it promotes adipogenesis by stimulating *PPAR-*γ nuclear translocation, meaning that it plays a role in body composition. This could be especially important for health issues such as obesity as well as competitive athletes playing sports with weight divisions.

As another gene concerned with metabolism, *RORA* interacts with most members of the core circadian oscillator as well as *NR1D1* (score 0.924). This finding matches those of Sato et al. (2004), who found that "Rora is required for normal [*ARNTL*] expression and consolidation of daily locomotor activity and is regulated by the core clock in the SCN". They suggest that "opposing activities of the orphan nuclear receptors Rora and Rev-erb α [*NR1D1*], which represses [*ARNTL*] expression, are important in the maintenance of circadian clock function". These findings correspond to those of Kim et al. (2018), who found that in liver,

> "Rev-erbα, a core repressive transcription factor of the clock, opposes functional loop formation between Rev-erbα,–regulated enhancers and circadian target gene promoters by recruitment of the NCoR-HDAC3 co-repressor complex, histone deacetylation, and eviction of the elongation factor BRD4 and the looping factor MED1" (Kim et al. 2018).

This means that it operates as "a repressive arm of the molecular clock" (Kim et al. 2018).

Another supporting study is that of Sen et al. (2018), who found that "Rev-erbα in the brain is involved in the temporal partitioning of feeding and sleep, whereas its effects on energy metabolism are mainly exerted through its peripheral expression". This paints an interesting picture considering that Liu et al. (2007) found that "PGC-1a (Ppargc1a), a transcriptional coactivator that regulates energy metabolism" is stimulating "the expression of clock genes, notably Bmal1 (*Arntl*) and Rev-erba (Nr1d1)" making it "a key component of the circadian oscillator that integrates the mammalian clock and energy metabolism". This could potentially have implications for a role of endurance training in circadian clock setting within the brain and liver.

Other interactions of note, not or not highly supported by experimental studies, are as follows:

- *DBP* has interactions with the core heterodimer (*CLOCK*/*ARNTL*) as well as *TEF* and *NR1D2*.
- *TEF* interacts with the forward arm of the core oscillator (0.483 and 0.436 with *ARNTL* and *CLOCK*) as well as *PER3* (0.458).
- *WEE1* interacts only with *ARNTL* (0.457)
- *DBT* interacts with *CSNK1D*, *CSNK1E*, *CLOCK*, *ARNTL*, *PER2*, as well as *IDH1*

HLF interacts only with the core heterodimer (*CLOCK*/*ARNTL*), and its connection with *ARTNL* has a reasonably high score (0.714) that is founded on some experimental data (0.324). However, I could not find relevant literature regarding this interaction.

DGAT2 and *PNPLA3* interact with each other and have a very high combined score (0.943). Looking further into the analysis, however, showed that this is because of association in curated databases and being co-mentioned in PubMed Abstracts, not due to experimental data.

All in all, my StringDB analysis was able to replicate the so far uncovered circadian loops including the core circadian oscillator. Other potentially interesting links between CCO and other genes as well as of genes outside the CCO could be shown and could provide potentially fruitful research opportunities.

3.2 ToppGene

To see whether the genes I study have common features, I employed the ToppGene Suite. This is an online portal for gene list functional enrichment and candidate gene prioritization of novel disease candidates. It uses functional annotations and network analysis to achieve this.

The first part of the ToppGene Analysis concerns itself with molecular functions. GO: Molecular function (Table 1) shows that the genes under investigation are part of the processes of binding of DNA, especially regarding transcription.

Table 1: Top 5 results for GO: Molecular Function

	ID	Name	Source	p-value	FDR B&H	FDR B&Y	Bonferroni	Genes from Input	Genes in Annotation
1	GO:0003690	double stranded DNA binding		6.363 E-13	1.432 E-10	8.584 E-10	1.432 E-10	15	823
2	GO:0001047	core promoter binding		4.283 E-12	4.819 E-10	2.889 E-9	9.637 E-10	9	164
3	GO:0044212	transcription regulatory region DNA binding		2.240 E-11	1.383 E-9	8.290 E-9	5.041 E-9	14	860
4	GO:0001067	regulatory region nucleic acid binding		2.458 E-11	1.838 E-9	8.290 E-9	5.531 E-9	14	866
5	GO:000976	transcription regulatory region sequence-specific DNA binding		3.095 E-11	1.393 E-9	8.351 E-9	6.965 E-9	13	704

31 input genes in category / 225 annotations before applied cutoff / 18661 genes in category

The biological processes described in GO: biological processes (Table 2) are part of what can best be surmised as the processes for regulating the circadian rhythm, which can be expected of a gene set consisting of circadian genes. They are, however, also part of pathways regarding steroid and corticosteroid hormones as well as lipid metabolism. The connection between circadian genes and corticosteroid was established by Tahara, Aoyama, and Shibata (2017), who suggest that “the common signaling pathways

of clock entrainment by stress and exercise involve sympathetic nervous activation and glucocorticoid release". The interaction of the clock genes and metabolism is already well established, for example by Bass and Takahashi (2010), who investigated crosstalk between the clock and metabolic transcription networks.

Table 2: Top 5 results for GO: Biological Process

	ID	Name	Source	p-value	FDR B&H	FDR B&Y	Bonferroni	Genes from Input	Genes in Annotation
1	GO:0032922	circadian regulation of gene expression		3.164 E-25	3.813 E-22	2.925 E-21	3.813 E-22	13	62
2	GO:0007623	circadian rhythm		1.353 E-23	8.152 E-21	6.254 E-20	1.630 E-20	16	214
3	GO:0048511	rhythmic process		2.062 E-23	8.282 E-21	6.354 E-20	2.485 E-20	18	266
4	GO:0042752	regulation of circadian rhythm		1.774 E-21	5.343 E-19	4.099 E-18	2.137 E-18	13	115
5	GO:2000323	negative regulation of glucocorticoid receptor signaling pathway		5.458 E-14	1.315 E-11	1.009 E-10	6.577 E-11	5	6

31 input genes in category / 1205 annotations before applied cutoff / 18623 genes in category

GO: Cellular Component (Table 3) shows that the cellular components of the genes investigated here are found in the cytoplasmic ribonucleoprotein granule and the ribonucleoprotein granule, "[a] non-membranous macromolecular complex containing proteins and translationally silenced mRNAs" (yeastgenome.org 2018), as well as the chromatoid body, which is a dense structure in the cytoplasm of male germ cells. All these structures are relevant for the regulation of gene expression.

Table 3: Top 3 results for GO: Cellular Component

	ID	Name	Source	p-value	FDR B&H	FDR B&Y	Bonferroni	Genes from Input	Genes in Annotation
1	GO:0036464	cytoplasmic ribonucleoprotein granule		9.530 E-5	6.108 E-3	3.192 E-2	9.911 E-3	4	144
2	GO:0035770	ribonucleoprotein granule		1.175 E-4	6.108 E-3	3.192 E-2	1.222 E-2	4	152
3	GO:0033391	chromatoid body		2.454 E-4	8.4.446 E-2	2.552 E-2	2	14	

32 input genes in category / 104 annotations before applied cutoff / 19061 genes in category

Mouse phenotypes that could be produced by manipulating the genes in my corpus (Table 4) show mostly abnormalities of the circadian rhythm regarding behavior as well as sleep/wake cycles. Outside of the Top 5 selection, there were also phenotypes with abnormal homeostasis and abnormal function of the cardiovascular system. Phenotypes also showed abnormalities of fatty acid metabolism. This is unsurprising

since many metabolic genes are present within the investigated corpus.

Table 4: Top 5 results for Mouse Phenotype

	ID	Name	Source	p-value	FDR B&H	FDR B&Y	Bonferroni	Genes from Input	Genes in Annotation
1	MP:0020468	abnormal circadian behavior period		1.036 E-22	1.253 E-19	9.617 E-19	1.253 E-19	12	37
2	MP:0020470	shortened circadian behavior period		1.877 E-19	1.134 E-16	8.706 E-16	2.269 E-16	10	28
3	MP:0020467	abnormal circadian behavior		1.625 E-18	6.547 E-16	5.025 E-15	1.964 E-15	12	76
4	MP:0020472	arrhythmic circadian behavior persistence		1.013 E-15	3.060 E-13	2.349 E-12	1.224 E-12	7	12
5	MP:0020471	abnormal circadian behavior persistence		2.189 E-15	5.294 E-13	4.036 E-12	2.647 E-12	7	13

29 input genes in category / 1209 annotations before applied cutoff / 9319 genes in category

The pathways found by ToppGene (Table 5) are mostly of circadian function, as it is to be expected when looking at the host of literature describing circadian genes, especially concerning those belonging to the circadian clock system (Shearman et al. 2000; Reppert and Weaver 2001; Lowrey and Takahashi 2011; Andreani et al. 2015). Other pathways are part of lipid metabolism or circadian entrainment, a topic I already discussed in the previous section. In short, McCarthy et al. (2007) identified various circadian genes involved in lipid metabolism, i.e. *NR1D1* (Cho et al. 2012; Kim et al. 2018), *RORA* (Lau et al. 2004), *CCRN4L* (Kawai et al. 2010) and *PPARGC1B* (Kamei et al. 2003).

Table 5: Top 5 results for Pathway

	ID	Name	Source	p-value	FDR B&H	FDR B&Y	Bonferroni	Genes from Input	Genes in Annotation
1	P00015	Circadian clock system	PantherDB	7.257 E-25	1.473 E-22	8.682 E-22	1.473 E-22	9	9
2	83084	Circadian rhythm	BioSystems: KEGG	1.619 E-22	1.587 E-20	9.353 E-20	3.286 E-20	11	31
3	1269871	Circadian Clock	BioSystems: REACTOME	2.346 E-22	1.587 E-20	9.353 E-20	4.762 E-20	13	71
4	1269872	BMAL1: CLOCK,NPAS2 activates circadian gene expression	BioSystems: REACTOME	4.218 E-20	2.141 E-18	1.261 E-17	8.562 E-18	11	48
5	137978	circadian rhythm pathway	BioSystems: Pathway Interaction Database	1.128 E-10	4.581 E-9	2.700 E-8	2.290 E-8	5	14

30 input genes in category / 203 annotations before applied cutoff / 12450 genes in category

ToppGene's automated PubMed search (Table 6) uncovered a host of disease links to psychological disorders such as depression, schizophrenia, bipolar disorder as well as alcoholism. These associations are already well established within the current scientific literature:

- Nievergelt et al. (2006) and Mansour et al. (2006) suspect a connection of *PER3* and *ARNTL* with bipolar disorder
- McClung (2007) hypothesizes that Seasonal Affective Disorder is rooted in circadian phase shifts
- Soria et al. (2010) linked *CRY1* and *NPAS2* with unipolar major depression

It also showed results concerned with furthering our understanding of the circadian rhythm functions.

Table 6: Top 5 results for Pubmed

	ID	Name	Source	p-value	FDR B&H	FDR B&Y	Bonferroni	Genes from Input	Genes in Annotation
1	19693801	PER2 variation is associated with depression vulnerability.	Pubmed	3.720 E-35	1.914 E-31	1.746 E-30	1.914 E-31	12	16
2	20174623	Systematic analysis of circadian genes in a population-based sample reveals association of TIMELESS with depression and sleep disturbance.	Pubmed	3.791 E-34	9.751 E-31	8.896 E-30	1.950 E-30	12	18
3	26332085	Chronic Alcohol Exposure and the Circadian Clock Mutation Exert Tissue-Specific Effects on Gene Expression in Mouse Hippocampus, Liver, and Proximal Colon.	Pubmed	3.295 E-33	5.651 E-30	5.156 E-29	1.695 E-20	15	72
4	20072116	Differential association of circadian genes with mood disorders: CRY1 and NPAS2 are associated with unipolar major depression and CLOCK and VIP with bipolar disorder.	Pubmed	2.666 E-30	3.429 E-27	3.129 E-26	1.372 E-26	11	19
5	19839995	Association study of 21 circadian genes with bipolar I disorder, schizoaffective disorder, and schizophrenia.	Pubmed	1.243 E-29	1.279 E-26	1.167 E-25	6.395 E-26	11	21

32 input genes in category / 5145 annotations before applied cutoff / 36332 genes in category

The interactions found by ToppGene analysis (Table 7) show *PER3* and *PER2* interactions as highest in rank, further illuminating their importance within the circadian gene system. Other interactions ranked highly are those of the forward arm (*ARNTL* and *CLOCK*) as well as the negative arm (*PER1*, *CRY1*, *CRY2*). Also ranked highly were interactions of *CSNK1E* and *BHLHE41*.

The interactions with *DEC1* are highly ranked, but its functions are unknown:

"The function of this gene is not known. This gene is located in a region commonly deleted in esophageal squamous cell carcinomas. Gene expression is reduced or absent in these carcinomas and thus this is a candidate tumor suppressor gene for esophageal squamous cell carcinomas" (genecards.org 2018a).

Table 7: Top 5 results for Interaction

	ID	Name	Source	p-value	FDR B&H	FDR B&Y	Bonferroni	Genes from Input	Genes in Annotation
1	int:PER3	PER3 interactions		8.742 E-22	1.168 E-18	9.080 E-18	1.168 E-18	9	17
2	int:PER2	PER2 interactions		5.072 E-19	3.388 E-16	2.634 E-15	6.777 E-16	9	30
3	int:DEC1	DEC1 interactions		9.922 E-19	4.419 E-16	3.436 E-15	1.326 E-15	9	32
4	int:ARNTL	ARNTL interactions		1.403 E-17	4.369 E-15	3.397 E-14	1.874 E-14	10	69
5	int:CLOCK	CLOCK interactions		1.635 E-17	4.369 E-15	3.379 E-14	2.184 E-14	10	70

30 input genes in category / 1336 annotations before applied cutoff / 1763 genes in category

Diseases associated with circadian genes (Table 8) include disorders regarding sleep (e.g. advanced sleep phase syndrome) as well as a host of psychological diseases such as seasonal affective disorder or winter depression.

Table 8: Top 5 results for Disease

	ID	Name	Source	p-value	FDR B&H	FDR B&Y	Bonferroni	Genes from Input	Genes in Annotation
1	C0085159	Seasonal Affective Disorder	DisGeNET Curated	1.883 E-13	1.254 E-10	8.880 E-10	1.254 E-10	7	32
2	C0854739	Advanced sleep phase	DisGeNET BeFree	1.254 E-11	4.176 E-9	2.957 E-8	8.353 E-9	4	4
3	C0393770	Delayed Sleep Phase Syndrome	DisGeNET BeFree	1.873 E-10	4.158 E-8	2.943 E-7	1.247 E-7	4	6
4	C3807327	ADVANED SLEEP PHASE SYNDROME, FAMILIAL, 1	DisGeNET Curated	7.007 E-9	1.167 E-6	8.259 E-6	4.667 E-6	3	3
5	C0851578	Sleep Disorders	DisGeNET BeFree	1.037 E-8	1.186 E-6	8.393 E-6	6.909 E-6	5	38

32 input genes in category / 666 annotations before applied cutoff / 16205 genes in category

One example of this would be a hereditary form of Delayed Sleep Phase Disorder (DSPD) associated with a dominant coding variation in the core circadian clock gene *CRY1*, found by Patke et al. (2017). Also, highly ranked among associated diseases are obesity and metabolic syndrome X. Froy (2009) already established "interconnection between the circadian clock and metabolism, with implications for obesity and how the circadian clock is influenced by hormones, nutrients, and timed meals".
To summarize, my ToppGene analysis showed enrichment of the genes in the corpus in a host of different fields, most importantly:

1. It revealed that the genes are transcriptionally active and influence the expression levels of other genes.
2. A high prevalence of functions regarding metabolism and homeostasis links the circadian genes to metabolic functions.
3. When dysfunctional, mutations of circadian genes can lead to a host of phase shift-based sleep disorders as well as psychological disorders such as depression or bipolar disorder.

3.3 Tissue Expression

I employed Gtex Portal (gtexportal.org 2018) to find out in which tissues the genes I study are expressed. My analysis focused on the different expression levels of circadian genes across different tissue, with a special focus on how much it is expressed in skeletal muscle relative to other tissues. This was done in order to determine how important the genes could be in regard to exercise. After careful examination of the different expression graphs provided by Gtex Portal, the broad results are as follows (Table 9):

Table 9: Distribution of the relative expression of clock genes in skeletal muscle

Skeletal Muscle Expression	**Number**	**Percentage**
Not expressed	6	18,75%
Low expression	15	46,88%
Medium expression	8	25%
High expression	1	3,13%
Exclusively in SM	2	6,25%
Total	32	

I found that most (15 out of 32) genes of the corpus I investigated have only low expression rates in skeletal muscle compared to other tissues. Some are not expressed there at all (six out of 32). Some genes are of average expression (eight out of 32), one gene is highly expressed in skeletal muscle and two are exclusively expressed there.
The gene that is highly expressed in skeletal muscle is *CRY2* (Figure 4), while *MYOD1* and *UCP3* are exclusively expressed in skeletal muscle (Figures 5a and 5b). Upon discovering similar results in their study, McCarthy et al. (2007) concluded that "[t]he observation that genes like Myod1, Atrogin1, and PGC-1beta are under circadian control strongly suggests circadian rhythms will have a central role in the daily

regulation of skeletal muscle function and phenotype".

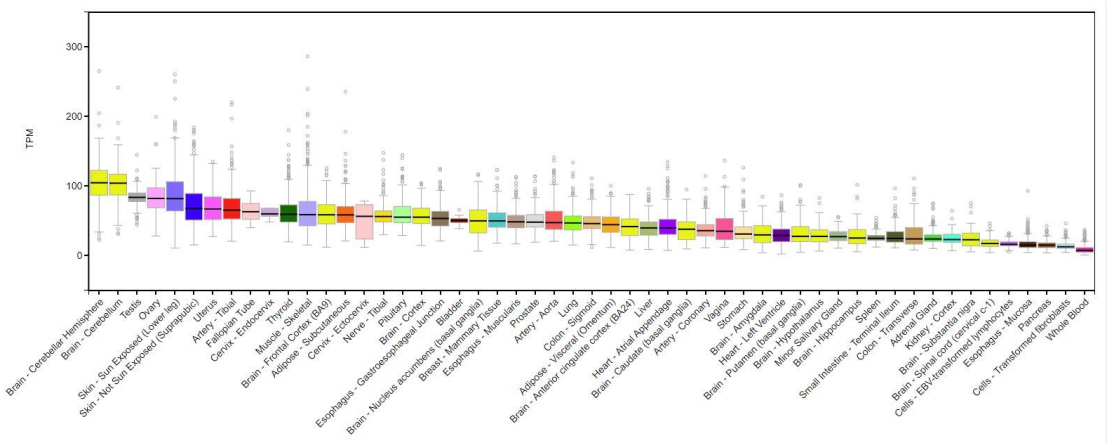
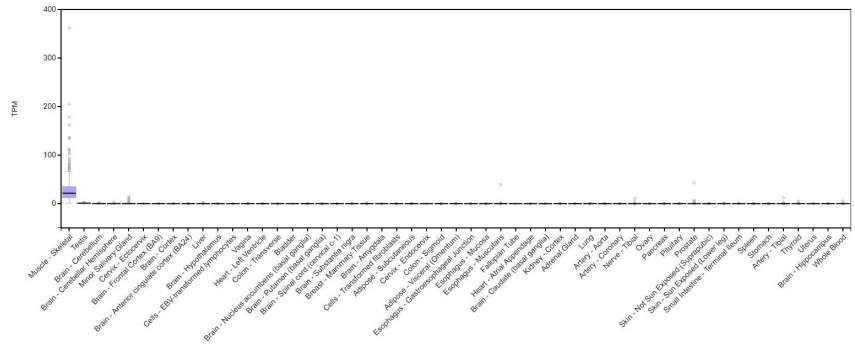
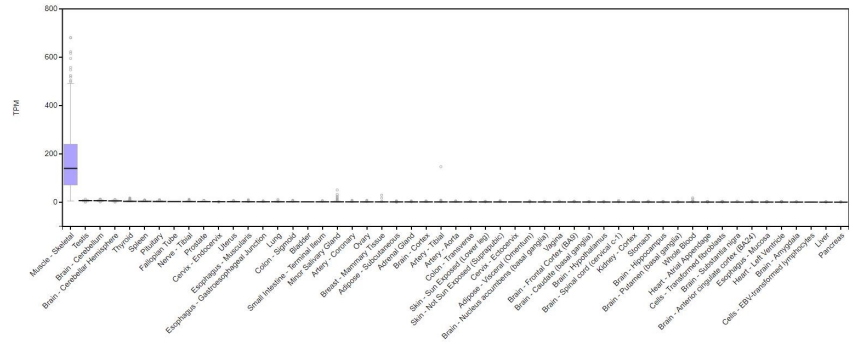


Figure 4: Expression of *CRY2* in different human tissues

Andrews et al. (2010) found that "MyoD is a direct target of the circadian transcriptional activators CLOCK and BMAL1, which bind in a rhythmic manner to the core enhancer of the MyoD promoter", concluding from their study that "the circadian clock may generate a muscle-specific circadian transcriptome in an adaptive role for the daily maintenance of adult skeletal muscle".



(a) Expression of *MYOD1* in different human tissues



(b) Expression of *UCP3* in different human tissues

Figure 5: Expressions of *MYOD1* and *UCP3* in human tissues

This implicates that it could be possible to optimize results of resistance training by timing it to circadian timing of *MYOD1* expression. Hayes, Bickerstaff, and Baker (2010) already suggested in a study regarding anabolic hormones such as testosterone that "(a) athletes are advised to coincide training times with

performance times, and (b) individuals may experience greater hypertrophy and strength gains when resistance training protocols are designed dependent on individual [testosterone] response".

I suppose that with enough research, similar recommendations could be given regarding the timing of resistance training with *MYOD1* peak expression. However, this notion clearly needs more research to become applicable.

Vidal-Puig et al. (1997) discerned *UCP3* from the other uncoupling proteins by its "abundant and preferential expression in skeletal muscle" and hypothesize it "may be an important mediator of adaptive thermogenesis". Indeed, McCarthy et al. (2007) suggest it "acts to facilitate fatty acid oxidation in skeletal muscle cells". Upon finding that it has a circadian pattern of expression, they suggest that "Myod1 may be involved in the regulation of Ucp3 circadian expression given that Ucp3 peak expression occurs 8 h after peak Myod1 expression" (McCarthy et al. 2007).

A link between those two genes would be most interesting, since it would suggest a connection between a muscle maker gene and a lipid metabolism gene – maybe this could lead to a better understanding of how to use circadian timing to improve upon methods of losing weight and building muscle. compared to other tissue, the rest of the corpus is averagely expressed in skeletal muscle.

While McCarthy et al. (2007) note that "[e]vidence from mice with a targeted deletion of PGC-1β clearly demonstrates the importance of PGC-1β in skeletal muscle", my results show that it is not more than averagely expressed compared to other tissues on Gtex Portal (Figure 6). However, this brings up the notion that average expression must not be confused with lack of importance.

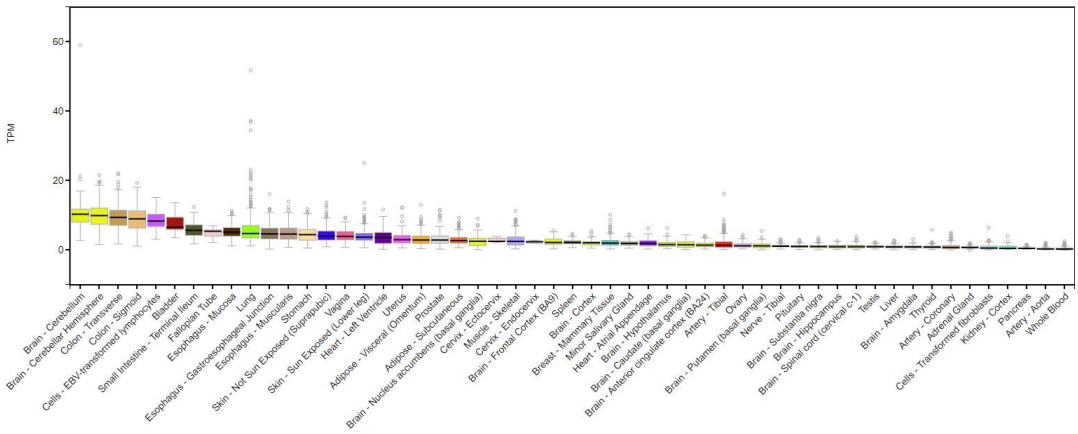


Figure 6: Expression of *PPARGC1B* in different human tissues

The fact that most genes I examined were in fact expressed in skeletal muscle was very much expected, since they were included for exactly this reason. McCarthy et al. (2007) provided much of my corpus by their examination of circadian genes found in muscle. However, it is of note that they are, in most cases, not most active in skeletal muscle. This is also true for the fact that six genes were not found in skeletal muscle at all.

Alas, the influence of the circadian genome does not always have to be a direct one. Many genes of the corpus were not only highly expressed in the adrenal gland but also within the cerebellar hemisphere of the brain as well as the cerebellum. One gene in particular is known for its function at this specific

brain area: *NR1D1* (Figure 7). Delezie et al. (2016) linked it to circadian food entrainment, i.e. "food-anticipatory behaviour and thermogenesis" and concluded that "Rev-erbα [NR1D1] is required for neural network-based prediction of food availability" (Delezie et al. 2016). Results like these further underline the importance of clock genes in the cerebellum as well as other brain areas in coordinating our body's clock, and especially, our muscle clock from within the brain. They also show that the influence of circadian genes does not only lie within the pure physiology but also can change behavior.

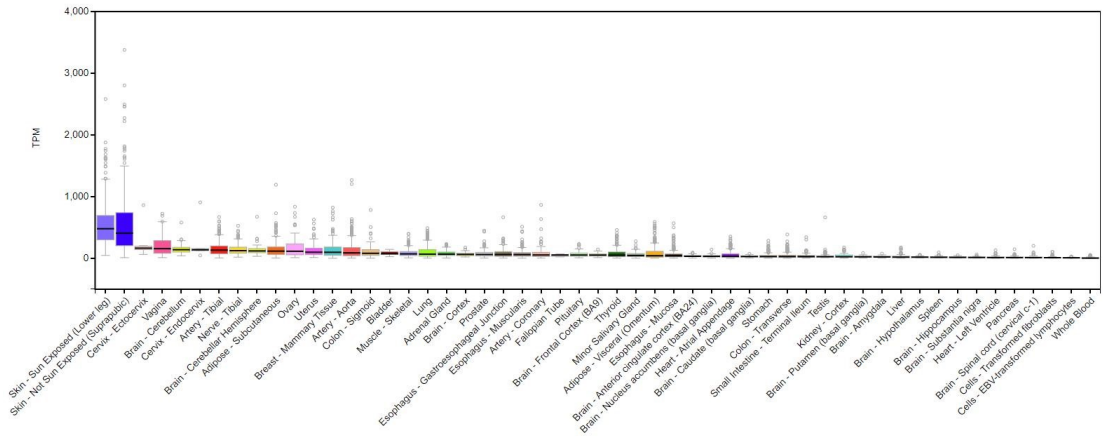


Figure 7: Expression of *NR1D1* in different human tissues

Another noteworthy result is an unexpectedly high prevalence of circadian genes in sun exposed and not sun exposed skin. Other studies explain this with the role of the circadian genome regarding the cell cycle (Bjarnason et al. 2001), the nucleotide excision repair system (Gaddameedhi et al. 2011) as well as the general "control of skin and cutaneous photodamage" (Desotelle, Wilking, and Ahmad 2012; Zanello, Jackson, and Holick 2000).

Seuter et al. (2013) found that the *BHLHE40* gene is "a dynamically regulated primary target of the vitamin D receptor", which coincides with its relatively high expression in sun exposed and not sun exposed skin (Figure 8). This might be relevant in the context of outdoor exercise.

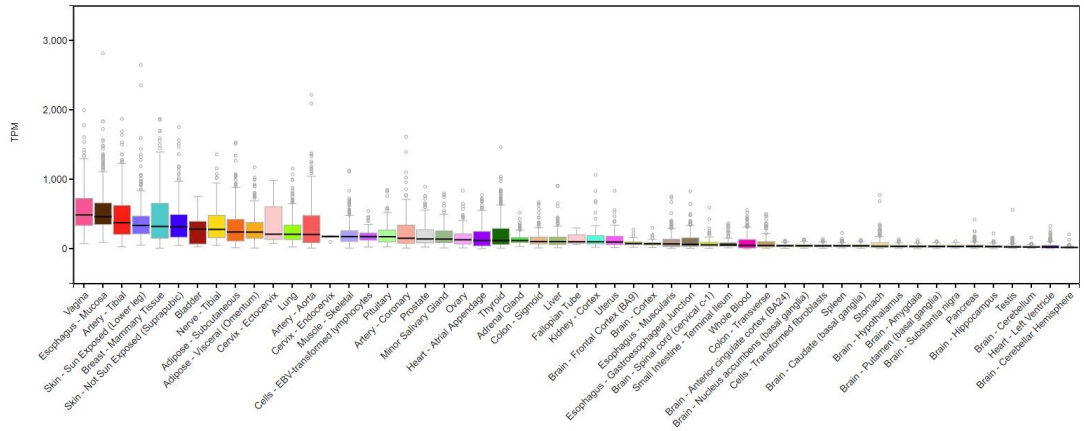


Figure 8: Expression of *BHLHE40* in different human tissues

To summarize, my tissue expression analysis on Gtex Portal showed that most circadian genes I investigated are indeed expressed in skeletal muscle. However, even though expression levels were mostly low compared to other tissues, I caution that this does not mean that they are not important in muscle. The circadian genes can also exert influence on skeletal muscle indirectly, for example by changing hormone levels. Another potentially interesting tissue could be revealed in the form of sun exposed and not sun exposed skin. This could be important in the context of outdoor exercise.

3.4 Cell Secretion

To find out whether the circadian genes I investigate are predicted to be secreted proteins or extracellular matrix genes, I employed the Human Protein Atlas (proteinatlas.org 2018c) and data from Naba et al. (2016) to compare their gene lists for overlap with mine. For this, I used an online overlapping tool (BaRC 2018). When I compared my circadian gene list with the list of secreted genes provided by the Human Protein Atlas, I found that only two of them were predicted to be secreted from the cell: *ARNTL* and *CRY2* (Table 10).

Table 10: Secreted circadian genes and circadian extracellular matrix genes

	Number of genes	Gene symbols
Secreted genes	2	*ARNTL*, *CRY2*
ECM[a] genes	0	

[a] extracellular matrix

The Human Protein Atlas reports that *ARNTL* is mainly localized to the nucleoplasm (Figure 9). In addition, it is localized to vesicles. However, evidence for the latter is uncertain (proteinatlas.org 2018a).

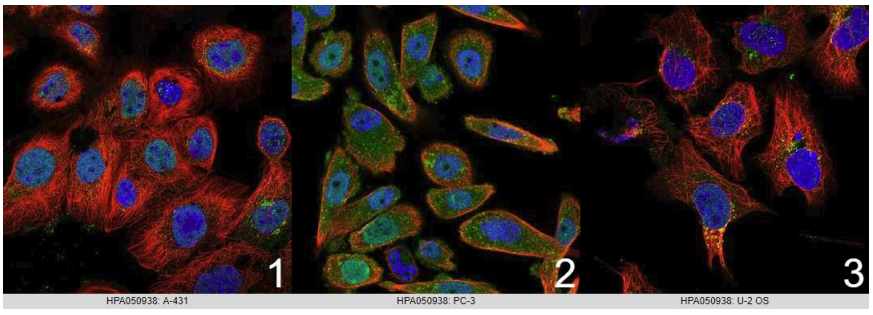


Figure 9: *ARNTL* localization in different human cell lines
Figure provided by proteinatlas.org (2018a)

I could not find current scientific literature concerned with the role of secreted *ARNTL*. In the case of *CRY2*, the Human Protein Atlas reports that *CRY2* is mainly localized to the nuclear speckles and additionally to the cytosol (Figure 10). Both assumptions are supported by evidence (proteinatlas.org 2018b).

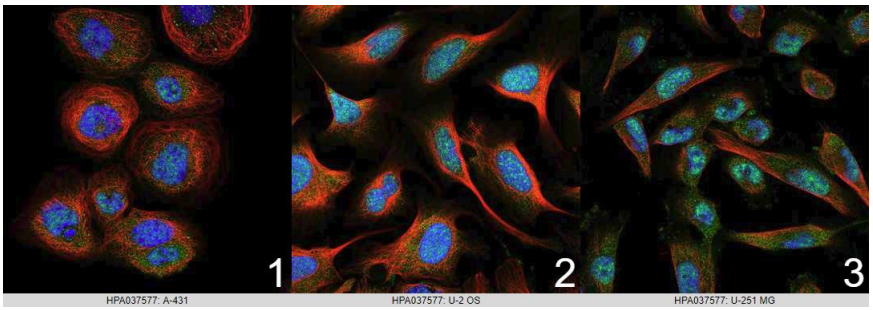


Figure 10: *CRY2* localization in different human cell lines
Figure provided by proteinatlas.org (2018b)

There are current scientific studies on the role of *CRY2* in the Arabidopsis plant, but I could find none about *CRY2* localization or secretion in humans. This would certainly be an interesting research topic for the future.

When I compared my circadian gene list with the list of extracellular matrix genes, I found that none of them were found to be part of the extracellular matrix. However, there is current scientific literature that shows while circadian genes may not be part of the extracellular matrix, they do interact with it.

Dityatev, Schachner, and Sonderegger (2010) investigated the role of the extracellular matrix in synaptic plasticity and homeostasis and found that it plays a significant role as "a promoter of structural and functional plasticity and as a degradable stabilizer of neural microcircuits" (Dityatev, Schachner, and Sonderegger 2010). In regard to circadian genes, they report circadian oscillations in the retina that allow the visual system to adjust to daily changes in ambient illumination (Dityatev, Schachner, and Sonderegger 2010). Mengatto et al. (2011) investigated cartilage extracellular matrix genes in the context of osseointegration, i.e. dental and orthopedic implants, which require the integration of bone and tissue. In their study, they found that "the placement of the implant into bone alters the expression of genes involved in circadian regulation" (Mengatto et al. 2011) and thus hypothesize that "the circadian rhythm system and cartilage extracellular matrix may be involved in the establishment of osseointegration under vitamin D regulation" (Mengatto et al. 2011).

In summary, my analysis showed that only two circadian genes are secreted (*ARNTL* and *CRY*). However, current scientific literature about this is lacking. No circadian gene was present in the extracellular matrix, but scientific literature is currently discussing indirect influences of circadian genes in the context of osseointegration.

3.5 OMIM

To find out more about the current state of scientific literature concerning the circadian genes of my corpus, i.e. what is known about the genes, especially in relation to disease, I consulted the Online Mendelian Inheritance in Man (OMIM) website. It is a freely available online compendium of human genes and genetic phenotypes (omim.org 2018a).

For reasons of clarity and brevity, I will go through the genes of my corpus in an alphabetic order and limit citation and discussion to recent findings with possible relation to circadian rhythms and exercise.

ARNTL OMIM states in the animal model section of *ARNTL*, that

> "Liu et al. (2007) showed that PGC1-alpha (604517), a transcriptional coactivator that regulates energy metabolism, is rhythmically expressed in the liver and skeletal muscle of mice. PGC1-alpha stimulates the expression of clock genes, notably Bma11 and Rev-erb-alpha (602408), through coactivation of the ROR family of orphan nuclear receptors. Liu et al. (2007) concluded that they identified PGC1-alpha as a key component of the circadian oscillator that integrates the mammalian clock and energy metabolism" (omim.org 2018b).

This finding highlights a direct connection between an important gene of endurance exercise pathways, *PPARGC1A*, and a core component of the clock, *ARNTL*. I speculate that it could provide valuable insights regarding the timing of exercise interventions to improve treatment outcomes in diabetes and cancer treatments.

BHLHE40 The animal model section of *BHLHE40* on OMIM reports that

> "Vercherat et al. (2009) demonstrated that Stra13 [BHLHE40] protected muscle cells from oxidative damage and that its absence led to muscle necrosis in response to injury in Stra13-deficient mice. The Stra13-null mutants expressed elevated levels of TNF-alpha (191160), reduced levels of heme oxygenase-1 (HMOX1; 141250), and displayed apparent signs of oxidative stress prior to muscle death. Moreover, Stra13-null muscle cells exhibited an increased sensitivity to prooxidants, and conversely, Stra13 overexpression provided resistance to oxidative damage. Treatment with antioxidant N-acetylcysteine ameliorated muscle necrosis in Stra13-null mice. Stra13 expression was elevated in muscles from dystrophin (DMD; 300377)-deficient (mdx) mice, and mdx/Stra13-null double mutants exhibited an early onset of muscle degeneration. Vercherat et al. (2009) noted that the findings revealed the contribution of Stra13 in maintenance of muscle integrity" (omim.org 2018c).

If *BHLHE40* has a part in maintaining muscle mass in the face of oxidative damage, its circadian pattern of expression could potentially be used to improve the efficacy of exercise and medicative interventions in the treatment of muscle ageing.

CLOCK The gene function section of *CLOCK* on OMIM shows:

> "Rutter et al. (2001) demonstrated that the DNA binding activity of the Clock:BMAL1 and NPAS2:BMAL1 heterodimers is regulated by the redox state of nicotinamide adenine dinucleotide (NAD) cofactors in a purified system. The reduced forms of the redox cofactors, NAD(H) and NADP(H), strongly enhance DNA binding of the Clock:BMAL1 and NPAS2:BMAL1 heterodimers, whereas the oxidized forms inhibit. Rutter et al. (2001) suggested the possibility that food, neuronal activity, or both may entrain the circadian clock by direct modulation of cellular redox state" (omim.org 2018d).

If it is possible to entrain the circadian clock directly by timing food and neuronal activity, the efficacy of diet and exercise interventions could be improved by appropriate circadian timing.

CRY1 The molecular genetics section of *CRY1* on OMIM states that

> "In affected members of 7 unrelated families, mostly of Turkish origin, with delayed sleep phase disorder (DSPD; 614163), Patke et al. (2017) identified a splice site mutation in the CRY1 gene (601933.0001). The mutation in the first family was found by a combination of candidate gene and whole-exome sequencing. The variant was found at a frequency of 0.6% in databases of human genetic variation: minor allele frequency of 0.0012 in the 1000 Genomes Project and 0.004335 in the ExAC databases; this frequency lies within the reported range of DSPD prevalence in the general population. In vitro functional expression studies showed that the mutant CRY1 increased circadian period via a gain of function. The mutant protein showed increased localization to the nucleus compared to wildtype, and had increased interaction with its target transcription factors *CLOCK* and *ARNTL*, resulting in increased transcriptional inhibition" (omim.org 2018e).

If *CRY1* has a role in the genetics of sleep, exercise and medication it may possibly be used to relieve symptoms of diseases like delayed sleep phase disorder or even improve the sleep of healthy individuals.

CSNK1D The animal model section of *CSNK1D* shows that

> "Zhou et al. (2010) found that mice with targeted overexpression of the *Csnk1d* gene in the forebrain and striatum exhibited hyperactivity, decreased anxiety, increased impulsivity, and defects in nesting behavior compared to wildtype mice. Mutant mice also showed paradoxical responses to dopamine receptor stimulation, showing hypoactivity following injection of amphetamine or methylphenidate and certain dopamine agonists. *Csnk1d* overexpression was also associated with downregulation of DRD1 (126449) and DRD2 (126450) receptor levels. The behavioral characteristics in these mice were reminiscent of the symptoms and drug responses observed in attention deficit-hyperactivity disorder (ADHD; 143465) in humans" (omim.org 2018f).

Although highly speculative, I propose further investigation of this connection might reveal possible uses of well-timed exercise interventions to improve symptoms in individuals experiencing symptoms of ADHD or problems with drug abuse.

NR1D1* and *NR1D2 The gene function sections of *NR1D1* and *NR1D2* report that

> "Sulli et al. (2018) showed that 2 agonists of REV-ERBs, SR9009 and SR9011, are specifically lethal to cancer cells and oncogene-induced senescent cells, including melanocytic nevi, and have no effect on the viability of normal cells or tissues. The anticancer activity of SR9009 and SR9011 affects a number of oncogenic drivers such as HRAS (190020), BRAF (164757), PIK3CA (171834), and others, and persists in the absence of p53 and under hypoxic conditions. The regulation of autophagy and de novo lipogenesis by SR9009 and SR9011 has a

> critical role in evoking an apoptotic response in malignant cells. Notably, the selective anticancer properties of these REV-ERB agonists impair glioblastoma growth in vivo and improve survival without causing overt toxicity in mice. Sulli et al. (2018) concluded that pharmacologic modulation of circadian regulators is an effective antitumor strategy, identifying a class of anticancer agents with a wide therapeutic window. Sulli et al. (2018) proposed that REV-ERB agonists are inhibitors of autophagy and de novo lipogenesis, with selective activity towards malignant and benign neoplasms" (omim.org 2018g).

This very recent research warrants further investigation about whether not only *NR1D1* and *NR1D2* antagonists in the form of drugs, but also exercise interventions with appropriate circadian timing could be used to fight cancer.

NFIL3 The gene function section of *NFIL3* states that

> "Wang et al. (2017) showed that intestinal microbiota regulates body composition through the circadian transcription factor NFIL3. Nfil3 transcription oscillates diurnally in intestinal epithelial cells, and the amplitude of the circadian oscillation is controlled by the microbiota through group 3 innate lymphoid cells, STAT3 (102582), and the epithelial cell circadian clock. NFIL3 controls expression of a circadian lipid metabolic program and regulates lipid absorption and export in intestinal epithelial cells. Wang et al. (2017) concluded that their findings provided mechanistic insight into how the intestinal microbiota regulates body composition and established NFIL3 as an essential molecular link among the microbiota, the circadian clock, and host metabolism" (omim.org 2018h).

The connection of circadian rhythm, metabolism and intestinal microbiota by *NFIL3* could prove a very interesting field for further study, especially regarding optimal circadian timing of diet and exercise interventions targeting this gene.

PER3 The molecular genetics section of *PER3* shows that

> "In 3 affected members of a family with familial advanced sleep phase syndrome-3 (FASPS3; 616882) and features of seasonal affective disorder (see 608516), Zhang et al. (2016) identified heterozygosity for 2 missense mutations in the PER3 gene on the same allele (P415A and H417R; 603427.0001). In vitro functional expression studies showed that the variant allele resulted in decreased protein expression and a loss of PER3 repressor activity. Transgenic mice expressing the variant allele showed altered phases of the sleep-wake cycle and depressive-like behavior similar to the patients with FASPS3. The findings suggested that PER3 plays a role in sleep and mood regulation, especially in response to seasonal changes in day length" (omim.org 2018i).

The connection between sleep and performance is an important and relevant finding. A well-timed exercise or drug intervention might be a possible solution for individuals suffering of familial advanced sleep phase syndrome or seasonal mood affective disorder.

PPARGC1B In the gene function section of *PPARGC1B*, OMIM reports that

> "Sahin et al. (2011) used transcriptomic network analyses in mice null for either Tert (187270) or Terc (602322), which exhibit telomere dysfunction, to identify common mechanisms operative in hematopoietic stem cells, heart, and liver. Their studies revealed profound repression of peroxisome proliferator-activated receptor-gamma (PPARG; 601487), PCG1-alpha (604517) and PGC1-beta, and the downstream network. Consistent with PGCs as master regulators of mitochondrial physiology and metabolism, telomere dysfunction was associated with impaired mitochondrial biogenesis and function, decreased gluconeogenesis, cardiomyopathy, and increased reactive oxygen species. In the setting of telomere dysfunction, enforced Tert or PGC1-alpha expression or germline deletion of p53 (191170) substantially restored PGC network expression, mitochondrial respiration, cardiac function, and gluconeogenesis. Sahin et al. (2011) demonstrated that telomere dysfunction activates p53 which in turn binds and represses PGC1-alpha and PGC1-beta promoters, thereby forging a direct link between telomere and mitochondrial biology. Sahin et al. (2011) proposed that this telomere-p53-PGC axis contributes to organ and metabolic failure and to diminishing organismal fitness in the setting of telomere dysfunction" (omim.org 2018j).

This link between PGC pathways, metabolism and longevity could warrant further investigation regarding possible exercise intervention (i.e. endurance training) timing to improve health and longevity.

In summary, recent literature found on OMIM shows a host of interesting connections of circadian genes with a wide span of topics relevant for staying healthy and performing well at exercise:

- Endurance Pathways
- Metabolism Pathways
- Muscle Ageing
- Sleep Disorders
- ADHD and drug abuse
- Psychological Disorders / Mood Regulation
- Gut Microbia
- Telomere Length and Longevity

3.6 Exercise and Gene Expression

To examine whether and how the circadian genes I study change their expression after endurance and strength exercise, I consulted the paper "Simplified data access on human skeletal muscle transcriptome responses to differentiated exercise" and its respective simplified access excel data sheet, published by Vissing and Schjerling (2014). They "investigated the human skeletal muscle transcriptome responses to differentiated exercise and non-exercise control intervention" and developed a "straightforward search tool to allow for easy access and interpretation" (Vissing and Schjerling 2014).

Table 11: Number of genes regulated per condition and relative to whole corpus

	2.5h	5h	2.5h – % of corpus	5h – % of corpus
regulated by any exercise	25	28	81%	90%
upregulated by RT	7	8	23%	26%
upregulated by ET	9	6	29%	19%
downregulated by RT	7	9	23%	29%
downregulated by ET	2	5	6%	16%

only significant changes, no data for *FABP9*

After careful analysis of the data, I can conclude that most of the genes in the corpus at hand were indeed significantly influenced by resistance or endurance exercise (with one gene not covered by the data) (Table 11).

- Depending on time frame, up to 90% of the genes investigated in this thesis were regulated by exercise in general
- Up to 26% were upregulated and up to 23% were downregulated by resistance training
- Up to 29% were upregulated and up to 16% were downregulated by endurance training

It is worth noting, that *ARNTL* was significantly upregulated by resistance training as well as endurance training with fold changes up to 3.63 after resistance training (Figure 11). Concurrent with this finding, Schiaffino, Blaauw, and Dyar (2016) report that "[*ARNTL*] circadian oscillation in the skeletal muscle is involved in adult muscle metabolism". They further hypothesize a "major role of the muscle clock in anticipating the sleep-to-wake transition, when glucose becomes the predominant fuel for the skeletal muscle" (Schiaffino, Blaauw, and Dyar 2016). The details as to how resistance training could influence *ARNTL* expression for circadian time-setting warrants further research. While *ARNTL* was clearly influenced by exercise, *CLOCK* was not influenced significantly. It is a noteworthy finding that only one component of the positive arm is impacted by exercise, while the other is not.

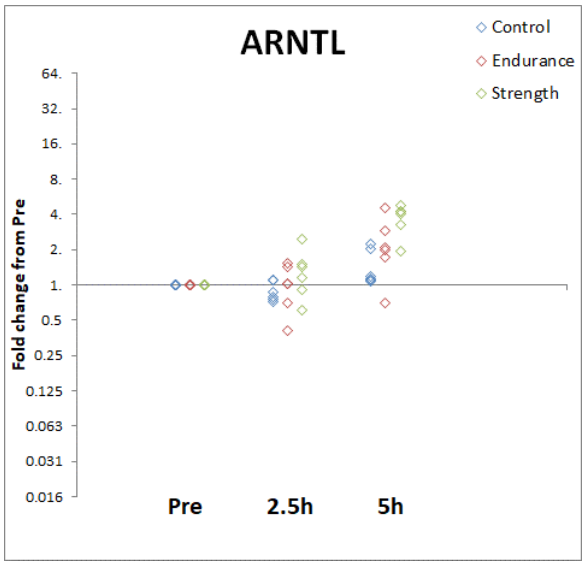


Figure 11: Skeletal muscle expression of *ARNTL* before and 2.5h and 5h after human strength (resistance) and endurance exercise

The components of the negative arm of the core circadian oscillator were mostly downregulated by both resistance and endurance exercise, except for *CRY1*, which was upregulated by both forms of training (Figure 12). This is an important finding since Jordan et al. (2017) reported that "CRY1 and CRY2 modulate exercise physiology by altering the activity of several transcription factors, including CLOCK/BMAL1 and PPARδ, and thereby alter energy storage and substrate selection for energy production".

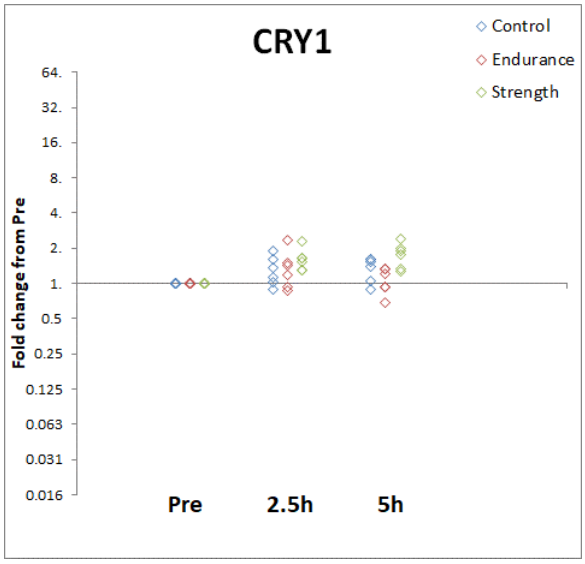
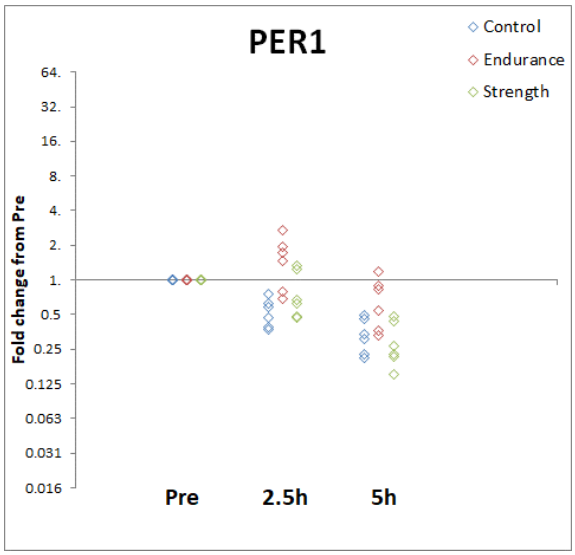
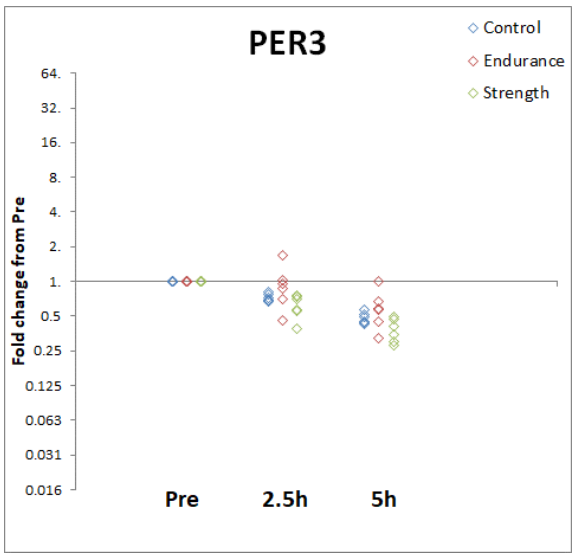


Figure 12: Skeletal muscle expression of *CRY1* before and 2.5h and 5h after human strength (resistance) and endurance exercise

The *PER* genes were especially downregulated with fold changes ranging up to -3.66 in *PER1* 5h after resistance training (Figure 13a). Downregulation of the negative arm via resistance training could thus be important for circadian clock setting. *PER3*, an important influence on most elements of the core oscillator, was significantly downregulated by both forms of training with downregulation up to a fold change of -2.67 5h after resistance training (Figure 13b).



(a) Skeletal muscle expression of *PER1* before and 2.5h and 5h after human strength (resistance) and endurance exercise



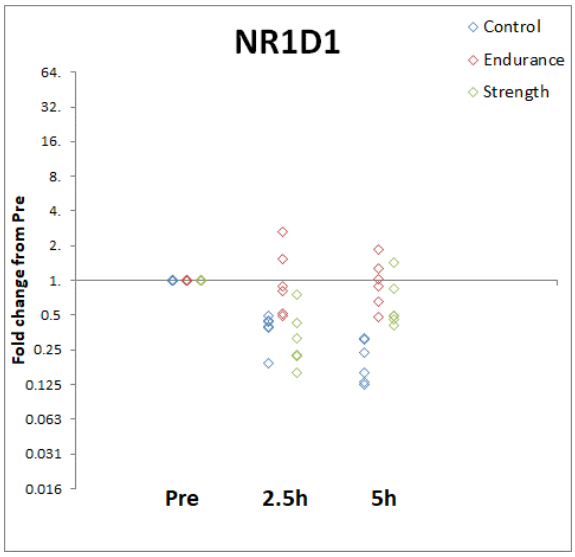
(b) Skeletal muscle expression of *PER3* before and 2.5 h and 5 h after human strength (resistance) and endurance exercise

Figure 13: Skeletal muscle expressions of *PER1* and *PER3* in the context of exercise

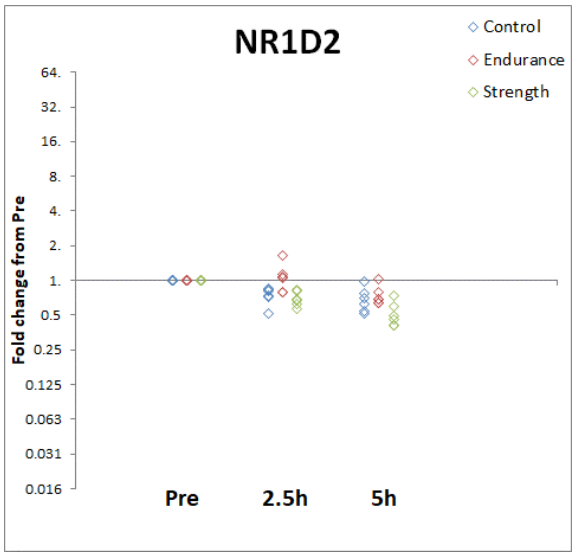
It seems that the components of the core circadian oscillators are indeed influenced by exercise, especially resistance exercise. This finding was also established by Zambon et al. (2003), when they concluded that "[t]he coordinated regulation of the circadian clock genes Cry1, Per2, and [...] suggest that RE [resistance exercise] may directly modulate circadian rhythms in human skeletal muscle".

The data provided by Vissing and Schjerling (2014) also show that circadian genes, which are known to regulate lipid metabolism, are also influenced by exercise. *NR1D1* is known for its role as "an integrator of circadian rhythms and metabolism" and seems to "behave as a gatekeeper to timely coordinate the circadian metabolic response" (Duez and Staels 2009). It is also shown to be strongly downregulated only by resistance exercise in the data, showing significant fold changes up to -3.29 at 2.5h after resistance

training (Figure 14a).



(a) Skeletal muscle expression of *NR1D1* before and 2.5h and 5h after human strength (resistance) and endurance exercise



(b) Skeletal muscle expression of *NR1D2* before and 2.5h and 5h after human strength (resistance) and endurance exercise

Figure 14: Skeletal muscle expressions of *NR1D1* and *NR1D2* in the context of exercise

The same is true for *NR1D2*, which is suspected “to form an accessory feedback loop that contributes to clock function” (Cho et al. 2012). It was downregulated significantly by a fold change up to -1.98 at 5h after resistance training (Figure 14b). However, it was significantly downregulated 5h after endurance training, challenging the singular role of resistance training. In general, this could mean that resistance training, or exercise in general, is a possible means of influencing circadian time of metabolism.

Another gene, *BHLHE40*, is upregulated 2.5h post endurance exercise (fold change 1,41) and downregulated 5h post resistance exercise (fold change -1.36) (Figure 15). This is an interesting result, because Lundberg et al. (2016) also reported upregulation of *BHLHE40* after resistance training that took place

after a bout of endurance training, which resulted in a fold change of 2.05.

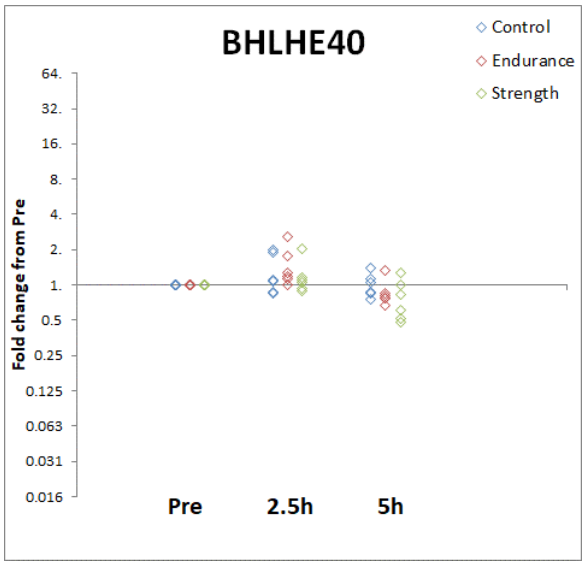


Figure 15: Skeletal muscle expression of *BHLHE40* before and 2.5h and 5h after human strength (resistance) and endurance exercise

Hsiao et al. (2009) found that "Bhlhe40 binds to the promoters of PGC-1α and myogenic genes in vivo and that Bhlhe40 represses the MyoD-mediated transactivation of these promoters", hypothesizing that "Bhlhe40 functions as a repressor of MyoD". I speculate that *BHLHE40* upregulation is part of one of the negative pathways suppressing protein biosynthesis if activated by endurance exercise. One might possibly increase muscle hypertrophy by matching resistance training timing with expression lows of *BHLHE40*. However, results of a recent study by Chung et al. (2015) also "suggest that Bhlhe40 is a novel regulator of PGC-1α activity repressing oxidative metabolism gene expression and mitochondrion biogenesis". They "further found that Bhlhe40-mediated repression can be largely relieved by exercise, in which its recruitment to PGC-1α -targeted cis elements was significantly reduced" (Chung et al. 2015). This could be a challenging new research question for further studies in the future.

In summary, I was able to identify that most of the circadian genes I investigate can be influenced by resistance and endurance exercise. My analysis showed multiple circadian genes as possible targets for exercise interventions, such as members of the core circadian oscillator as well as the circadian rhythm integrator *NR1D1* and the strength and endurance pathway modulator *BHLHE40*.

3.7 Exercise and Phosphoproteomics

I turned to two studies conducted by Hoffman et al. (2015) and Potts et al. (2017) to find out whether the proteins encoded by the genes in my corpus become phosphorylated in skeletal muscle or other tissues in response to exercise. Both studies provided simplified data access via an excel spreadsheet, which I used to examine whether proteins encoded by circadian genes of my corpus are phosphorylated either by endurance or resistance exercise. To get a clearer view of the data, I applied a coloured formatting to it, creating a color scale indicating increase (green) and decrease (red) of phosphorylation (Table 12 and

Table 13).

Table 12: Phosphoproteomic analysis of the effects of electrically evoked maximal-intensity contractions on circadian protein phosphorylation in mouse skeletal muscle

Gene symbol	Log2 (contraction control)[a]	p-value
Csnk1e	-0.17	0.27
Plin4	-0.109	0.80
Plin4	-0.171	0.69
Plin4	-0.470	0.16
Plin4	-0.481	0.06

phosphorylation sites were omitted for the sake of a comprehensive view, for exact phosphorylation sites see supplemental data

[a] negative values were highlighted using a color formatting

Hoffman et al. (2015) indeed measured phosphorylation of two proteins expressed by circadian genes: *CSNK1E* was phosphorylated at one site with a change of -0.17, while *PLIN4* was phosphorylated at four sites with changes up to -0.48 (Table 12). However, none of these measurements reached significance and differences to the control group were not always high. Potts et al. (2017) found phosphorylation of proteins expressed by the following circadian genes after high intensity resistance training (Table 13):

- *PLIN4* was phosphorylated at 14 sites, fold change 0.7 to -0.4
- *HNRNPU* was phosphorylated at one site, fold change -0.05
- *DBT* was phosphorylated at one site, fold change -0.12
- *CSNK1E* was phosphorylated at one site, fold change -0.06
- *CSNK1D* was phosphorylated at four sites, fold change -0.06 to -0.36

Table 13: Phosphoproteomic analysis of human skeletal muscle biopsies from untrained healthy males before and after a single high-intensity exercise bout

Gene symbol	Log2fold change[a]	p-value
PLIN4	0.406	na
PLIN4	-0.017	na
PLIN4	0.323	na
PLIN4	0.162	na
PLIN4	-0.431	na
PLIN4	0.718	na
PLIN4	0.036	0.918
PLIN4	0.228	na
PLIN4	0.120	na
PLIN4	-0.042	na
PLIN4	-0.045	na
PLIN4	-0.036	na
PLIN4	-0.132	na
PLIN4	0.204	na
HNRNPU	-0.051	na
DBT	-0.121	na
CSNK1E	-0.064	na
CSNK1D	-0.103	na
CSNK1D	-0.359	na
CSNK1D	-0.229	na
CSNK1D	-0.255	na

phosphorylation sites were omitted for the sake of a comprehensive view, for exact phosphorylation sites see supplemental data

[a]color formatting: green = positive; red = negative

However, while some changes to the control group were of note, no p-values were available for most of these measurements except one, which was not significant.

In the context of circadian gene phosphorylation, Takahashi et al. (2008) report:

> "Recent work shows that post-translational modification and degradation of circadian clock proteins are crucial steps for determining circadian periodicity of the clock. As seen previously in Drosophila, mammalian PeR1 and PeR2 proteins are progressively phosphorylated as they accumulate during the late afternoon and night".

CSNK1D and *CSNK1E* are part of the circadian clock system and is regarded as key kinase for *PER* and *CRY* phosphorylation. They have a key role in regulating circadian periodicity in mammals (Takahashi et al. 2008). Even though I could detect phosphorylation of both genes in the context of high intensity resistance and endurance training, there were no p-values given. Thus, I can only speculate that *CSNK1D* and *CSNK1E* might be manipulated by high intensity exercise to change periodicity of the clock. But this notion clearly needs more research.

PLIN4 is a circadian gene involved with lipid metabolism, which was phosphorylated after both endurance and resistance training. Dalen et al. (2004) uncovered that it is among the peroxisome proliferator–activated receptor-γ (PPAR-γ) target genes and that "[t]he direct regulation of perilipin and [PLIN4] by PPAR-γ therefore is likely to be an important mediator of the in vivo effects of prolonged treatment with PPAR-γ activators: insulin sensitization, fatty acid trapping in adipose tissue, reduced basal adipose lipolysis, and weight

gain". However, the data was once again not significant or lacked p-values. A significant role of exercise in phosphorylation of *PLIN4* and, following that assumption, an influence on lipid accumulation has still to be established. Chen et al. (2013) found that "inactivation of Plin4 downregulates Plin5 and reduces cardiac lipid accumulation in mice", but this conclusion was not reached by using an exercise intervention.
In summary, during my search for phosphorylated genes in the context of high intensity strength and endurance exercise, I could identify *PLIN4*, *CSNK1D* and *CSNK1E* as potentially interesting candidates for further examination. These genes were phosphorylated at multiple sites and could be targets for post-translational phosphorylation, and thus creating signals for adaptation to high intensity exercise. However, due to lack of data about significance, I can only recommend further research in this area.

3.8 Synergist Ablation

A study by Chaillou et al. (2013) was examined to find out whether the circadian genes belonging to the gene corpus at hand change their expression in the context of hypertrophy following mechanical overload. There, they subjected young mice to synergist ablation to induce muscle hypertrophy. This was done to identify signaling pathways that are active during hypertrophic response in young muscle.
I used the data provided by Chaillou et al. (2013) to conduct my own analysis of the expression of circadian genes after synergist ablation (Table 14). To get a clearer view of the data, again, I applied a coloured formatting to it, creating a color scale indicating increase (green) and decrease (red) of expression.

Table 14: Circadian gene expression after synergist ablation surgery

Gene symbol	% Day 0	% Day 1/ Day 0	% Day 3/ Day 0	% Day 5/ Day 0	% Day 7/ Day 0	% Day 10/ Day 0	% Day 14/ Day 0
Wee1	100	268	199	246	294	182	196
Nfil3	100	258	120	90	84	74	76
Bhlhe40	100	219	98	80	66	59	67
Per1	100	197	102	96	97	107	115
Per1	100	197	102	96	97	107	115
Myod1	100	157	239	215	181	141	93
Csnk1e	100	126	84	81	83	80	82
Csnk1d	100	118	122	121	121	112	106
Cry1	100	113	77	70	64	74	73
Cry2	100	113	60	65	68	86	96
Rora	100	112	50	48	46	54	76
Ces3a	100	110	48	22	14	17	19
Per2	100	108	74	73	87	83	103
Fabp9	100	106	82	68	75	71	79
Arntl	100	94	65	67	64	80	67
Idh1	100	86	129	142	158	135	129
Clock	100	83	91	90	93	88	94
Nr1d1	100	61	41	49	54	59	85
Pnpla3	100	60	60	51	54	67	70
Plin4	100	59	39	22	19	26	43
Nr1d1	100	57	42	41	33	57	66
Tef	100	56	51	55	55	71	88
Dgat2	100	53	33	27	22	35	44
Hlf	100	47	38	35	36	39	45
Pank1	100	45	28	32	32	35	49
Dbt	100	44	46	48	49	55	60
Ppargc1b	100	42	62	65	67	79	91
Per3	100	42	47	60	72	77	109
Ucp3	100	36	31	31	27	41	65

[a] relative to day 0 (%), a coloured formatting was applied to highlight positive (green) and negative (red) changes from baseline

Chaillou et al. (2013) described changes of the expression of the genes after synergist ablation by distinguishing between three distinct phases:

1. Early response: day 1
2. Intermediate response: days 3, 5 and 7
3. Late response: days 10 and 14

The expression of 13 genes of the corpus was increased during the early response to synergist ablation, but only two of them continued this trend until day 14 (*WEE1* and *CSNK1D*).

As I outlined in the previous section about phosphoproteomics, *CSNK1D* is part of the circadian clock system and are regarded as key kinases for *PER* and *CRY* phosphorylation. It has a key role in regulating circadian periodicity in mammals. It has, to my knowledge, not yet been implicated in hypertrophy signaling pathways. The moderate increase of expression of up to 22% might be a byproduct of other processes. Without further research, I can only hypothesize that it could be involved in the process hypertrophy due to mechanical overload.

During the cell cycle, *WEE1* "negatively regulates the entry into mitosis by catalyzing the inhibitory tyrosine phosphorylation of the Cdc2 protein" (Tang, Coleman, and Dunphy 1993). This finding is also supported by McGowan and Russell (1995), who found that "WEE1 kinase negatively regulates entry into mitosis". Thereby, it plays an antagonistic role in cell replication. However, Kellogg (2003) uncovered that "Wee1 is part of a cell-size checkpoint that prevents entry into mitosis before cells have reached a critical size", thus making it a repressor only until sufficient size is reached. Therefore, I conclude that the high increase in *WEE1* expression after synergist ablation could be due to its activity during muscle cell replication, where it ensures that the new muscle cells do not start mitosis until sufficient size is reached.

I could not find current scientific literature explaining the strong upregulation of *Nfil3* during the early response phase as well as its downregulation during the intermediate and late response phases. Its involvement in the immune response (omim.org 2018h) could potentially be relevant as a response to the ablation process rather than muscular hypertrophy. However, at this point, this is mere speculation.

While the expression *IDH1* is reduced at day 1, it starts to become more expressed at day 3 (beginning of intermediate response) and continues this trend until day 14 (late response). When I searched the literature for possible reasons for this, I could not find any studies regarding *IDH1* and hypertrophy.

From the current research one can only hypothesize about its role in cancer. In their study concerned with the role of mutated *IDH1* in glioma, Dimitrov et al. (2015) describe its normal physiological function as "formation of non-mitochondrial NADPH", which in turn "is thought to be an important mechanism for limiting cellular oxidative damage". They further explain: "The product of the IDH1 forward reaction, a-KG, is an intermediate in the tricarboxylic acid cycle (TCA) and is also involved in nitrogen transportation, oxidation reactions, and amino acid formation" (Dimitrov et al. 2015). Interestingly, they also note that "IDH1 regulates glucose-stimulated insulin secretion" (Dimitrov et al. 2015), making it possibly relevant for glucose uptake in muscle. The role of *IDH1* during muscle hypertrophy seems very under researched and needs further study to allow any conclusions.

BHLHE40 reacted with a high degree of upregulation but became downregulated during intermediate response and continued this trend into its late response (day 14). As already discussed in a previous chapter, *BHLHE40* seems to have an influence on muscle hypertrophy by repressing *MYOD1* (Hsiao et al. 2009) and that it is downregulated after endurance training (data by Vissing and Schjerling 2014). Other researchers found that it is upregulated after resistance training done after an endurance bout (Lundberg et al. 2016). The different expression levels of *BHLHE40* between early response (219% upregulation) and late response (downregulated to 67%) might provide an explanation for these contradictory findings. However, I could only speculate about how these different responses might work over time and rather conclude that this question warrants further research in the future.

MYOD1 is a powerful driver of myogenesis and a proven circadian gene by McCarthy et al. (2007). It is highly upregulated by synergist ablation during its early up to its late response, returning to baseline (and possibly below) at day 14 (late response). The early to intermediate responses are easily explained by the role of *MYOD1* as "muscle maker gene". Its return to baseline (and a little below) on day 14, however, is curious. I can only speculate that in interaction with myogenin and satellite cell proliferation (Adams, Haddad, and Baldwin 1999; Adams 2006), hypertrophy due to mechanical overloading was no longer

dependent on expression of *MYOD1*.

The expression of 13 other genes of the corpus was decreased from day 1 on until day 14 (early to late response), the most downregulated being *UCP3*, *PANK1* and *CES3a*. *UCP3* is part of a group of uncoupling proteins potentially involved in thermogenesis and reaction to oxidative stress induced for example by endurance exercise (Brand and Esteves 2005). Its strong downregulation until day 14 (late response) could potentially be explained by antagonistic mechanisms because of skeletal muscle hypertrophy. The same might be true for other genes involved with lipid metabolism:

- *PANK1*, which is involved in lipid metabolism as well (Zhou et al. 2001)
- *DGAT2*, which is responsible for the synthesis of triglycerides (Cases et al. 2001) and
- *CES3*, which is responsible for hydrolysis of triglycerides McCarthy et al. (2007)
- *PLIN4*, which coats nascent lipid droplets (Wolins et al. 2003)
- *PPARGC1B*, which is relevant for activation of medium chain acyl *CoA* dehydrogenase (*MCAD*) gene transcription McCarthy et al. (2007)

It would be interesting to see if other studies could validate this hypothesis.

PER3, an important influence on the core circadian oscillator, was downregulated in response to synergist ablation as well, but returns to baseline of expression at day 14 (end of late response). I could not find literature that could explain this finding.

In summary, my analysis showed that in response to hypertrophy following synergist ablation in mice, about half of the circadian genes investigated were upregulated, while the other half was downregulated. The upregulated genes consisted mostly of very different genes, including both members of the circadian clock system such as *CSNK1D* as well as *IDH1*, a gene frequently mutated in cancer, as well as *WEE1*, a gene involved in the cell cycle. As to be expected, the gene *MYOD1*, which are active in myogenesis, was also upregulated. The downregulated circadian genes were almost all metabolically active. I suspect antagonistic mechanisms of myogenesis to be responsible for this. The downregulation of *PER3*, a known modulator of the core circadian oscillator, remained without explanation and warrants further research.

3.9 Cardiotoxin Injury

Another factor to be considered regarding exercise and circadian genes is the response to muscle injury. To investigate this, I looked into the study “Genomic Profiling Reveals That Transient Adipogenic Activation Is a Hallmark of Mouse Models of Skeletal Muscle Regeneration” (Lukjanenko et al. 2013). There, Lukjanenko and colleagues induced muscle injury either by cardiotoxin or glycerol injection and mapped gene expression changes. I analyzed their data to find out more about how circadian genes react to cardiotoxin injury (Table 15).

Table 15: Changes of circadian genes in response to cardiotoxin-induced muscle injury

Gene symbol	% 3 day vs control
Csnk1d	161
Idh1	136
Acat2	132
Arntl	128
Bhlhe40	114
Csnk1e	100
Ppargc1b	98
Cry2	78
Ccrn4l	78
Clock	74
Per1	72
Rora	64
Cry1	61
Plin4	54
Nr1d2	49
Nr1d1	49
Tef	45
Pank1	45
Dbt	43
Hlf	39
Dgat2	35
Ucp3	34
Nfil3	32

color formatting: <100=red; >100=green

After examining the data, I found out that only five genes of my corpus were upregulated after cardiotoxin injury:

- *CSNK1D*
- *IDH1*
- *ACAT2*
- *ARNTL*
- *BHLHE40*

While *ARNTL* is part of the forward arm of the CCO and upregulates activity of other genes, in this case during muscle injury response, *CSNK1D* encodes a casein kinase active in eukaryotic cells (Kusuda et al. 1996)[1]. I speculate that these two genes positively regulate the rebuilding of muscle cells. However, the downregulation of *Clock* contrasts this. Further research will be needed to dissolve this contradiction.
IDH1 and *ACAT2*, on the other hand, are mostly concerned with muscle energy metabolism. They could be more expressed to either induce fatty regeneration as a response to muscle damage (Lukjanenko et al. 2013) or supply the remaining muscle with more energy to compensate for the injury.
BHLHE40 has been implicated in repressing *MYOD1* (Hsiao et al. 2009) as well as *PPARGC1A* (Chung et al. 2015). I speculate that in the case of muscle injury, it actively represses the *PPARGC1A* pathway to increase myogenesis driven by *MYOD1*. Due to the absence of date for *MYOD1*, this is, however, highly speculative.

[1] cited summary from https://www.omim.org/entry/600864

Looking at the other end of the spectrum, I found that 16 of the investigated genes were downregulated, 9 of them by more than 50%:

- *NR1D2*
- *NR1D1*
- *TEF*
- *PANK1*
- *DBT*
- *HLF*
- *DGAT2*
- *UCP3*
- *NFIL3*

I assume that, because most of these genes are involved in metabolism, they are suppressed by myogenic pathways triggered by the muscle injury.
HLF and *TEF* were implicated in cardiac hypertrophy in a study conducted by Wang et al. (2010). Together with the finding of Drolet et al. (1991), that *TEF* and *DBT* are “coexpressed in a pituitary cell line” during embryogenesis, thus influencing thyroid development, I can not see a clear connection of these genes to muscle injury response. Hopefully, further research can uncover this.
The strong downregulation of *NFIL3* is unexpected, since Kashiwada et al. (2011) found that it controls type 2 T helper cell cytokine expression. Because of its role in immune response, one would expect it to be more active in response to muscle injury. This apparent contradiction should be subjected to further research.
In summary, I found that after cardiotoxin muscle injury, a member of the forward arm (*ARNTL*) as well as a member of the circadian clock system (*CSNK1D*) were upregulated. I believe them to be involved in the upregulation of myogenic processes. However, research in this area is contradictory. Two muscle metabolism genes, *IDH1* and *ACAT2*, were also upregulated. I assume them to be part of fatty regeneration processes as described by Lukjanenko et al. (2013). *BHLHE40*, a known modulator of both *MYOD1* and *PPARGC1A* pathways, was also upregulated. Even though this is highly speculative due to lack of data about *MYOD1*, I believe it to suppress endurance pathways to enhance myogenesis. The genes that were downregulated are mostly metabolically active. I think that they are suppressed by myogenic pathways triggered by the muscle injury.

3.10 Berlin High vs Berlin Low

To further illuminate the impact of the circadian genes within my gene corpus on muscle mass, I studied the data collected in the study “Genetic and genomic analyses of musculoskeletal differences between BEH” (Lionikas et al. 2013). There, the authors analyzed the transcriptome of two distinct breeds of mice: The Berlin high (BEH) and Berlin low (BEL), who were “selected for divergent growth and differ threefold

in body weight" (Lionikas et al. 2013). The Berlin high strain is bred for high body size and muscle mass and is homozygous for the murine myostatin mutation known as compact allele. The Berlin low stain was selected for small body size and muscle mass, allowing me to see whether the genes in the corpus are differently expressed between those two. By analyzing the data Lionikas et al. (2013) collected about circadian rhythm genes, I was be able to see how differently they are expressed between the two strains.

Table 16: Musculoskeletal differences between BEH and BEL strains regarding circadian gene expression

Gene symbol	BEH[a]	BEL[b]	DE-Seq[c] (adjusted p-value)	% difference[d] (BEH vs BEL)	expressed in gastrocnemius?[e]
Ces3a	1	0	1.00	nv	no
Nrd1d	6417	1220	**0.00**	526	yes
Arntl	2097	650	**0.01**	323	yes
Bhlhe40	5505	2643	**0.00**	208	yes
Clock	3203	1659	**0.00**	193	yes
Pparc1b	465	338	0.35	138	yes
Acat2	105	81	0.48	130	yes
Cry1	1142	952	0.40	120	yes
Pank1	1026	873	0.94	117	yes
Myod1	1119	1033	1.00	108	yes
Fabp9	1	1	1.00	101	no
Csnk1d	3537	3557	1.00	99	yes
Dbt	1740	1755	1.00	99	yes
Csnk1e	2142	2290	0.99	94	yes
Dgat2	1507	1861	0.99	81	yes
Nr1d2	3795	4909	0.19	77	yes
Nfil3	2092	2718	0.07	77	yes
Rora	5635	7429	**0.05**	76	yes
Pnpla3	747	1014	0.21	74	yes
Idh1	1394	1998	**0.00**	70	yes
Cry2	1486	2440	**0.00**	61	yes
Tef	10186	16901	**0.00**	60	yes
Ccrn4l	910	1557	**0.01**	58	yes
Plin4	15347	32042	**0.00**	48	yes
Ucp3	1925	4265	**0.04**	45	yes
Per1	2342	5336	**0.01**	44	yes
Hlf	371	1237	**0.00**	30	yes
Wee1	88	343	**0.00**	26	yes
Per3	440	2272	**0.00**	19	yes
Per2	209	2281	**0.00**	9	yes

[a] Berlin High

[b] Berlin Low

[c] significance was highlighted by bold formatting

[d] a coloured formatting was applied to highlight differences

[e] a coloured formatting was applied to highlight absence in gastrocnemius muscle

Looking at the data (Table 16), there were several circadian genes significantly ($p < 0.1$) expressed in BEH.

Table 17: Circadian genes most expressed in BEH strain

Gene symbol	p-value[a]	% Difference
Nr1d1	**0.00**	526
Arntl	**0.01**	323
Bhlhe40	**0.00**	208
Clock	**0.00**	193

[a] significant ($p < 0.05$) values highlighted by bold formatting

[b] BEH/BEL

It is interesting to note that in the BEH strain, both genes of the forward arm of the CCO are significantly more expressed than in BEL (Table 17). I hypothesize that the increased activity of *Arntl* and *Clock* was involved in creating the increased mass of the BEH strain by upregulating other, i.e. myogenic genes.

Nr1d1 on the other hand shows the highest difference regarding expression levels between the strains. Even though it was first found to be concerned with adipocyte differentiation (Chawla and Lazar 1993) we now know it is strongly influencing *Arntl* activity (Preitner et al. 2002). I conclude that it is most likely a major driver of forward arm activity during BEH myogenesis.

Bhlhe40 is a transcriptionally active gene that reacts to a variety of training stimuli, as discussed in previous chapters. I speculate that, while it may not play a direct role in BEHs increased hypertrophy, it could be interesting to further investigate its influence under exercise conditions.

Myod1 was surprisingly not significantly more expressed in BEH. Since samples were taken at 70 days of age, I assume that at this stage, *Myod1* activity was no longer the (only) driving force behind the maintenance of muscle mass.

On the other end of the spectrum were clock genes significantly ($p < 0.5$) more expressed in BEL (Table 18).

Table 18: Circadian genes most expressed in BEL strain

Gene symbol	p-value[a]	% difference[b]
Per2	**0.00**	91
Per3	**0.00**	81
Wee1	**0.00**	74
Hlf	**0.00**	70
Per1	**0.01**	56
Ucp3	**0.04**	55
Plin4	**0.00**	52
Ccrn4l	**0.01**	42
Tef	**0.00**	40
Cry2	**0.00**	39
Idh1	**0.00**	30
Rora	**0.05**	24
Nfil3	0.07	23

[a] significant ($p < 0.05$) values highlighted by bold formatting

[b] BEH/BEL

In the BEL strain, members of the negative arm (*Per1*, *Per2* and *Per3*) of the CCO are much more expressed than in BEH. Following this observation, I hypothesize that because these members of the

negative arm were more active in the BEL strain, forward arm activity during myogenesis was inhibited, preventing BEL from reaching the same amount of mass as BEH.

Apart from the *Per* genes, *Wee1* was also much more active in the BEL strain. As discussed in a previous chapter, *Wee1* is active during muscle cell replication, where it ensures that the new muscle cells do not start mitosis until sufficient size is reached. I speculate that in the case of BEL, increased *Wee1* activity suppressed myogenesis in many muscle cells even after sufficient size for mitosis was reached.

The two circadian PAR bZip transcription factors *Hlf* and *Tef* were also more expressed in BEL strain mice. Together with *Dbp* (which was not significantly different between the strains in this study), they were implicated in cardiac hypertrophy by Wang et al. (2010), who found “Cardiac hypertrophy, low blood pressure, and low aldosterone levels in mice devoid of the three circadian PAR bZip transcription factors DBP, HLF, and TEF”. I can only speculate that the increased expression level of *Hlf* and *Tef* in BEL mice might have suppressed myogenesis or prevented them from developing enough heart muscle mass to support more muscle. However, this assumption clearly needs more research.

Other genes more expressed in BEL (i.e. *Ucp3*, *Rora*, *Ccrn4l*, *Plin4*) are part of pathways regarding lipid metabolism and might also have a role in antagonistic mechanisms toward muscle hypertrophy. However, this speculative assumption needs further investigation and any causality cannot be supported purely by this data.

In summary, the BEH strained displayed a significantly stronger activity of both forward arm clock genes. A known driver of *Arntl* activity, *Nr1d1*, was also revealed to be much more active. *Bhlhe40*, a gene regulating both strength and endurance adaptation pathways, was likewise upregulated compared to BEL.

The BEL strain on the other hand showed increased activity of multiple members of the negative arm, *Per1* and *Per2*, as well as of a known negative arm regulator, *Per3*. Other metabolic genes were more active in BEL, too, leading me to conclude that more active metabolic genes might have led to reduced hypertrophy in BEL.

The analysis showed a pattern of more active forward arm genes in the more hypertrophied strain in contrast to more active negative arm genes in the less hypertrophied strain, accompanied by more active metabolic genes in the same less hypertrophied strain. This leads me to hypothesize that the core circadian oscillator might have contributed to the increased hypertrophy in BEH. However, this hypothesis needs further research to be reliable.

3.11 Type 1 versus Type 2b

Chemello and colleagues approached the topic of differential gene expression in single muscle fibers in their publication “Microgenomic Analysis in Skeletal Muscle: Expression Signatures of Individual Fast and Slow Myofibers” (Chemello et al. 2011). They investigated the differences in gene expression between type 1 and type 2b mouse muscle on the single fiber level, creating “the first wide catalogue of gene expression in type 1 and type 2B fibers”. I used this catalogue to investigate how the expression of the gene differs in-between type 1 and type 2B fibres in mice (Table 19).

Table 19: Differences in circadian gene expression between mouse type 1 and type 2b fibers

Gene symbol	logFC[a]	p-value[b]
Dgat2	2.64	**0.000**
Pank1	1.77	**0.000**
Per1	1.74	0.055
Cry2	0.95	0.343
Arntl	0.92	0.685
Nr1d2	0.91	0.408
Csnk1d	0.81	0.070
Nfil3	0.79	**0.027**
Hnrnpu	0.59	0.181
Plin4	0.54	0.433
Ccrn4l	0.54	0.304
Wee1	0.42	0.865
Fabp8	0.41	0.766
Clock	0.34	0.673
Idh1	0.18	0.802
Dbt	0.12	0.790
Per2	0.06	0.913
Cry1	-0.01	0.993
Csnk1e	-0.11	0.945
Per3	-0.12	0.823
Dbp	-0.26	0.662
Bhlhe40	-0.27	0.555
Myod1	-0.32	0.744
Rora	-0.51	0.798
Tef	-0.51	0.694
Ucp3	-2.15	0.265

[a] Positive logFC values indicate that the gene is more expressed in type 1 fibers, negative logFC values that the gene is more expressed in type 2b fibers

[b] significant ($p < 0.05$) values highlighted by bold formatting

Half of the genes of my corpus was more expressed in type 1 fibers (16 out of 32), while another eight genes were more expressed in type 2b fibers. Two genes (*PER2*, *CRY1*) were not differently expressed between the fibers (logFC$< |0.10|$) (Table 20).

Table 20: Numbers of circadian genes more expressed in either type 1 or type 2b mouse fiber

expression	number of genes	% of corpus
more expressed in Type 1	16	50.00%
more expressed in Type 2b	8	25.00%
no difference	2	6.25%
no data	6	18.75%

Of those genes more expressed in type 1 fiber, DGAT 2, *PANK1* and *PER1* were highly expressed in this fiber type (LogFC 2.64, 1.77 and 1.74 respectively), while only *UCP3* was highly expressed in type 2b fiber (LogFC -2.15). However, the p-values of most measurements were insignificant ($p>0.05$). Only the measurements of *DGAT2*, *PANK1*, and *NFIL3* reached significance. I will therefore limit further discussion to these genes.

DGAT2, which is most expressed in Type 1 fibers among my gene corpus, is responsible for the synthesis

of triglycerides (Cases et al. 2001). Thus, it does not defy expectation that this gene is more active in a slow twitch muscle fiber. *PANK1*, which is involved in lipid metabolism (Zhou et al. 2001) is also more expressed in slow type 1 fiber, following the same pattern as *DGAT2*. *NFIL3* is not part of muscle metabolism pathways but was implicated in the immune response by Kashiwada et al. (2011). It is unclear why it is more expressed in type 1 muscle fibers.

All in all, data indicates that about half of the circadian genes investigated were more expressed in type 1 slow muscle fiber, while about a quarter of them were more expressed in Type 2b muscle fast fiber. However, most measurements were statistically insignificant. Of the three significant measurements, two were of metabolically active genes more active in type 1 fiber. The other gene was also more active in type 1 fiber but is involved in immune response pathways.

3.12 Muscle young and old proteome

To further illuminate the role of circadian genes in muscle, I looked at two studies: one that analyzed the differences between gene expression between whole muscle and C2C12 myotubes and another that compared concentrations between young and old type 1 and 2A muscle fibers. The first study was conducted by Deshmukh et al. (2015) and I used the data to further analyze my own gene corpus regarding differences between expression levels in C2C12 myotubes and muscle tissue.

Table 21: Proteomic detection of circadian genes in triceps muscles of 15-week old female C57Bl/6 mice and 8 day-differentiated C2C12 myotubes

Gene symbol	Muscle 1	Muscle 2	Muscle 3	C2C12 1	C2C12 2	C2C12 3	% Difference
Per1	113.77	96.42	92.73	2.36	0.90	1.11	6926,97%
Ucp3	11.25	11.60	17.35	0.16	0.84	0.00	4002.39%
Plin4	345.27	334.96	345.84	62.31	6.75	26.37	1075.12%
Csnk1e	1.06	1.31	2.55	3.70	0.09	0.95	103.87%
Dbt	160.05	148.38	127.92	272.42	163.83	318.76	57.79%
Tefm	7.91	7.56	6.76	14.48	7.96	18.02	54.93%
Myod1	0.23	1.24	0.00	1.31	0.06	2.03	43.20%
Nr1d1	0.00	0.14	0.36	0.38	0.00	1.89	22.17%
Idh1	126.16	97.84	147.20	1052.54	1072.06	907.27	12.24%
Csnk1d	2.65	3.48	3.86	47.17	16.16	39.72	9.70%
Acat2; *Acat3*	20.02	12.50	9.79	468.05	436.78	483.48	3.05%
Cry2	0.00	0.00	0.05	1.09	0.00	1.55	1.96%
Arntl	0.06	0.00	0.00	2.72	0.02	5.73	0.70%
Pank1	0.01	0.00	0.00	2.47	1.21	4.83	0.10%
Ccrn4l	0.00	0.00	0.00	1.76	0.00	1.10	0.00%
Clock	0.00	0.00	0.00	2.23	0.00	1.10	0.00%
Cry1	0.00	0.00	0.00	0.44	0.00	0.74	0.00%
Nfil3	0.00	0.00	0.00	2.11	0.91	3.73	0.00%
Per2	0.00	0.00	0.00	17.87	2.91	6.90	0.00%
Wee1	0.00	0.00	0.00	0.94	0.14	0.00	0.00%
Bhlhe40	0.00	0.00	0.00	0.26	0.00	0.49	0.00%

There was no data for ten of the genes relevant for this thesis. Most of those genes that Deshmukh et al. (2015) could measure were more expressed in the C2C12 myotube model (15 out of 22), while a minority was much more expressed in the muscle model. Two genes were not expressed in either model. However,

of those genes only seven were of noteworthy different expression (Table 21).
Three genes were much more active in the muscle model:

- *Ucp3*
- *Per1*
- *Plin4*

Four genes were much more active in the C2C12 model:

- *Dbt*
- *Idh1*
- *Csnk1d*
- *Acat2*

Both *Ucp3* and *Plin4* are involved in metabolic pathways involved in either thermogenesis (Brand and Esteves 2005) or lipid metabolism (Wolins et al. 2003). It is not unexpected to see them more active in the muscle model, which encompasses muscle fibers within their tissue context. I could not find scientific literature to explain the strong presence of *Per1* in the muscle model.
We also see multiple metabolically active genes that are more active in the C2C12 model:

- *Dbt*, which is involved in the breakdown of the branched-chain amino acids (Friedrich et al. 2011)
- *Idh1* "regulates glucose-stimulated insulin secretion" and is concerned with the "formation of non-mitochondrial NADPH" (Dimitrov et al. 2015)
- *Csnk1d* "encodes an isoform of casein kinase I, which is a ubiquitous serine/threonine-specific protein kinase that constitutes most of the kinase activity in eukaryotic cells, where it is distributed in the nucleus, cytoplasm, and membrane fractions" (Kusuda et al. 1996)[2]
- *Acat2* "encodes cytosolic acetoacetyl-CoA thiolase (EC 2.3.1.9), which is important in the utilization of ketone bodies" (Groot et al. 1977)[3]

I believe that while the former set of genes (*Ucp3*, *Per1*, *Plin4*) is most active in the muscle periphery, the latter (*Dbt*, *Idh1*, *Csnk1d*, *Acat2*)operates more within the muscle fiber. The second study I analyzed was done by Murgia et al. (2017). I used their data to search for age differences regarding the circadian genes of my corpus.

[2]cited summary from https://www.omim.org/entry/600864
[3]cited from https://www.omim.org/entry/100678

Table 22: Proteomic detection of circadian gene expression in vastus lateralis in individual muscle fibers of younger (22–27 years) and older (65–75 years) male volunteers, absolute protein abundances

Gene symbol	slow younger	slow older	fast, 2B younger	fast, 2B older
Acat2	52.0	10.4	14.6	5.9
Dbt	52.6	79.9	23.9	33.2
Idh1	17.4	22.9	31.5	21.5
Plin4	289.3	354.4	279.4	431.9

Values were highlighted using a colored formatting to make differences more visible

Table 23: Proteomic detection of circadian gene expression in vastus lateralis in individual muscle fibers of younger (22-27 years) and older (65-75 years) male volunteers, relative values

Gene symbol	% Slow[a]	% Fast 2B[a]
ACAT2	498.85	249.33
DBT	65.80	71.82
IDH1	76.09	146.56
PLIN4	81.62	64.70

[a] young vs old

As you can see from the tables, only data for four of the genes could be provided (Table 22 and Table 23). Those four were:

- *ACAT2*
- *DBT*
- *IDH1*
- *PLIN4*

ACAT2 was much more expressed in young slow (5-fold higher expression) and young fast (2.5-fold higher expression) muscle. This is of note since, as explained before, it is involved in the utilization of ketone bodies (Groot et al. 1977). I speculate that younger muscle fiber could potentially make better use of ketone bodies than old fiber, making a diet low on carbohydrates more effective for younger athletes than older ones. At the same time, one could make use of the circadian pattern of *ACAT2* as well regarding exercise and diet interventions. Of course, this will need further research.

DBT was more expressed in old slow (25% higher expression) and old fast (29% higher expression) muscle fiber. A study done by Friedrich et al. (2011) showed that *DBT* breaks down branched-chain amino acids. This could be interesting in the context of sarcopenia of older muscle fibers. A potential link should be investigated in further studies.

IDH1 was more expressed in old slow and young fast muscle fibers. Since the gene is involved in supplying energy to our muscles, it could be interesting to further investigate its different expression levels in ageing muscle to find out more about how differently slow and fast muscle fibers age within the human body.

PLIN4 showed more expression in older muscle fiber for both types, it was more active in both old slow and old fast muscle fiber, the latter having a very strong percentage difference towards activity in older fiber. Interestingly, (Conte et al. 2013) found that a closely related gene of the PAT family, *Plin2*, is associated

with sarcopenia and muscle weakness. The fact that *Plin4* is also much more expressed in older fiber could suggest a similar role for *Plin4*. However, this needs more research.

In summary, my analysis of circadian gene activity between fiber and periphery showed that the circadian genes of my corpus could generally be found in both fiber and periphery. Most genes detected were metabolically active and could be responsible for energy metabolism both within the fibers themselves as well as between fiber and periphery. The finding that a gene of the negative arm of the CCO, *Per1*, was much more active in the muscle model, could not be explained and needs further research.

Regarding age differences between slow and fast muscle fibers, my analysis detected four circadian genes, *ACAT2*, *DBT*, *IDH1* and *PLIN4* that were very different in their respective activity patterns. These metabolically active genes were, in general, more active in older muscle fiber and in part already associated with sarcopenia and muscle ageing. However, *ACAT2* is the exemption, being more active in both younger slow and fast muscle fiber. Its role in the use of ketone bodies could be an interesting field of study in the context of low-carb diets and circadian rhythms in muscle.

3.13 CircaDB

CircaDB is a database of circadian transcriptional profiles from time course expression experiments from mice and humans (see circadb.org (2018a)). It was created in 2012 by Angel Pizarro, Katharina Hayer, Nicholas F. Lahens and John B. Hogenesch to "systematically collect, analyse and visualize circadian expression profiles for bench researchers in a simple and straightforward fashion" (Pizarro et al. 2012). I used CircaDB to find out if and how the genes I investigated are expressed in muscle in a circadian fashion. I restricted the search to skeletal muscle to maintain my focus on exercise and to be able to analyze more clearly and concisely. To cover all tissues in a gene set consisting completely of circadian genes would produce massive amounts of data I could not cover appropriately within the context of this thesis.

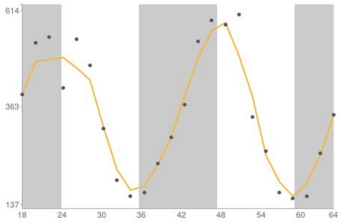
For my corpus of circadian genes, I could find data of circadian expression patterns in skeletal muscle for a total of 24 out of 32 genes. For eight genes there was no data about circadian expression in skeletal muscle. However, this does not necessarily mean that they are not expressed in a circadian pattern there. I analyzed expression over time in a qualitative, descriptive manner to contrast the activity patterns of the genes I research. Please note that this manner of analysis is severely limited due to subjective assessment, especially when it comes to blurry transitions between late day and early night or late night and morning. It is meant solely as an overview, not a precise representation of purely quantitative data.

I will now discuss my findings, beginning with qualitative description before moving on to literature comparison. Generally, I found that most genes peaked in skeletal muscle during the late hours of day and had their lowest point during the late hours of night (Table 24).

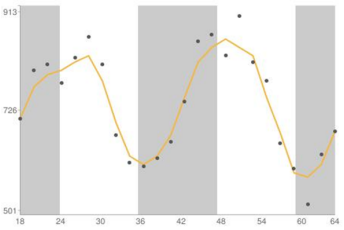
Table 24: Overview of Circadian Gene Peaks and Lows

	Number of genes	Gene symbols
Diurnal peak	14	
morning	4	*ARNTL*, *CLOCK*, *PANK1*, *UCP3*
midday	1	*NR1D1*
late	9	*CRY2*, *DBP*, *NR1D2*, *HLF*, *PER1*,*PER2*, *PER3*, *RORA*, *TEF*
Nocturnal peak	10	
early	4	*DBT*, *DGAT2*, *IDH1*, *WEE1*
midnight	3	*CES3*, *MYOD1*, *CCRN4L*
late	3	*CRY1*, *NFIL3*, *PNPLA3*
Diurnal low	14	
morning	7	*DBP*, *DBT*, *DGAT2*, *PER1*, *PER2*, *PER3*, *WEE1*
midday	4	*CRY1*, *MYOD1*, *CES3*, *CCRN4L*
late	3	*ARNTL*, *NFIL3*, *PNPLA3*
Nocturnal low	12	
early	2	*CLOCK*, *UCP3*
midnight	2	*NR1D1*, *PANK1*
late	8	*CRY2*, *DBP*, *HLF*, *IDH1*, *NR1D2*, *PER3*, *RORA*, *TEF*

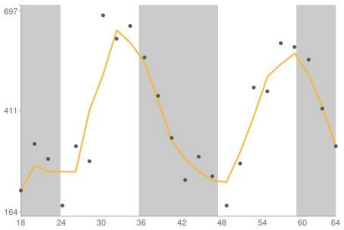
Both genes of the forward arm of the CCO, *Arntl* and *Clock* (Figures 16a and 16b), display a circadian pattern in muscle. The same is true for members of the negative arm, *Per1* (Figure 16c), *Per2*, *Cry1* and *Cry2*.



(a) *Arntl* expression in mouse 1.
OST skeletal muscle



(b) *Clock* expression in mouse 1.
OST skeletal muscle



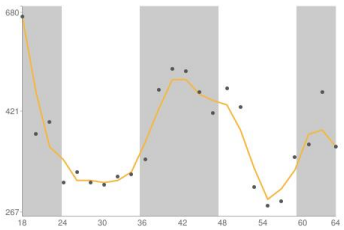
(c) *Per1* expression in mouse 1.
OST skeletal muscle

Figure 16: *Arntl*, *Clock* and *Per1* expression in mouse 1. OST skeletal muscle

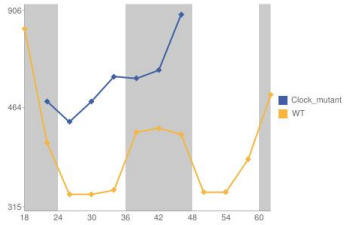
Consistent with scientific literature, the pattern of *Arntl* is opposite to that the *Per* genes (Shearman et al. 2000). Qualitative analysis clearly showed the opposite patterns of the forward and negative arms of the CCO: *ARNTL* and *CLOCK* peaked in the early morning and reached their low in the late hours of day or early hours of night, respectively. Opposite to this, most genes of the negative arm peaked during the later hours of day and were lowest in the morning.

Another important gene for exercise is the muscle maker gene *Myod1*. It was already reported as a circadian gene with peak expression during subjective night by McCarthy et al. (2007). It would be interesting to see how the circadian expression pattern of *Myod1* changes in the context of a diurnal resistance training intervention. This could allow researchers to find out how resistance training changes the circadian rhythm itself as well as how to optimally time resistance training sessions within the circadian rhythm.

As you can see in Figure 17b, a clock mutation in mice disrupts the circadian pattern of *Myod1*. I also found that the circadian expression of *Ucp3* peaks shortly after *Myod1*.



(a) *Myod1* expression in mouse 1. OST skeletal muscle



(b) *Myod1* expression in mouse wild type muscle

Figure 17: *Myod1* expression in mouse 1. OST skeletal muscle and in mouse wild type muscle

This was also observed by McCarthy et al. (2007):

> "Myod1 may be involved in the regulation of Ucp3 circadian expression given that Ucp3 peak expression occurs 8 h after peak Myod1 expression [...]. This notion is further supported by the finding that Myod1 is capable of activating Ucp3 transcription in a peroxisome proliferator-activated receptor-dependent manner" .

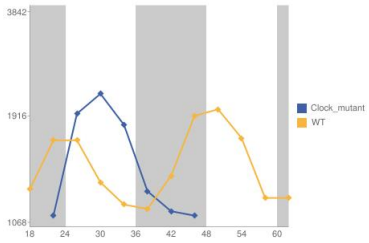


Figure 18: *Ucp3* expression in mouse wild type muscle

A clock mutation also disrupts the circadian pattern of *Ucp3* (Figure 18), which is an observation also shared by McCarthy et al. (2007):

> "Finally, in the Clock mutant, Ucp3 expression is dramatically phase shifted, suggesting circadian expression of Myod1 is required for proper phasing of Ucp3 peak expression (Fig. 4D). If Myod1 is involved in the circadian regulation of Ucp3 expression, this indicates that it is the phasic/peak expression of Myod1 that is critical as there was no difference in the average expression of Myod1 in muscle of wild-type and Clock mutant mice".

The notion that *Myod1* and *Ucp3* circadian expression is interdependent is also shared by other researchers like Solanes et al. (2003), who attest that “MyoD and PPAR-dependent pathways mediate human UCP3 gene regulation and that acetylase activity elicited by coregulators is implicated in the functional interaction between these regulatory pathways”.
In the context of circadian zeitgebers, Dudek and Meng (2014) report that in the case of *Myod1*, while “It was found that although the core clock components are unaffected in mice fasted for 24 h, the circadian rhythm of MyoD and Myf4 was disrupted and their expression was reduced” (Dudek and Meng 2014). Thus, the timing of meals could also play a key role in this interaction.
All in all, this connection could provide “a molecular mechanism for skeletal muscle specificity and fatty acid regulation of human UCP3 gene” (Solanes et al. 2003) in the context of diet and exercise interventions.

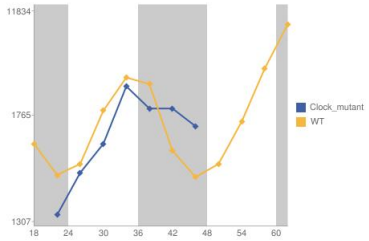


Figure 19: *Rora* expression in mouse wild type muscle

Interestingly, the circadian patterns of *Rora* (Figure 19) and *Myod1* (Figure 17) also seem to be related, as *Myod1* peaks only shortly before *Rora* reaches its minimum. This finding makes sense in the context of the findings of Downes, Carozzi, and Muscat (1995), who reported that “Rev-erbα functions as a negative regulator of myogenesis by targeting the expression of the myoD gene family”. Cross-talk between *RORA* and *NR1D1* was also established by Forman et al. (1994).
In summary, I found that in skeletal muscle, my selection of circadian genes mostly peaks during the late hours of day and is expressed lowest during the late hours of night. The opposite patterns of the forward and negative arm of the CCO could be replicated by qualitative analysis. My analysis also identified the genes *Myod1*, *Ucp3* and *Rora* as potentially linked targets for interventions regarding circadian expression in skeletal muscle. Interventions could include drug therapy for disturbed muscle metabolism as well as possible methods of optimizing concurrent strength and endurance exercise.

3.14 IMPC Mouse Phenotypes

The IMPC (see mousephenotype.org 2018) is a mass scale mouse phenotyping project, that plans to “expand the ongoing phenotyping effort from hundreds of mouse lines to thousands, with the ultimate aim of tackling the whole genome” (Brown and Moore 2012). I used their data to further analyze the genes of my corpus in relation to possible mouse mutant phenotypes.

Table 25: Changes in mouse phenotype due to circadian gene knockout

Phenotype changes	Number of genes	Gene symbols
preweaning lethality	3	*Acat2*, *Nr1d1*, *Wee1*
metabolism	4	*Bhlhe40*, *Nr1d1*, *Idh1*, *Nfil3*
body size/mass	2	*Acat2*, *Bhlhe40*
bone structure	2	*Bhlhe40*, *Per2*
sensory organs	2	*Nr1d1*, *Nfil3*
heart	2	*Nr1d1*, *Nr1d2*

Circadian gene knockout led to increased preweaning lethality and mostly negative changes of metabolism, body size or mass, bone structure, sensory organs or the heart (Table 25). There was no data for 19 genes of the corpus.

Most interestingly, many circadian genes involved with metabolism in a wider sense were affecting preweaning lethality. But due to lack of data about other circadian genes, I cannot make further conclusions about many genes that are part of the core oscillator or myogenic processes.

Ikonen (2006) reports that "ACAT2-deficient mice lack cholesteryl esters in the liver and small intestine [...]. They were protected from diet-induced hypercholesterolemia and cholesterol gallstones because of the reduced capacity to obtain cholesterol from the diet", making the implications about the role of *Acat2* inconsistent.

Bhlhe40 was implicated in both phenotype changes to metabolism, body size/mass and bone structure, highlighting its many roles in the genome of humans and mice. This is in line with current scientific literature connecting *Bhlhe40* with the *PPARGC1A* pathway (Chung et al. 2015) and organ microenvironments (Jarjour et al. 2018).

Nr1d1 was similarly involved in many phenotype changes regarding preweaning lethality, metabolism and changes to organs. Shen et al. (2017) also implicate it in organ-specific alterations by vertical sleeve gastrectomy, further supporting its role not only in metabolism but in shaping internal organs.

In summary, many circadian metabolic genes produce phenotype changes in mice resulting in mostly changes to metabolism, preweaning lethality and changes to body structure. Regarding exercise capacity and effects on performance, changes to body structure and metabolism or organs such as the heart could be factors that would make further investigation worthwhile. It could even be relevant in relation to the possible improvement of the timing of drug treatment of embryos during pregnancy.

3.15 GWAS Human Phenotypes

The GWAS Catalog (see ebi.ac.uk (2018b)) is a project by the National Human Genome Research Institute (NHGRI), which was created in 2008 to catalogue and summarize published genome-wide association studies. The catalog is freely available and provides data access to published SNP-trait associations. I searched the catalog for data about the genes in my corpus to find out more about diseases or other phenotypes that are SNPs in or near circadian genes associated with them. If there were associated diseases or phenotypes, I also checked how big the effects are.

Table 26: Reported traits of circadian gene SNPs[a]

Reported trait category	Number of genes	Gene symbols
metabolism	6	*ARNTL*, *CRY2*, *DGAT2*, *PANK1*, *PPARGC1B*, *RORA*
chronotype	5	*FABP9*, *HLF*, *MYOD1*, *PER2*, *PER3*
digestive tract disease	3	*NFIL3*, *PER3*, *TEF*
personality traits	3	*ARNTL*, *CLOCK*, *TEF*
cancer	2	*DBT*, *IDH1*
sleep	2	*FABP9*, *MYOD1*
immune system	2	*CSNK1D*, *NR1D1*
reproductive age	2	*ARTNL*, *RORA*
psychological disorders	1	*NFIL3*
intelligence	1	*NR1D2*
cognitive ability	1	*NR1D2*

[a] single nucleotide polymorphisms

There was no data for seven genes (*ACAT2*, *CES3*, *CSNK1E*, *CCRN4L*, *PER1*, *UCP3*, *WEE1*). Most reported traits for the rest of the genes were regarding changes to metabolism (6 genes), changes to chronotype (5 genes) and changes regarding digestive tract diseases (3 genes) (Table 26).

While the high incidence of metabolism trait changes and chronotype changes do not defy expectations of a gene corpus consisting of many metabolically relevant circadian genes, there are multiple interesting findings of gene SNPs leading to unexpected trait changes, some with very high effect scores.

A SNP of *ARNTL* was associated with a 7.545 z score decrease in the personality trait neuroticism, which had a Risk/Effect allele frequency (RAF) of 55,8%. This association was also found by Hill et al. (2017) but was left unexplained. Further research is clearly needed here. A SNP of *CLOCK* was also associated with changes to personality dimensions, however the effect was small (0.14 s.d. decrease), and even though it had a rather high RAF of 73%, I could not find further clarification for this. A rather high beta score increase of 8.723 z in Neuroticism was also reported for a SNP in *TEF*. RAF was not high with 30%. It will be interesting to see in further research how the positive arm of the CCO and the circadian gene *TEF* might influence personality.

Another gene that was associated with multiple unexpected trait changes is *NR1D2*. Known mainly for its involvement in circadian and metabolic pathways, multiple of its SNPs surprisingly are associated with trait changes to cognitive ability and intelligence. Changes range from a 6.016 z score increase to intelligence in one SNP to a 5.71 z score decrease to cognitive ability in another. It could therefore be beneficial to investigate precisely how this gene influences these traits. While Goto et al. (2017) already reported that "Nr1d1 was found to play a pivotal role in corticogenesis via regulation of excitatory neuron migration and synaptic network formation", implicating a related gene in the process of brain development and as a potential cause of autism spectrum disorders, I could not find current research on *NR1D2* in the context of cognitive ability and intelligence. A rather high odds ratio of 11.67 was also reported for one of its SNPs regarding the development of severe acne, highlighting that *NR1D2* fulfills several roles within the human genome.

The muscle maker gene *MYOD1* was associated with changes to night sleep phenotypes, albeit only with a beta score of a 0.09052-unit increase. While the circadian regulation of *MYOD1* is already being researched (Andrews et al. 2010), a potential role in sleep phenotype does not seem to have been investi-

gated yet. It would be interesting to see if this link could be used to influence sleep with resistance training interventions.
Another noteworthy finding was the association of multiple *PER3* SNPs with inflammatory bowel diseases. This was also described by Mazzoccoli et al. (2012), who reported:

> "Altered body rhythmicity and deregulated clock gene expression may cause circadian disruption, which can lead to immune dysregulation and chronic inflammatory diseases. PERIOD3 (PER3) polymorphisms have been associated with circadian disruption and changed secretion of cytokines involved in chronic inflammation" (Mazzoccoli et al. 2012).

Swanson, Burgess, and Keshavarzian (2011) suspect that sleep disturbances might be the reason for the development of colitis, reporting that "preliminary human studies have shown that patients with [inflammatory bowel disease] are at increased risk for altered sleep patterns" (Swanson, Burgess, and Keshavarzian 2011).
In conclusion, I found that SNPs of the circadian genes I investigated are mostly associated with changes to metabolism and chonotype. Other rather unexpected findings were associations with personality traits as well as inflammation-related bowel diseases.

3.16 EXAC LoF Tolerance

The Exome Aggregation Consortium (ExAC) (see broadinstitute.org (2018b)) provides an extensive dataset, allowing researchers to investigate "3,230 highly LoF-intolerant genes, 72% of which have no established human disease phenotype in the OMIM or ClinVar databases of observed human genetic mutations" (Lek et al. 2016). I used the data provided by ExAC to further research how the circadian genes of my corpus behave related to loss-of-function mutations.
For these mutations, ExAC assumes the following:

> "[T]here are three classes of genes with respect to tolerance to LoF variation: null (where LoF variation is completely tolerated), recessive (where heterozygous LoFs are tolerated), and haploinsufficient (where heterozygous LoFs are not tolerated). We used the observed and expected variants counts to determine the probability that a given gene is extremely intolerant of loss-of-function variation (falls into the third category). The closer pLI is to one, the more LoF intolerant the gene appears to be. We consider pLI $>=$ 0.9 as an extremely LoF intolerant set of genes" (broadinstitute.org 2018a).

Following this assumption, I divided the genes I examined into these three categories (Table 27)

Table 27: Broad categorization of loss-of-function intolerance of circadian genes

LoF[a] variance tolerance	Number of genes	Gene symbols	pLI[b]
null	13	*ACAT2*, *CES3*, *CRY1*, *DBT*, *DGAT2*, *FABP9*, *IDH1*, *MYOD1*, *PANK1*, *PER3*, *PLIN4*, *PNPLA3*, *UCP3*	$\text{pLI} = 0$
recessive	3	*NR1D2*, *CCRN4L*, *TEF*	$0 < \text{pLI} < 0.9$
haploinsufficient	16	*ARNTL*, *BHLHE40*, *CLOCK*, *CRY2*, *CSNK1D*, *CSNK1E*, *DBP*, *NR1D1*, *HLF*, *HNRNPU*, *NFIL3*, *PER1*, *PER2*, *PPARGC1B*, *RORA*, *WEE1*	$\text{pLI} > 0.9$

[a] loss-of-function

[b] probability of being LoF intolerant

Most of the gene corpus is either tolerant (null) or intolerant (haploinsufficient) to loss-of-function mutation. Only three genes were tolerant towards heterozygous loss of function mutation (recessive). Thus, it can generally be assumed that most circadian genes either serve functions necessary for life that can not be compensated by other genes, or they do not. Therefore, only the nonessential circadian genes seem to tolerate homozygous loss-of-function mutation.

Table 28: Overview over loss-of-function-intolerance of circadian genes

Gene symbol	Expected LoF[a]	Observed LoF[a]	pLI[b]
CLOCK	36.3	4	1.00
HNRNPU	30.4	1	1.00
WEE1	21.3	1	1.00
ARNTL	26.2	3	0.99
CRY2	23.5	2	0.99
CSNK1D	14.1	0	0.99
CSNK1E	14.7	1	0.97
NFIL3	9.1	0	0.95
NR1D1	17.8	2	0.95
RORA	27	4	0.95
BHLHE40	8.7	0	0.94
PER2	39.2	7	0.93
PPARGC1B	30.4	5	0.93
HLF	6.6	0	0.88
PER1	41.8	8	0.88
DBP	5.4	0	0.82
NR1D2	15.9	3	0.63
CCRN4L	10.3	2	0.52
TEF	5	1	0.42
ACAT2	12.1	9	0.00
CES3	21.2	15	0.00
CRY1	28.2	15	0.00
DBT	17.5	9	0.00
DGAT2	16.8	13	0.00
FABP9	4.6	5	0.00
IDH1	14.8	7	0.00
MYOD1	4.2	6	0.00
PANK1	15.7	7	0.00
PER3	38.8	22	0.00
PLIN4	12.3	16	0.00
PNPLA3	16.6	9	0.00
UCP3	8.7	8	0.00

[a] loss-of-function

[b] probability of being LoF intolerant, values indicating extreme tolerance or intolerance were highlighted in green and red

It comes to no surprise that most of the members of the core circadian oscillator are haploinsufficient (Table 28). Only the *CRY1* gene proved to be tolerant toward loss of function mutation. However, this does not mean that life for an individual without the *CRY1* gene will be without impaired circadian functionality, just that it is possible (Van Der Horst et al. 1999).

CSNK1D and *CSNK1E* are also haploinsufficient, which makes sense considering they phosphorylate core clock proteins of the core circadian oscillator (Etchegaray et al. 2009).

Other genes that are haploinsufficient are typically found at key junctions of vital genetic pathways:

- *BHLHE40* represses transcription of *MYOD* (Hsiao et al. 2009) and *PPARGC1A* (Chung et al. 2015) and is implicated in a host of other functions
- *DBP* modulates important clock output genes (genecards.org 2018b)
- *NR1D1* functions as “gatekeeper to timely coordinate the circadian metabolic response” (Duez and Staels 2009).

- *HLF* “binds sequence-specific promoter elements to activate transcription” (genecards.org 2018c)
- *PPARGC1B*, is a PGC-1 coactivator. They “play a critical role in the maintenance of glucose, lipid, and energy homeostasis” (Lin, Handschin, and Spiegelman 2005).
- *RORA* is part of “a family of orphan nuclear receptors that are believed to play a role in numerous physiological processes, including circadian rhythm and bone metabolism” (genecards.org 2018d)
- *WEE1* “is part of a cell-size checkpoint that prevents entry into mitosis before cells have reached a critical size” (Kellogg 2003).

Only three genes were reported as recessive, meaning they are tolerant to being heterozygous:

- *NR1D2*
- *CCRN4L*
- *TEF*

In the case of *NR1D2*, Cho et al. (2012) report that “REV-ERB-a [NR1D1] and REV-ERB-b [NR1D2] together function as integral drivers of the circadian clock”. I therefore speculate that *NR1D2* knockout may still allow life, but with serious disruptions to circadian rhythmicity.
Stubblefield, Terrien, and Green (2012) report *CCRN4L* as “a potential key post-transcriptional mediator in the circadian control of many metabolic processes”. In their experiments, mice lacking it lived, and even displayed resistance to diet-induced obesity and hepatic steatosis.
TEF, according to Drolet et al. (1991), is expressed in a pituitary cell line. I speculate that while life without *TEF* is possible, a loss of it may lead to problems associated with pituitary gland development.
The rest of the gene corpus proved very tolerant to loss of function variation. The genes *ACAT2*, *CES3*, *DGAT2*, *PANK1*, *PLIN4* and *UCP3* are involved in metabolic pathways but seem to not be of essential function for these pathways. I suspect, however, that loss-of-function in these genes could lead to impaired metabolic functions. I already discussed these functions in previous chapters.
CRY1 is part of the core circadian oscillator (CCO), which makes its non-vital role surprising, However, even though Van Der Horst et al. (1999) report that *CRY1* and *CRY2* are essential for maintaining the circadian rhythm and Vitaterna et al. (1999) stress the “complexity in the role of cryptochromes in mammals”, individuals lacking the *CRY1* gene can still live, albeit with potential loss of rhythmicity.
PER3 regulates many members of the core circadian oscillator (CCO), but it is not essential for life. Bae et al. (2001) reported that mice lacking the gene do not lose rhythmicity and that “mPer3 is not essential for the core clock loops”. It may, however, “be involved in regulating circadian outputs other than locomotor activity”.
DBT is involved in the breakdown of the branched-chain amino acids (Friedrich et al. 2011). Due to its LoF-tolerance, it seems to be nonessential for life. However, potentially impaired function cannot be ruled out by this finding.
MYOD1 is a known myogenic gene. However, it is redundant in function to myogenin (Rudnicki et al. 1992; Weintraub 1993). It therefore comes to no surprise that it is tolerant towards loss-of-function mutations.

In summary, EXAC browser analysis revealed the corpus to be about evenly distributed between being tolerant or haploinsufficient towards loss-of-function mutations. As to be expected, most members of the core circadian oscillator proved to be very intolerant toward loss-of-function mutation, while many metabolic and myogenic genes were found more tolerant, most likely due to other genes being redundant in function.

3.17 Cancer Mutations

To find out whether the circadian genes of my corpus are frequently mutated in human cancer, I consulted two online databases: TumorPortal (see tumorportal.org (2018)) and cBio Portal (see cbioportal.org (2018)). TumorPortal is a comprehensive database consisting of data by Lawrence et al. (2014), who "analyzed somatic point mutations in exome sequences from 4,742 human cancers and their matched normal-tissue samples across 21 cancer types". They found that "large-scale genomic analysis can identify nearly all known cancer genes in these tumour types" (Lawrence et al. 2014). cBio Portal is an open-access portal, where cancer genomics data sets can be explored freely and interactively. Its data sets include "data from more than 5,000 tumor samples from 20 cancer studies" (Cerami et al. 2012). It was "specifically designed to address the unique data integration issues posed by large-scale cancer genomics projects and to make the raw data generated by largescale cancer genomic projects more easily and directly available to the entire cancer research community" (Cerami et al. 2012). I used both these research portals to find out more about cancer mutations in circadian genes of my corpus.

My analysis showed that, out of 32 genes, only one gene was significantly mutated in cancer (Table 29): *IDH1*. It is significantly mutated in acute myeloid leukemia, glioblastoma multiforme, significantly mutated in multiple myeloma and mutated near significance in chronic lymphocytic leukemia.

Table 29: Number of significantly mutated circadian genes of the corpus

Mutation in cancer	Number of genes
significantly mutated in cancer	1
not significantly mutated in cancer	31

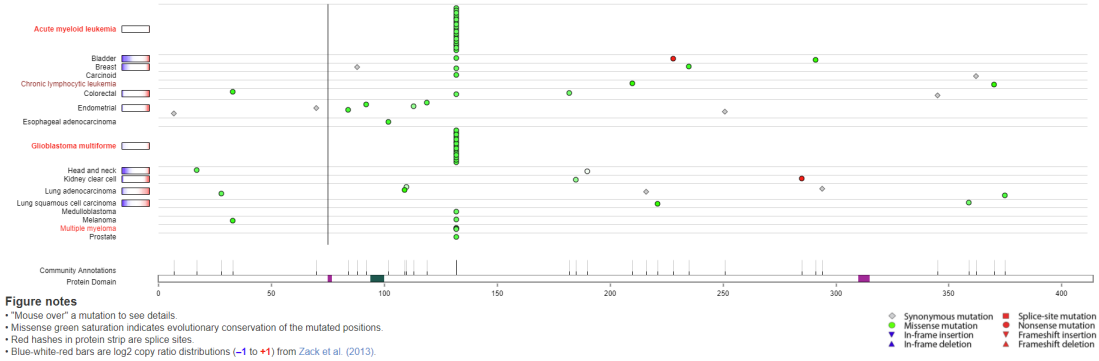


Figure 20: Mutations of *IDH1* in human cancers (TumorPortal)

TumorPortal shows that *IDH1* is often mutated in one specific region (Figure 20), namely codon 132. Mutations that occur in this codon are exclusively missense mutations. Kang et al. (2009) hypothesize that

“The recurrent nature of the IDH1 mutations in the same amino acid strongly suggests that the mutations may play important roles in the pathogenesis of glial tumors” (Kang et al. 2009). They concluded their study:

> “The data indicate that IDH1 codon 132 mutations occur not only in GBM, but also in prostate cancers and B-ALL. This study suggests that despite the infrequent incidence of the IDH1 mutations in prostate cancers and B-ALL, mutated IDH1 could be therapeutically targeted in these cancers and in glial tumors with the IDH1 mutations” (Kang et al. 2009).

cBio Portal reports high alteration frequencies (Figure 21) of:

- 100% in Glioma (UCSF)
- 77% in Glioma (TCGA)
- 74% in Glioma MSK (Neuro Oncol 2017)
- 50% in LGG-GBM (TCGA 2018)
- 20% in Cholangiocarcinoma (TCGA)

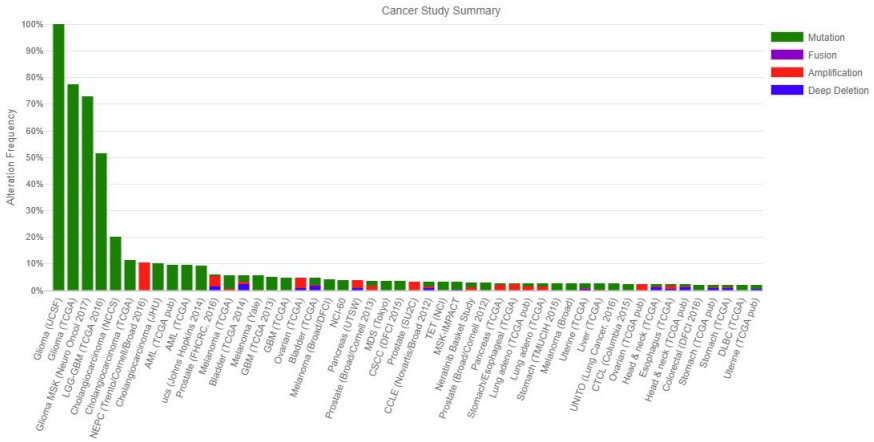


Figure 21: Mutations of *IDH1* in human cancers (cBioPortal)

Metallo et al. (2012) showed that the *IDH1*-dependent pathway of “reductive metabolism of alpha-ketoglutarate (αKG) [synthesizes] AcCoA for lipid synthesis” and that “renal cell lines deficient in the von Hippel-Lindau (VHL) tumor suppressor protein preferentially utilize reductive glutamine metabolism for lipid biosynthesis even at normal oxygen levels.” (Metallo et al. 2012). The role of *IDH1* in metabolism makes it a very attractive target gene for intervention, since under cancer conditions it would be a strong driver of cancer cell multiplication. The data from cBio Portal is therefore relevant especially in the context of the conclusions by Kang et al. (2009) and Metallo et al. (2012).

If *IDH1* could also be targeted in cancers with low incidence of *IDH1*, it would be a highly relevant field of cancer study in the context of circadian rhythms-based timing of medication and exercise not only in glioma, but a host of other cancers.

To sum up, my analysis about which circadian genes of the corpus are frequently mutated in cancer uncovered frequent mutation only in one gene: *IDH1*. It has already been identified as such in current scientific literature and a prime target for possible interventions. Circadian drug chronotherapy could potentially be used to treat or even prevent cancer outbreak more effectively in the future.

3.18 Cancer Gene Expression & Survival

To find out more about how the expression of the genes of my corpus is associated with survival in cancer, I analyzed the dataset provided by Gentles et al. (2015). They performed a meta-analysis of expression signatures from around 18,000 human tumors with overall survival outcomes across 39 malignancies. The key indicator in their analysis is the "unweighted meta-z (all cancers)" value. Positive values indicate that high expression is associated with poor survival. Negative values suggest that high expression is associated with good survival.

Table 30: Association of the expression of genes with good or poor survival in 18,000 cases of human cancer

Rank	Gene symbol	unweighted meta-z (all cancers)	p-value[a] (all cancers)
666	*WEE1*	4.86	**0.000**
923	*CSNK1D*	4.29	**0.000**
1681	*CCRN4L*	3.19	**0.001**
2126	*NFIL3*	2.76	0.006
3746	*NR1D1*	1.73	0.084
3876	*CRY1*	1.67	0.095
4182	*ACAT2*	1.52	0.129
4304	*PNPLA3*	1.47	0.142
4657	*CES3*	1.32	0.186
5413	*IDH1*	1.06	0.290
5781	*DGAT2*	0.94	0.347
5970	*CSNK1E*	0.89	0.373
7490	*PANK1*	0.55	0.584
7606	*BHLHE40*	0.52	0.600
12598	*PPARGC1B*	-0.04	0.972
15293	*MYOD1*	-0.37	0.713
15634	*FABP9*	-0.43	0.665
15824	*PLIN4*	-0.47	0.640
16049	*UCP3*	-0.51	0.607
16300	*PER1*	-0.57	0.567
18518	*CLOCK*	-1.21	0.228
21553	*TEF*	-2.83	0.005
21881	*NR1D2*	-3.13	**0.002**
21979	*DBT*	-3.25	**0.001**
22388	*ARNTL*	-3.77	**0.000**
22477	*PER2*	-3.91	**0.000**
22754	*PER3*	-4.41	**0.000**
23023	*CRY2*	-5.40	**0.000**
23145	*RORA*	-6.06	**0.000**
23254	*HLF*	-7.56	**0.000**

[a] significant ($p < 0.05$) values highlighted by bold formatting

The circadian genes of my corpus are listed between ranks 666 and 23254 and seem to influence survival mostly in various forms of brain cancer (Brain cancer neuroblastoma and glioma), both with a positive and negative impact on survival (Table 30). Other cancers that were strongly influenced were cancers of the lung (z-score -5.97of *HLF*, z-score -4.49 of *PER3*) and pancreas (z-score -5.15 of *PER2*). The most positive impact was found in *RORA* with a z-score of -8.24 in brain cancer neuroblastoma. The most negative impact was found in *WEE1* with a z-score of 8.11 in brain cancer neuroblastoma.
The circadian genes most detrimental to survival were:

- *WEE1* (4.86), ranked 666
- *CSNK1D* (4.29), ranked 923
- *CCRN4L* (3.19), ranked 1681
- *NFIL3* (2.76), ranked 2126

The data for most of these genes was significant ($p<0.05$). They regulate either the cell cycle (*WEE1* and *CSNK1D*) or metabolism (*CCRN4L* and *NFIL3*). In the case of *CSNK1D*, Schittek and Sinnberg (2014) attest that it "regulate[s] key signaling pathways known to be critically involved in tumor progression". It thus comes as no surprise that, if these genes get overregulated by cancer, cancerous cells can spread even faster, decreasing the likelihood of survival.
Interestingly, *IDH1*, a gene commonly implicated in mutation leading to cancer, did not rank very high in z-score (1.06). Given its role in cancer onset due to point mutations it is surprising that its potential impact regarding cancer survival did not have a higher meta-z score. However, Bent et al. (2010) report that, while *IDH1* mutations may serve a prognostic purpose, they are not in themselves predictive of outcome (i.e. survival).
The circadian genes most beneficial to survival were:

- *HLF* (-7.56), ranked 23254
- *RORA* (-6.06), ranked 23145
- *CRY2* (-5.40), ranked 23023
- *PER3* (-4.41), ranked 22754
- *PER2* (-3.91), ranked 22477
- *ARNTL* (-3.77), ranked 22388

The data for all these genes was significant ($p<0.05$). Most of them are part of the core circadian oscillator (CCO) (*ARNTL*, *PER2*, *CRY2*) or interact tightly with it (*PER3*, *RORA*).
The finding that increased expression of *ARNTL* increases cancer survival rates makes sense in the context of the research of Mullenders et al. (2009), who found that it regulates the p53 tumor suppressor pathway. Interestingly, Schittek and Sinnberg (2014) report that *CSNK1D* also can regulate p53-signalling, albeit in the other direction. It could be interesting to see whether drugs or exercise interventions could, with appropriate circadian timing, be used to strengthen the p53 pathway from both ends of the equation.

Chen et al. (2005) suggest that “disturbances in PER gene expression may result in disruption of the control of the normal circadian clock, thus benefiting the survival of cancer cells and promoting carcinogenesis”, explaining the positive impact of increased *PER2* and *PER3* expression in cancer survival.
CRY2, like the other members of the negative arm, seems to also be helpful as a tumor suppressor. It was reported by Mao et al. (2015) that “knockdown of CRY2 causes the epigenetic dysregulation of genes involved in cancer-relevant pathways”. However, there are also other findings by Ozturk et al. (2009), who report a detrimental effect of *CRY2* expression in p53 mutant mice.
RORA is found on a common fragile site (CFS). “Common fragile sites (CFSs) are large genomic regions present in all individuals that are highly unstable and prone to breakage and rearrangement, especially in cancer cells with genomic instability” (Zhu et al. 2006). Zhu et al. (2006) further explain that

> “RORA appears to play a critical role in responses to cellular stress, lending further support to the idea that the large CFS genes function as part of a highly conserved stress response network that is uniquely susceptible to genomic instability in cancer cells”.

Thus, high *RORA* expression could indicate a stable stress response network, which would be beneficial to survival in cancer.
HLF had the lowest meta-z score of -7.56, proving most beneficial to survival in cancer. I could not find current scientific literature explaining this finding. Thus how *HLF* can increase survival in cancer is an area in need of further scientific study.
All in all, my analysis of the influence of circadian genes on survival showed that they were relevant mostly in cases of brain cancer. Looking at the significant results, the four genes most detrimental to survival were part of metabolic or cell cycle processes. I hypothesize that their upregulation could increase the rate with which cancerous cells are multiplied.
I found that six genes, among them one gene of the positive arm and three members of the negative arm, were beneficial to survival. These were already associated with tumour suppressor pathways in current scientific literature. The metabolic gene *RORA* was also identified and is part of a common fracture site already uncovered in the literature, which is a mechanism of cellular stress response. The positive impact of the *HLF* gene could not be explained.

3.19 Cancer Cell Expression

To find out in which cell lines the genes of my corpus are expressed, I consulted the Cancer Cell Line Encyclopaedia (CCLE). It is a compilation of gene expression, chromosomal copy number and parallel sequencing data from 947 human cancer cell lines. When combined with pharmacological profiles for 24 anticancer drugs across 479 of the cell lines, this collection allows identification of genetic, lineage, and gene-expression-based predictors of drug sensitivity (Barretina et al. 2012).

Table 31: Expression of circadian genes in cancer cell lines

Expression in cancer cell lines	Number of genes	Gene symbols
Upregulated in most cancer cell lines	23	*ACAT2*, *ARNTL*, *BHLHE40*, *CLOCK*, *CRY1*, *CRY2*, *CSNK1D*, *DBP*, *DBT*, *DGAT2*, *NR1D1*, *NR1D2*, *NHRPNPU*, *IDH1*, *NFIL3*, *CCRN4L*, *PANK1*, *PER1*, *PER2*, *PER3*, *TEF*, *WEE1*
Downregulated in most cancer cell lines	7	*CES3*, *FABP9*, *HLF*, *MYOD1*, *PLIN4*, *RORA*, *UCP3*
Mixed	2	*PNPLA3*, *PPARGC1B*

My analysis showed that most circadian genes that I investigated were upregulated in most cancer cell lines (Table 31). Cancer lines of T- and B- Cell cancers were most present in the higher ranges of upregulation. Other cancers present were cancers of the brain, blood, cartilage, lung, nerve tissue, prostate and skin.

Among these genes are all genes of the core circadian oscillator. The forward arm, consisting of *CLOCK* (Figure 22) and *ARNTL*, upregulate many other genes. It makes sense for them to be upregulated by cancer to increase cancer cell growth. Gorbacheva et al. (2005) found:

> "CLOCK/BMAL1 may be directly involved in the regulation of expression of cell-cycle- and apoptosis-related genes, in which case daily and genotype-dependent variations in balance of their products will determine response to stress at the individual cell level" (Gorbacheva et al. 2005)

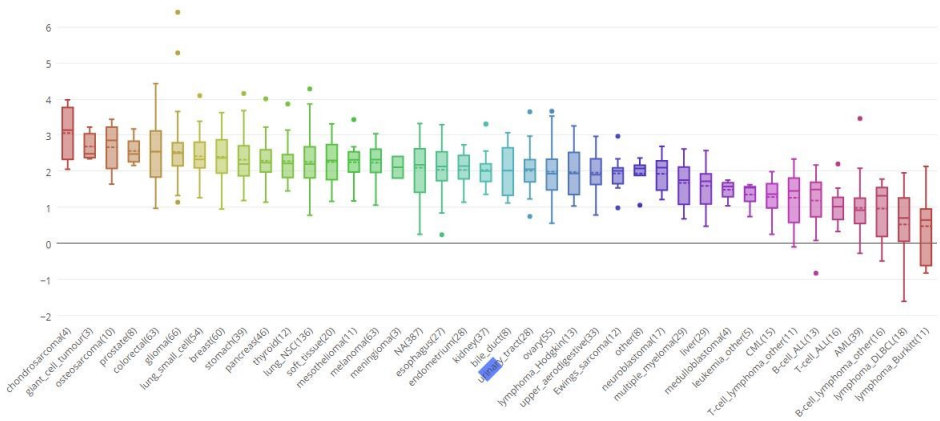
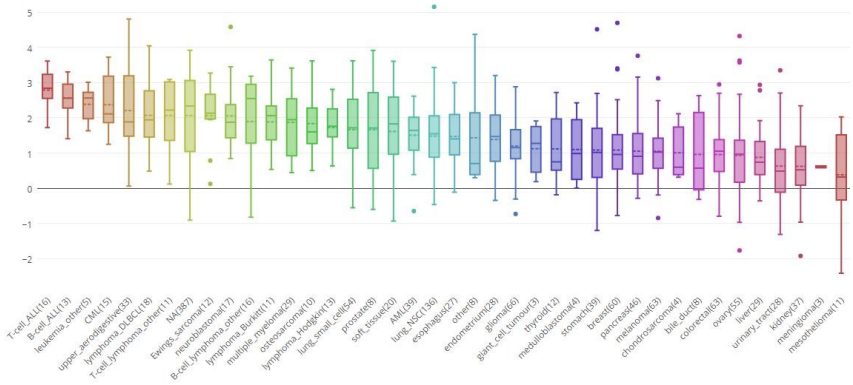


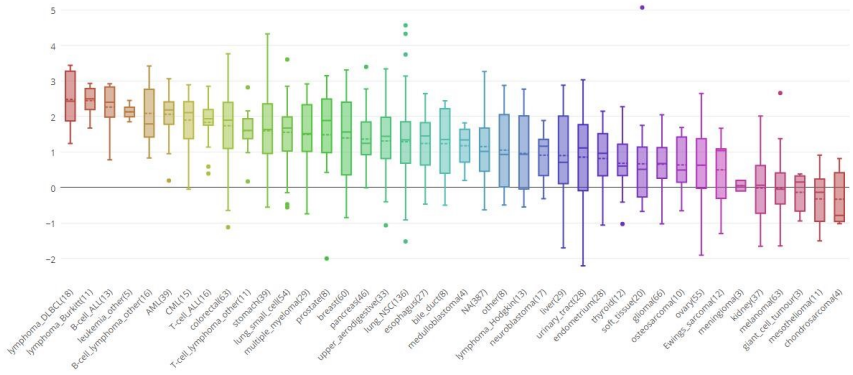
Figure 22: Expression of *CLOCK* in different cancer cell lines

This finding could be used to reduce negative reactions to chemotherapy agents by using appropriate circadian chemotherapy timing.

Two members of the negative arm were also upregulated: For *PER1* (Figure 23a), Gery et al. (2006) report that "Overexpression of Per1 sensitized human cancer cells to DNA damage-induced apoptosis" and that it is part of a network of checkpoint pathways that initiate either cell cycle arrest and DNA repair or apoptosis. It seems that *PER1* overexpression is not a consequence of the cancer itself but potentially part of the body's defense mechanism against it.



(a) Expression of *PER1* in different cancer cell lines



(b) Expression of *PER2* in different cancer cell lines

Figure 23: Expressions of *PER1* and *PER2* in different cancer cell lines

The same might be true for *PER2* (Figure 23b). Fu et al. (2002) showed that “loss of mPer function results in increased tumor development and deficiencies in DNA damage response in mice”. They hypothesize that it can be regarded as a tumor suppressor.

As for the other genes, Sahar and Sassone-Corsi (2009) give a very concise overview, stating that “Several cell cycle genes, such as Wee1, myc and cyclin D1, are regulated by the circadian clock”. And, that “cellular metabolism is altered in cancer. Several key metabolic genes are under circadian regulation”. Only seven genes were mostly downregulated in cancer cell lines. *MYOD1*, one of the muscle maker genes, is among the downregulated genes (Figure 24).

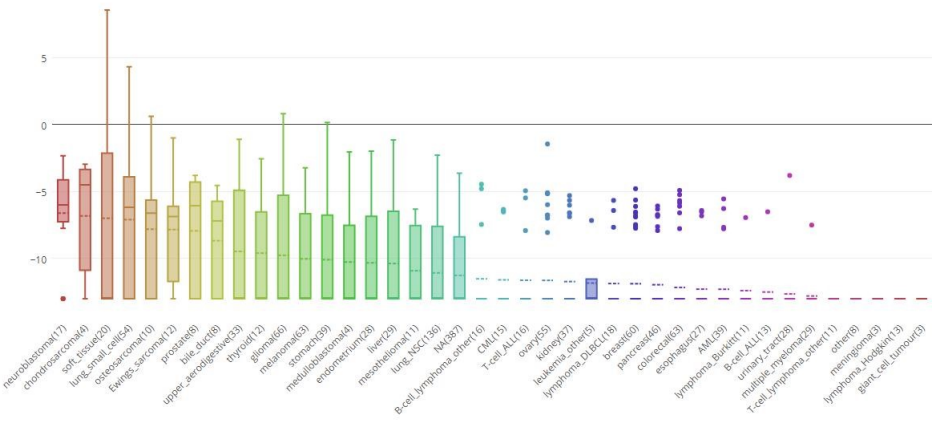


Figure 24: Expression of *MYOD1* in different cancer cell lines

It is at first confounding that a gene that promotes myogenesis would be downregulated in cancer instead of being severely upregulated. However, considering that *MYOD1* induces myogenic differentiation of muscle precursor cells into muscle cells, which do not proliferate anymore, this finding makes sense. Abbas and Dutta (2009) found that it "control[s] CDKN1A transcription and upregulate[s] p21 in response to a plethora of stimuli and anticancer agents". According to OMIM, "CDKN1A plays a critical role in the cellular response to DNA damage, and its overexpression results in cell cycle arrest" (omim.org 2018k).
Interestingly, *FABP9* was not regulated in most cancer cell lines (Figure 25). In those cancers that regulated *FABP9*, it was strongly downregulated with prostate cancer being the most relevant.

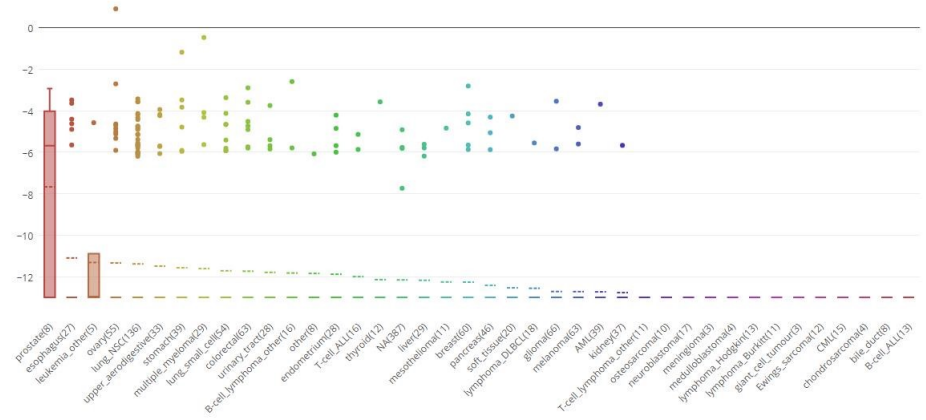
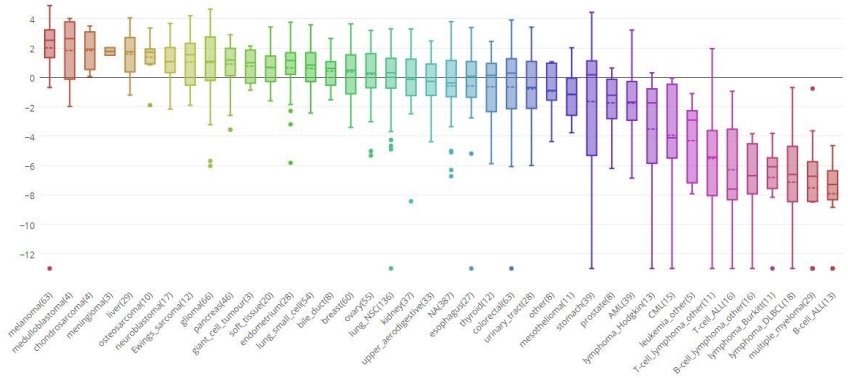


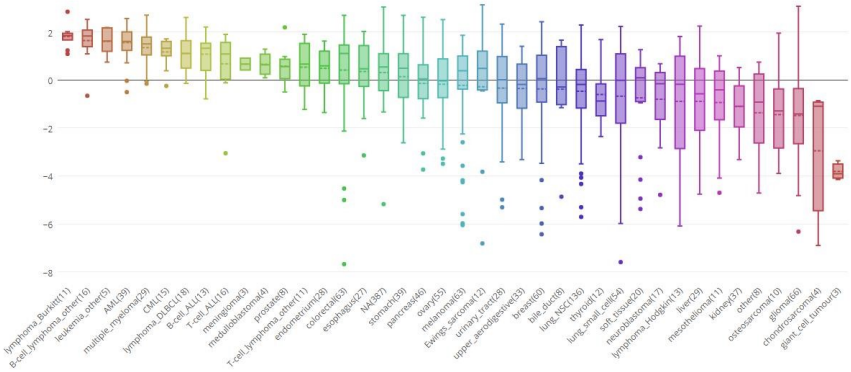
Figure 25: Expression of *FABP9* in different cancer cell lines

Contrary to this finding, Al Fayi et al. (2016) report that *FABP9* expression is strongly increased in prostate cancer. They conclude "FABP9 is a valuable prognostic marker to predict the outcomes of prostate cancer patients, perhaps by playing an important role in prostate cancer cell invasion" (Al Fayi et al. 2016).
These contradictory findings are curious because they are both centered on prostate cancer. I suspect that they are the result of different methodologies and therefore conclude that the scientific community should undertake more research to be able to reconcile their findings.
Two genes, *PNPLA3* (Figure 26a) and *PPARGC1B* (Figure 26b), were both up- and downregulated across

cell lines. I suspect that their activities are not essential for cancer cell development in all cancer cell lines and/or that their up- or downregulation is a byproduct of other pathways.



(a) Expression of *PNPLA3* in different cancer cell lines



(b) Expression of *PPARGC1B* in different cancer cell lines

Figure 26: Expressions of *PNPLA3* and *PPARGC1B* in different cancer cell lines

In summary, I found that cancer upregulates most circadian genes (23 out of 32) and that most genes of the core circadian oscillator were among these upregulated genes. Cancer was found to upregulate the *PER* genes of the negative arm, who are regarded as potential cancer suppressors in the scientific literature and I suspect that this pattern shows a defensive reaction of the genes rather than one beneficial to the cancer. However, this clearly needs more research. All in all, the fact that cancer upregulates circadian genes in cancer cells makes them potentially effective targets for circadian interventions (chronotherapy) in the form of exercise or medication.

3.20 YAP-ERMS Expression

To find out whether the circadian genes of my corpus are differently expressed between embryonal rhabdomyosarcomas caused by *YAP1 S127A* expression in activated satellite cells compared to control skeletal muscle, I consulted the data created by Tremblay et al. (2014).

> "Embryonal rhabdomyosarcoma (ERMS) usually affects children in their first 5 years of life, but it is the most common type of RMS at all ages. The cells of ERMS look like the developing

muscle cells of a 6- to 8-week-old embryo. ERMS tends to occur in the head and neck area, bladder, vagina, or in or around the prostate and testicles" (cancer.org 2018).

Tremblay et al. (2014) investigated the role of the Hippo pathway effector *YAP1* in soft tissue sarcomas and were able to identify *YAP1* as a potent ERMS oncogenic driver and a promising target for differentiation therapy. To do so, they assessed the abundance and cellular localization of *YAP1* protein in 78 ARMS and 196 ERMS samples. I used their collected data and reanalyzed it to find out more about my selection of circadian genes under the influence of *YAP*-induced rhabdomyosarcoma.

Table 32: Differential expression of circadian genes between embryonal rhabdomyosarcomas caused by *YAP1 S127A* expression in activated satellite cells versus control skeletal muscle

Gene symbol	**log2 fold change[a] (ERMS versus muscle)**	**adj. p-value[b]**
Wee1	1.44	**0.00**
Idh1	1.03	**0.00**
Acat2	0.64	**0.00**
Csnk1d	0.62	**0.00**
Myod1	0.52	**0.00**
Arntl2	0.23	0.09
Csnk1e	0.14	0.05
Clock	0.05	0.52
Ces3a	-0.14	0.18
Fabp9	-0.19	0.11
Noct/Ccrn4l	-0.55	**0.00**
Bhlhe40	-0.65	**0.00**
Nfil3	-0.71	**0.00**
Per1	-0.78	**0.00**
Per2	-0.88	**0.00**
Cry1	-1.06	**0.00**
Pank1	-1.26	**0.00**
Dbt	-1.34	**0.00**
Nr1d2	-1.35	**0.00**
Ppargc1b	-1.36	**0.00**
Hlf	-1.45	**0.00**
Pnpla3	-1.55	**0.00**
Cry2	-1.75	**0.00**
Nr1d1	-1.75	**0.00**
Per3	-1.83	**0.00**
Tef	-2.09	**0.00**
Dgat2	-2.94	**0.00**
Rora	-3.02	**0.00**
Plin4	-4.04	**0.00**
Ucp3	-4.36	**0.00**

[a] colored formatting was applied to highlight differences

[b] bold formatting was applied to highlight significance

Most of the data concerning the circadian genes of my corpus was significant ($p<0.05$). It showed one quite upregulated gene in embryonal rhabdomyosarcoma (ERMS) as well as a host of extremely down-regulated ones (Table 32).

The two upregulated genes were *WEE1* and *IDH1*. Kahen et al. (2016) identified *WEE1* as a potential treatment target for rhabdomyosarcoma:

> "The Wee1 kinase maintains cells in G2/M arrest, providing time for DNA repair prior to mitosis. When the Wee1 kinase is inhibited, CDK1/2 activity proceeds unchecked, and cells progress prematurely through G2/M leading to mitotic progression, DNA strand breaks, mitotic catastrophe, and cell death."(Kahen et al. 2016)

The high expression of *WEE1* in the data of Tremblay et al. (2014) further supports this as a potentially viable treatment option.
In the case of *IDH1*, a survey by Shukla et al. (2011) could not identify any mutations of the gene in pediatric solid tumors. They further explain that "IDH mutation studies of pediatric gliomas reveal that they occur in children 14 years of age or older and not in younger children". Ward et al. (2010) found that "somatic mutations in cytosolic isocitrate dehydrogenase 1 (IDH1) observed in gliomas can lead to the production of 2-hydroxyglutarate (2HG)", which is an agent of tumorigenesis. The mechanism of increased *IDH1* in ERMS, however, seems so far unresearched.
The genes of the negative arm of the CCO as well as their modulator were all downregulated in ERMS:

- *PER1* (fold change -0.78)
- *PER2* (fold change -0.88)
- *CRY1* (fold change -1.06)
- *CRY2* (fold change -1.75)
- *PER3* (fold change -1.83)

This makes sense, as the negative arm is associated with tumor suppressor functions (Gery et al. 2006; Fu et al. 2002). These functions seemingly need to be ineffective for ERMS development to take place.
Other genes that were extremely downregulated were:

- *UCP3* (fold change -4.36)
- *PLIN4* (fold change -4.04)
- *RORA* (fold change -3.02)
- *DGAT2* (fold change -2.94)
- *TEF* (fold change -2.09)

The finding that ERMS strongly downregulates *UCP3* is contrary to Argilés et al. (2014), who reported that *UCP3* was upregulated in both experimental animals and human subjects affected by cachectic tumors, resulting in "decreased oxidative capacity, disrupted protein synthesis, changes in membrane fluidity and oxidatively modified mitochondrial proteins; all of these result in impaired mitochondrial function" (Argilés et al. 2014). More research will be needed to illuminate the different effects between cancers.
Collins et al. (2002) report similar findings about *UCP3* in weight loss associated with gastrointestinal adenocarcinoma in humans, stating that "Elevations in muscle uncoupling proteins-3 activity may enhance energy expenditure and this in turn could contribute to tissue catabolism". I suspect that in ERMS, there are other factors at play. However, I could not find current scientific literature to further illuminate this.

I recommend that more studies should be undertaken to allow a more differentiated view of the role of *UCP3* in ERMS as opposed to other cancers.

PLIN4, like *UCP3*, is a circadian gene involved with lipid metabolism. No current scientific literature could be found regarding its role in ERMS. However, Tirinato et al. (2017) explain that "tumor cells are able to uptake larger amount of lipids, as well as to enhance lipogenesis and CH production, and to increase FA β-oxidation" compared to normal cells. *PLIN4*, a member of the PAT (perilipin-ADRP-TIP47) family and is "involved in the control of the intracellular neutral lipid packaging and trafficking" (Tirinato et al. 2017), making it a potential cancer target. However, as in *UCP3*, this is contradictory to the finding of ERMS downregulating *PLIN4*. Due to lack of current scientific literature about its role in ERMS, I can only speculate that ERMS regulates *UCP3* differently than other cancers.

RORA "appears to play a critical role in responses to cellular stress, lending further support to the idea that the large CFS genes function as part of a highly conserved stress response network that is uniquely susceptible to genomic instability in cancer cells" (Zhu et al. 2006). In the case of ERMS, this cellular stress response seems to be downregulated to allow cancer cell development. I could not find current scientific literature about its role in ERMS to confirm this speculation.

Currie et al. (2013) explain that the metabolic gene *DGAT2* "catalyzes the only dedicated step in TG [triacylglycerols] formation and thus provides a key target for decreasing available lipids by increasing lipid storage". In cancer, the energy stored in fat is used to produce more and more cancer cells, which is why Currie et al. (2013) hypothesize that "limiting fatty acid availability can control cancer cell proliferation".

I speculate that *DGAT2* downregulation in ERMS could potentially be a defense mechanism of the body, however I could not find scientific literature to support this claim. Thus, this might also be attributed to the specific nature of ERMS.

Other noteworthy downregulated genes were *ACAT2* (fold change 0.64), *CSNK1D* (fold change 0.62) and *MYOD1* (fold change 0.52). *ACAT2*, together with the ACLY enzyme, is involved in acetyl-CoA synthesis. Zaidi, Swinnen, and Smans (2012) found out that

> "increasing evidence supports the importance of ACLY in tumor cell growth. The fact that ACLY is upregulated in several types of cancer cells and that, upon inhibition of ACLY, cancer cells undergo a proliferation arrest both in vivo and in vitro indicates that this enzyme plays an important role in cancer cell progression." (Zaidi, Swinnen, and Smans 2012)

Yet, evidence linking these findings to ERMS does not exist so far in the literature.

Rosenberg et al. (2015) were able to identify *CSNK1D* as a treatment target for breast cancer and reported that "CK1δ is a vulnerability of human breast cancer subtypes overexpressing this kinase" (Rosenberg et al. 2015). The data of Tremblay et al. (2014) potentially suggests that this might be true for ERMS as well. Nonetheless, this clearly needs more research to rise beyond speculation.

MYOD1, an important gene for the differentiation of muscle precursor cells into fully developed muscle cells that do not proliferate further, plays a key role in ERMS. As Tremblay et al. (2014) themselves explain: "YAP1-TEAD1 upregulate pro-proliferative and oncogenic genes and maintain the ERMS differentiation block by interfering with MYOD1 and MEF2 pro-differentiation activities".

Yang et al. (2009) propose that “rhabdomyosarcomas represent an arrested progress through a normal transitional state that is regulated by the relative abundance of heterodimers between MyoD and the full-length E2A proteins”. Increasing the abundance of MyoD:E-protein heterodimers seems to be a possible solution to this predicament, however, all the mechanisms involved are still poorly understood. I speculate that circadian patterns of *MYOD1* expression could be exploited to further optimize potential treatment methods.

In conclusion, I found that rhabdomyosarcomas caused by *YAP1 S127A* expression downregulated most circadian genes in activated satellite cells compared to control skeletal muscle. Among them were the *PER* genes, which are known tumour suppressors (Gery et al. 2006; Fu et al. 2002). They seemingly need to be ineffective for ERMS development to take place. The role of other metabolically active genes (*UCP3*, *PLIN4*, *RORA*, *DGAT2*, *TEF*) in ERMS remained speculative, while the role of the myogenic gene *MYOD1* is well explained by Yang et al. (2009), who found that in ERMS, it can no longer differentiate muscle precursor cells due to downregulation, allowing ERMS to take place. It is a potential treatment target.

However, ERMS upregulated two genes:

- *WEE1*, a gene important for the cell cycle that was already identified as a potential ERMS treatment target by Kahen et al. (2016).
- *IDH1*, a gene implicated in cancer onset due to point mutation, was so far not found to be implicated in ERMS onset (Shukla et al. 2011). The mechanism of increased *IDH1* in ERMS seems under researched to date.

4 Conclusions

I will now begin answering the initial research questions based on my results and current scientific literature. To do so in a structured and comprehensive manner, I will answer related questions together in the form of 6 sections. Within these sections, I will summarize related results as well as discuss them in relation to each other.

4.1 The general structure and localization of the circadian genes in question

During the first five steps of my analysis, I focused on the following research questions:

1. Are the genes/proteins of the corpus functionally linked? Do they belong to the same signalling pathway?
2. Do the circadian genes of the corpus share common features?
3. Are the genes/proteins predicted to be secreted from cells?
4. In what tissues are the genes expressed?
5. What is current research about the genes?

During my analysis I found that there is indeed evidence that my set of genes is functionally linked. However, this is not surprising as I selected genes where I knew at the outset that they were functionally related. StringDB found that they have more interactions than what would be expected of a random set of genes.
The analysis was able to display the most researched functional structures of circadian genetics: The forward arm of the core circadian oscillator and the negative arm of the core circadian oscillator.
It also showed well known interactions in the form of secondary loops, such as the interactions between *PER3* and all genes of the negative arm and the interactions between *ARNTL*, *RORA* and *NR1D1*.
Apart from known interactions, StringDB analysis also highlighted some less researched connections:

- *CCRN4L* interacts with both components of the forward arm
- *HLF* interacts with the core heterodimer (*CLOCK/ARNTL*),
- *DGAT2* and *PNPLA3* interact with each other

All of these interactions are not yet well researched and could potentially bear good research opportunities.
ToppGene enrichment analysis of molecular function, biological processes and cellular components showed that my set of circadian genes is transcriptionally active and influences the expression levels of other genes to control circadian rhythmicity mostly from within the cytoplasm. They also repress glucocorticoid signalling pathways. A high prevalence of functions regarding metabolism and homeostasis links the circadian genes to metabolic functions.
When dysfunctional, mutations of circadian genes can lead to a host of phase-shift-based sleep disorders as well as psychological disorders such as depression or bipolar disorder. In mice, mutated specimens show loss or disturbances of the circadian rhythm.
No gene of the investigated corpus was found in the extracellular matrix. However, current scientific literature shows interactions between circadian genes and genes of the extracellular matrix in the context of osseointegration. Only two genes were predicted to be secreted from cells: *ARNTL* and *CRY2*. I could not find current scientific literature concerned with this result and suggest further research in this area.
Gtex Analysis showed that the majority of the circadian genes I investigated were expressed in skeletal muscle. Expression levels compared were, however, mostly below medium level. The *CRY2* gene was found to be highly expressed in skeletal muscle, while *MYOD1* and *UCP3* were exclusively expressed there. Current scientific literature hypothesizes that ”circadian rhythms will have a central role in the daily regulation of skeletal muscle function and phenotype” (McCarthy et al. 2007).
Current research in general currently is focused on the following themes in regard to circadian genes of the corpus:

- Endurance Pathways
- Metabolism Pathways
- Muscle Ageing
- Sleep Disorders
- ADHD and drug abuse

- Psychological Disorders / Mood Regulation
- Gut Microbia
- Telomere Length and Longevity

In summary, I was able to recreate the known core oscillator pattern of circadian genes as well as secondary loops via String-DB analysis. ToppGene enrichment analysis provided information that these genes are transcriptionally active and have metabolic functions. Mutations lead to phase-shifts resulting in disorders of sleep and psychological functions. In mice, mutant phenotypes also showed disturbed or lost rhythms. On the cellular level, only two genes were active in the form of secretion while the rest acts solely from within the nucleus. On the tissue level, I found that most circadian genes were expressed in skeletal muscle, albeit only seldomly expressed in a high or exclusive manner. Current scientific research is concerned with the relation of circadian genes with endurance and metabolism pathways as well as their role in muscle ageing, psychology, gut health as well as longevity.

4.2 The effect of exercise on the researched circadian genes

During the next two steps of my analysis, I focused on the following research questions:

1. Does the genes change their expression after strength or endurance exercise?
2. Do the proteins encoded by your genes become phosphorylated in skeletal muscle or other tissues in response to exercise?

My analysis of the data Vissing and Schjerling (2014) provided showed that a majority, namely up to 90% of the circadian genes studied, were influenced in their expression through exercise. Resistance training upregulated up to 26% of the corpus and downregulated up to 23%. Endurance training upregulated up to 29% of the corupus and downregulated up to 16%.
The data showed multiple candidate genes for potential circadian timing exercise interventions:

- *ARNTL*: part of the forward arm, upregulated by resistance training as well as endurance training
- *PER1*, *PER2*, *PER3*: part or modulator of the negative arm, especially downregulated after resistance training
- *NR1D1* and *NR1D2*: metabolism genes, downregulated after resistance training
- *BHLHE40*: a modulator of both *MYOD1* and *PPARGC1A*, upregulated after endurance exercise and downregulated after resistance exercise

Phosphorylation data after either endurance or resistance high-intensity exercise provided by Hoffman et al. (2015) and Deshmukh et al. (2015) on the other hand showed that only few genes are subject to posttranslational modifications due to exercise.
Among them, even fewer overlapped with genes regulated by exercise:

- *Plin4*: upregulated after endurance exercise and phosphorylated after high intensity endurance and resistance exercise

- *Csnk1d*: upregulated after resistance exercise and phosphorylated after high intensity endurance exercise
- *Hnrnpu*: upregulated after resistance exercise and phosphorylated after high intensity endurance exercise
- *Dbt*: upregulated after endurance exercise and phosphorylated after high intensity endurance exercise

These findings have not yet been put in context with each other in the current literature. To do so, I consulted the studies "Scheduled Exercise Phase Shifts the Circadian Clock in Skeletal Muscle" (Wolff and Esser 2012) and "Circadian Rhythms, skeletal muscle molecular clocks and exercise" (Schroder and Esser 2013), where the authors also investigated the effects of exercise on the circadian clock.
Among the central findings of Wolff and Esser (2012) was the fact that in mice, scheduled exercise during a lights-on condition could produce significant phase shifts in the circadian rhythm of muscle tissue, but not of the SCN. They found similar effects with restricted feeding. In the context of these findings they elaborate:

> "In skeletal muscle, genes involved in fatty acid metabolism such as PGC-1β, UCP3, Dbt, Dgat2, and Acat2 have been shown to oscillate in a circadian manner (period $\tilde{2}$4hrs) and are important for metabolic functions such as fatty acid oxidation [...], fatty/cholesterol synthesis [...]. Because daily activities for most species occur with some circadian rhythmicity (active phase-inactive phase every 24hrs) and skeletal muscle is a highly metabolic tissue, it is reasonable that genes involved in energy storage/utilization would also be under circadian control in order to anticipate the daily demands on the system" (Wolff and Esser 2012).

Schroder and Esser (2013) found that exercise studies in mice suggest there is an optimal time of day to make clock function more robust, however at this time it is not possible to extrapolate this to humans. Until then, they recommend consistently timed daily exercise. I conclude with the notion that my own findings about how metabolic circadian genes in particular are influenced by different modes of exercise could be used to further differentiate possible methods of using scheduled exercise. This could allow us to synchronize circadian rhythms in humans, which could be especially relevant in the context of metabolic disease.
In summary, my analysis could show known and new possible exercise intervention targets among the circadian genes I investigated. In general, almost all genes proved to be influenced by exercise. Both resistance and endurance training influenced a host of genes in the form of both up- and downregulation. Multiple metabolic genes were influenced in their phosphorylation by high intensity resistance and endurance exercise. Current research is concerned with uncovering practical applications to make circadian rhythms more stable or phase shift them through entrainment, however so far there are only general recommendations to be made.

4.3 The role of the circadian genes in myogenesis after mechanical overload or injury

I conducted two analyses concerned with answering the following questions about the role of circadian genes in myogenesis:

1. Do the genes change their expression during synergist ablation in mice?
2. Does the expression of the genes change in response to cardiotoxin-induced muscle injury?

After mechanical overloard due to syergist ablation, the circadian genes reacted both with strong up- and downregulation and acted differently during the respective early, intermediate and late stages of response. One of the most upregulated genes was *Wee1*, a gene involved in the cell cycle regulating entry into mitosis. *Myod1*, a known myogenic gene, was also highly upregulated. The role of the also upregulated genes *Idh1*, *Bhlhe40* and *Nfil3* remains unclear until further research.
In response to cardiotoxin injury, the genes *Bhlhe40*, *Csnk1d*, *Csnk1e*, and *Idh1* were also upregulated. I suspect that these genes serve mechanistically similar purposes in response to both conditions.
The involvement of the positive arm during response to cardiotoxin injury could not be found in response to synergist ablation. The same was true for the gene *DGAT2*, which I speculate is involved in fatty regeneration as a response to muscle damage (Lukjanenko et al. 2013).
Multiple known metabolic genes were downregulated in response to synergist ablation. I hypothesize that their expression was decreased because of antagonistic myogenic pathways, but this needs further research. The same was true in response to cardiotoxin injury. I suspect there could be mechanistic similarities between both cases. However, further research will be needed to prove this hypothesis.
In short, both hypertrophic and regenerative processes after injury influenced circadian genes. There seems to be overlap in the activity of many metabolic genes during both processes. I hypothesize that their downregulation could be a result of angagonistic myogenic pathways, however there is need for further research to confirm this.

4.4 The role of the circadian genes in regard to muscle fiber properties

I looked for the involvement of circadian genes in relation to multiple research questions concerning different muscle fiber properties:

1. Are the genes differently expressed in mice that have been selectively bred for high body size and muscle mass (Berlin high) or small body size and muscle mass (Berlin low)?
2. Does the expression of the genes differ in-between type 1 and type 2B fibres in mice?
3. Are the genes detectable as proteins in mouse muscle and C2C12 myotubes? Do their concentrations differ between young and old type 1 and 2A muscle fibres?
4. Are the genes expressed in a circadian fashion within mouse skeletal muscle?

Looking at gene activity of the two differently hypertrophied mouse strains, berlin high and berlin low, bioinformatic analysis revealed a pattern of increased forward arm activity in the former and increased

negative arm activity in the latter. The less hypertrophied strain, BEL, also showed more active metabolic genes compared to BEH. The role of circadian genes in more or less hypertrophied muscle fiber should be subjected to further research to uncover, if the described role of these genes is causally linked to increased or decreased myogenesis.

Comparing type 1 and type 2b mouse muscle fiber, I observed that 50% of the circadian gene corpus was more active in slower type 1 fibers and that those genes were generally metabolically active. However, only the data of the genes *Dgat2*, *Pank1* and *Nfil3* was significant, putting this assumption in need of further research.

The two muscle models whole muscle and C2C12 myotubes showed that circadian genes can be found active in both muscle fiber and muscle periphery. All genes detected were part of metabolic pathways. However, increased activity of *Per1* in the whole muscle model remained confounding and in need of further research.

While investigating differences between fast and slow muscle fibers in younger and older human muscle fibers, I found four circadian genes that were involved in metabolic processes. The majority of these genes was more active in older muscle fibers, regardless of contractile property. They also were in part already associated with muscle breakdown and sarcopenia in current scientific literature. The role of circadian genes in muscle ageing should be investigated further.

When I analyzed the circadian rhythms of the genes in mouse skeletal muscle, I found that they peaked mostly during the late hours of the day or early hours of night and had their lowest expression late in the night or during the morning. The analysis was able to replicate the well-known patterns of the forward and negative arms of the core circadian oscillator. It also showed possible interactions between the circadian rhythms of the important muscle genes *MYOD1* and *UCP3*. It also showed interesting interactions between *MYOD1*, *RORA* and *NR1D1*. These connections could offer interesting research opportunities regarding the interactions of myogenesis and metabolism in a circadian context.

Bhlhe40, a regulator of both strength and endurance adaptation pathways, was shown to be more active both in more hypertrophied mice as well as in fast 2b mouse muscle fiber. At the same time, however, data about circadian activity in muscle was lacking. It could provide interesting new research opportunities in the context of using circadian timing to optimize hypertrophy, especially in the context of concurrent resistance and endurance training.

To sum up, bioinformatic analysis of the role of circadian genes in regard to muscle fiber properties showed that in a more hypertrophied mouse phenotype, genes of the forward arm were more active. At the same time, in a less hypertrophied mouse phenotype, metabolic genes showed more activity. Between slower and faster mouse muscle fibers, there was more circadian gene activity in the slower fiber type. Looking at the difference between a model of isolated mouse muscle fiber and muscle fiber within the tissue context, I found that circadian genes were active in both models. All detected genes were metabolically active. However, increased *Per1* activity in the whole muscle model remained unexplained. When I analyzed differences between older and younger muscle fibers, I detected increased metabolic gene activity within older fibers. These genes were generally associated with processes of muscle breakdown and sarcopenia. Lastly, when directly analyzing circadian gene rhythms within muscle fiber, my analysis showed peak

activity mostly during the late hours of the day and lowest expression late in the night or in the morning. Thereby, both known interactions as well as potentially new interlocked osciallation patterns could be uncovered.

4.5 Phenotypes and diseases associated with the circadian genes of the corpus

Another area of research covered by my bioinformatic analysis was the relationship of circadian genes and phenotypes and diseases. I tried to answer the following research questions:

1. Does mutation of the genes cause phenotypes in mice?
2. Are there SNPs near the genes associated with human phenotypes in genome-wide association studies (GWAS)?
3. How many heterozygous and homozygous loss-of-function mutations exist in 60,607 humans?

In mice, mutations of circadian genes mostly caused preweaning lethality or changes to metabolism, body and organ structure. Data for 19 genes was unavailable, so these categories might not cover all possible influences of circadian genes. However, the genes *Nr1d1* and *Bhlhe40* were implicated in most of these changes, highlighting them as interesting target genes for studies and interventions. More research is clearly needed.
In humans, data was more readily available and showed changes mostly to metabolism and chronotype. Other changes were related to sleep, cognitive ability, personality and different diseases of the digestive tract, immune system and cancer. Thus, circadian genes seem to be implicated in a host of different phenotypes in humans. Here, the relevant genes were much more widely distributed than in the IMPC analysis. Genes of the CCO as well as metabolic and myogenic genes were implicated across the spectrum of phenotypes.
In regard to loss-of-function mutations in humans, the corpus was about evenly distributed between being tolerant or haploinsufficient towards loss-of-function mutations. Most members of the core circadian oscillator proved to be very intolerant toward loss-of-function mutation, while many metabolic and myogenic genes were found tolerant towards loss-of-function mutation, most likely due to other genes being redundant in function. Once again, being among the haploinsufficient circadian genes, the *BHLHE40* gene was among noteworthy genes of the analysis.
In conclusion, my analysis of phenotypes and diseases showed that circadian genes are mostly associated with preweaning lethality as well as changed metabolic activity and body structure in mutated mice. However, data for many genes was lacking, which possibly occluded other potential phenotypes. In humans, my analysis also showed changes to metabolism. Changes in chonotype were also observed next to changes in sleep, cognitive abilities and personality. Other associations connected circadian genes to different diseases of the immune system, digestive tract and cancer. About half of the corpus proved intolerant towards loss of function mutation in humans. This included most members of the core circadian oscillator, which makes sense as it provides the essential framework for circadian gene patterns. Many metabolic and myogenic genes on the other hand were shown to be very tolerant towards loss of function mutation, possibly due to redundancy with other genes.

4.6 The role of the circadian genes of the corpus regarding cancer risk, expression and survival

I also put the circadian genes I investigated through cancer-related analyses, trying to answer the following research questions:

1. Are the genes frequently mutated in human cancer?
2. Is the high expression of the genes associated with survival in 18,000 cases of cancer?
3. How high is the expression of the genes in cancer cell lines?
4. Are the genes differently expressed between embryonal rhabdomyosarcomas caused by *YAP1 S127A* expression in activated satellite cells versus control skeletal muscle?

My analysis of frequently mutated circadian genes identified a prime intervention target: *IDH1*. Studies by Kang et al. (2009) and Metallo et al. (2012) are suggesting that it has high rates of mutation in the same codon over many different cancers and that its metabolic role would make it a very promising target for interventions.

Looking at the influence of circadian genes in cancer survival, I could identify brain cancer neuroblastoma and glioma as being most influenced in survival rates. Four genes regulating the cell cycle and metabolism (*WEE1*, *CSNK1D*, *CCRN4L*, *NFIL3*) were implicated in bad survival rates. The gene *IDH1* was not highly ranked, making it less important in the context of survival compared to cancer onset.

On the other hand, five genes that are part of the core circadian oscillator or interact thightly with it (*RORA*, *CRY2*, *PER3*, *PER2*, *ARNTL*) could be identified as beneficial to survival in most cancer cases. The gene *HLF* could be identified as most beneficial, however, scientific literature was lacking in regard to potential explanations.

I found that in the cancer cell, most circadian genes of the corpus (23 out of 32) are upregulated. Cancers of T- and B- Cells were found to upregulate circadian genes the most. Cancers of the brain, blood, cartilage, lung, nerve tissue, prostate and skin also upregulated circadian genes. Among the upregulated genes were most members of the core circadian oscillator. The forward arm seems to expedite cancer growth, but could also be used in interventions to improve response to chemotherapeutic agents (Gorbacheva et al. 2005). The upregulation of the *PER* genes of the negative arm could be seen as a defensive reaction given their implicated role as cancer suppressors in current scientific literature (Gery et al. 2006; Fu et al. 2002).

Rhabdomyosarcomas caused by *YAP1 S127A* expression downregulated most circadian genes in activated satellite cells compared to control skeletal muscle. Interestingly, it downregulated the *PER* genes, thus disabling these known tumour suppressors. ERMS also downregulated *MYOD1*, a gene normally responsible for differentiation of muscle precursor cells. Yang et al. (2009) suggest it as a potential target for treatment and I hypothesize that treatment could be optimized by good circadian timing.

The *WEE1* gene, which is involved in the cell cycle, was among the two upregulated genes. It was identified as another potential treatment target by Kahen et al. (2016) as its inhibition can lead to ”mitotic progression, DNA strand breaks, mitotic catastrophe, and cell death” (Kahen et al. 2016) in cancer cells.

In short, the only circadian gene that is frequently mutated is *IDH1*. It is already known as a gene with a high rate of mutation within the same codon, causing a host of different cancers. The metabolic gene is a prime target for interventions. When it comes to cancer survial, the investigated circadian genes influence outcomes in cases of brain cancer neuroblastoma and glioma the most. Four cell cycle genes (*WEE1*, *CSNK1D*, *CCRN4L* and *NFIL3*) were associated with bad outcomes, while genes within or close to the core circadian oscillator (*ARNTL*, *PER2*, *PER3*, *CRY2* and *RORA*) generally had positive influence on survival if expressed highly in cancer. The positive role of *HLF* remained unexplained. In cells, cancer upregulates most circadian genes I investigated. The most upregulating cancers were found in T- and B-Cells. The genes of the forward arm of the core circadian oscillator were most upregulated, indicating a role of cancer growth. On the other hand, I hypothesize that the upregulation of *PER* genes, which are known tumour suppressors, could be a defensive reaction. Lastly, in embryonal rhabdomyosarcoma, I discovered that the *PER* genes were downregulated. I hypothesize that the inactivity of these tumour suppressors could be a potential cause for cancer outbreak along with downregulation of the myogenic differentiation gene *MYOD1*. The cell cycle gene *WEE1* was also identified as a potential treatment target due to it being upregulated in rhabdomyosarcoma.

4.7 Practical recommendations in regard to circadian genes and exercise

I regard the scientific method not only as a tool for fundamental research, but also as responsible for giving evidence-based practical recommendations. I will therefore try to answer the following additional research question: Which practical recommendations can you give on the basis of what we know about the circadian genes of your corpus in relation to exercise?
The results of my analyses show a considerable possible influence of exercise on circadian rhyhms over a broad range of exercise- and health-related topics. However, while modern scientific research shows the possibility of using circadian timing to optimize exercise and drug interventions, concrete practical recommendations are still lacking. I will therefore be unable to give recommendations in that direction.
Nevertheless, I consulted scientific literature regarding circadian rhythms in exercise performance to be able to give at least broad recommendations. These recommendations will focus on two distinctive topics:

1. Exercise & Performance Rhythms
2. Countermeasures against performance-decrease due to out-of-phase competitions

1. Exercise & Performance Rhythms There is extensive scientific literature about different rhythms of performance. Most of these measures are the result of in- and out of competition testing across many individuals and must therefore be put into context with chonotype and current rhythm to be useful for the specific individual. Facer-Childs and Brandstaetter (2015) found that “the evaluation of an athlete’s personal best performance requires consideration of circadian phenotype, performance evaluation at different times of day, and analysis of performance as a function of time since entrained awakening” (Facer-Childs and Brandstaetter 2015). Thus, time since entrained awakening and not time of day must be used as the major predictor of peak performance times. I will therefore be forced to limit my recommendations to

general tendencies of time of day. I recommend coaches at least evaluate individual athletes' chonotypes by using an athlete-specific chronometric test like the RBUB chronometric test (available from Dr. Roland Brandstaetter) before putting my recommendations into practice.
In their review "Circadian rhythms in exercise performance: Implications for hormonal and muscular adaptation" Teo, Newton, and McGuigan (2011) summarized the following peak performance times for different exercise attributes according to a collection of other studies:

- Aerobic: Increased physical performance capability and VO2max in the later part of the day
- Anaerobic/Strength/Power: Increased anaerobic fitness, strength and power during the day
- Agility/Coordination: Soccer specific skills were consistent with time-of-day variation and showed increased ability to dribble and more accurate shots later in the day
- Psycho-social: Studies are limited, results demonstrate better performance in physical activity during the time of an individual's chronotypological preference (Teo, Newton, and McGuigan 2011)

Traditionally, core temperature is considered an indicator of circadian performance rhythmicity, in part due to the hypothesis that "muscle glycogen use is augmented by increases in intramuscular temperature" (Starkie et al. 1999). Using extended warmup periods is recommended to counter lower body temperatures due to earlier time of day.
More recent research identified a host of underlying mechanisms with their own distinct circadian profile. However, most of them seem to peak in the later parts of the day (Teo, Newton, and McGuigan 2011). In the context of Facer-Childs and Brandstaetter (2015) I therefore recommend timing workouts to the later individual time of day. Facer-Childs and Brandstaetter (2015), who categorize individuals as either early, intermediate or late circadian phenotypes, place this time:

- At 5.5 and 6 hours after awakening for early and intermediate chronotypes
- At about 11 hours after awakening for late chronotypes

If this is not possible due to an early or conflicting schedule, I recommend using extended warmup periods as traditionally prescribed.
In summary, it seems beneficial to exercise roughly at the same time, with the best time point being dependent on individual circadian timing, i.e. chronotype. Those times are 5.5 hours past awakening for early chronotypes, 6 hours past awakening for intermediate chronotypes and 11 hours past awakening for late chronotypes. If it is not possible to adhere to this recommendation, the literature recommends an extended warmup period to increase core temperature.

2. Countermeasures against performance-decrease Another important factor is performance decrease due to out-of-phase training or competition. Many competive events are scheduled in the morning or have timetables making it hard to discern when one will actually perform.
However, out-of-phase competitions happens especially in cases where the scheduled event is in another time zone, i.e. requires longer travel by airplane, which can in itself cause problems in the form of jetlag.

First of all, I will discuss the case of the competition not requiring international travel, i.e. the effects of jetlag will not be an issue.

2a Competition within the same time zone Chtourou and collegues have studied the topic of time of day effects of training extensively and I consulted four of their studies (Chtourou et al. 2011; Chtourou and Souissi 2012; Chtourou et al. 2012; Chtourou et al. 2014) to gain insight as to how one could entrain oneself optimally for competitions.
In the case of scheduling training times close to competitons, Chtourou et al. (2012) elaborate:

> "For practical considerations, athletes required to compete at a certain time of day (i.e., when the time of competition is known) may be advised to coincide training hours with the time of day at which one's critical performance is planned. If this is impractical (i.e., the time of competition in not known) resistance training program should be scheduled in the morning to minimize the time-of-day effect on physical performances" (Chtourou and Souissi 2012).

They further distinguish between knowing the time of competiton or not knowing it as well as between being able to train once or twice a day (Chtourou et al. 2014):
If the time of competition is known and two sessions per day are possible, athletes are advised to

1. Train anerobically in the morning
2. Train aerobically in the afternoon

If the time of competition is unknown and two sessions per day are possible, athletes are advised to

1. Train anaerobically in the morning
2. Train aerobically in the morning or afternoon

If the time of competition is known and only one session per day is possible, athletes are advised to

1. Train anaerobically at the same time of the competition
2. Train aerobically at the same time of the competition

If the time of competition is unknown and only one session per day is possible, athletes are advised to

1. Train anaerobically in the morning
2. Train aerobically in the morning or afternoon

These different recommendations are based on their finding that in general, morning training reduced circadian rhythm amplitude of sport performances while afternoon training tended to increase it. If the competition takes place in another time zone, i.e. travel might produce jetlag, there are also other factors to consider.

2b Competition within another time zone There was no standardized study approach towards jet-lag, until Spitzer et al. (1997) developed the “Columbia Jet Lag Scale”. It has a high internal consistency with the items fatigue, difficulty in concentrating, clumsiness, decreased alertness in the daytime, difficulty with memory, general weakness, dizziness, lethargy, and daytime sleepiness. These symptoms are developed due to asynchrony of the internal clocks with the external zeitgebers of a travel destination outside ones normal time zone. Waterhouse, Reilly, and Atkinson (1997) already identified light and melatonin levels as the most important factors in this regard, noting that in general:

> “The effect of light depends on the time at which it is presented. Thus, bright light pulses of 3 h centred immediately after the trough of the body temperature rhythm produce a phase advance, pulses centred immediately before the temperature minimum produce a phase delay, and those centred away from it by more than a few hours have little effect” (Waterhouse, Reilly, and Atkinson 1997).

In regard to melatonin, which has a robust circadian rhythm of secretion, starting at about 2100 h and ending at about 0800 h, Waterhouse, Reilly, and Atkinson (1997) explain:

> “Melatonin ingestion produces general effects on the body clock similar to those of light (phase advance, phase delay, no shift), but the timing of its effects is opposite to that of light. Thus, melatonin given in the afternoon tends to advance the body clock, whereas given in the early morning the hormone tends to delay the clock. Since bright light suppresses melatonin secretion, the effects of light and melatonin reinforce each other” (Waterhouse, Reilly, and Atkinson 1997).

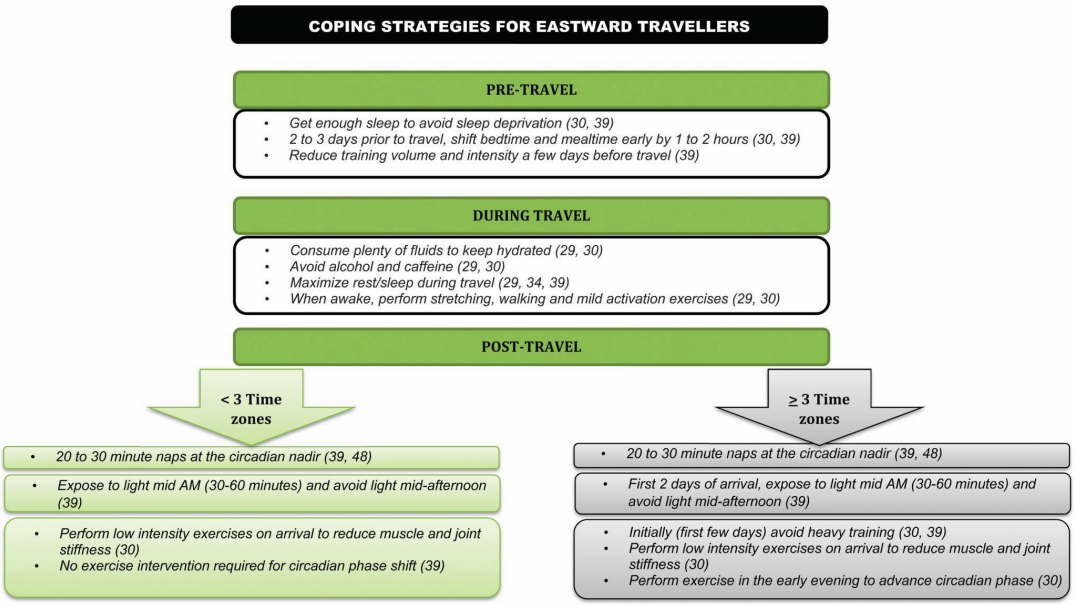
Modern research now knows that the rhythm of melatonin is linked to the activity of the central clock in the SCN (Arendt 2000) and therefore linked to its genetic machinery. Harrington (2010) contextualize the findings of Yamazaki et al. (2000), who found that jet lag varies with tissue location:

> “In the central brain clock, the SCN, Per1 and Per2 lead the system in shifting. Conversely, in the pancreas, Nr1d1 appears to take the lead. The liver circadian clock shows an intermediate pattern” (Yamazaki et al. 2000).

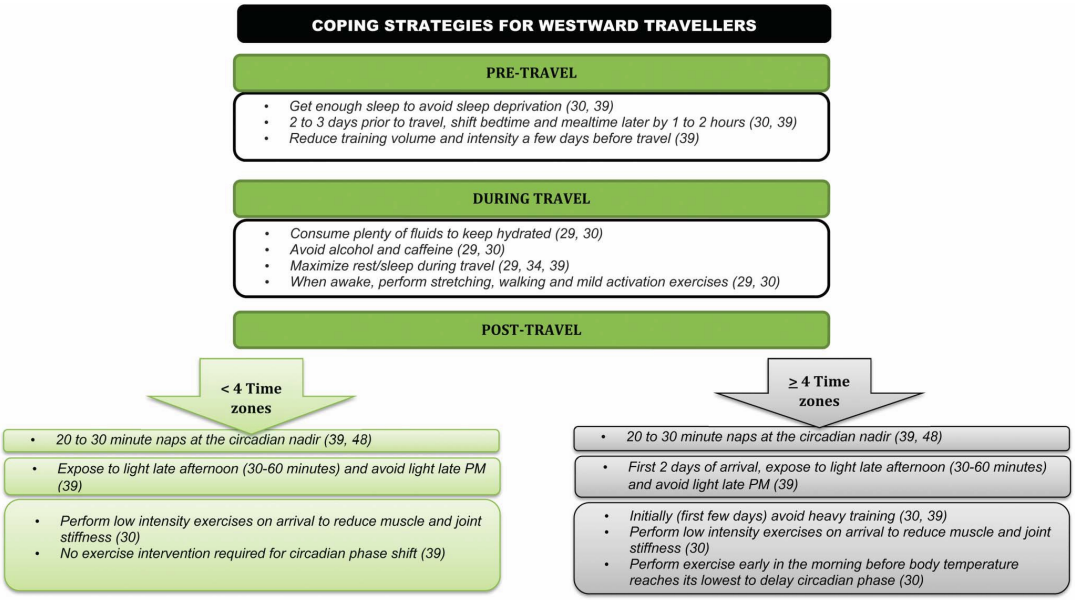
This gives us a genetic explanation why jet lag affects travellers with such different symptoms across our bodies system. To give concrete advice on how to lessen these symptoms, I consulted the review “Managing Performance Throughout Periods of Travel” by Williams et al. (2017), who give extensive evidence-based practical recommendations for the travelling athlete and coaching staff. In general, Williams et al. (2017) explain that “irrespective of the direction of travel, the body’s circadian rhythm can be resynchronized at the rate of approximately one time zone per day”. They further differentiate between eastward and westward travel:

> “During eastwardly travel, there is a need for a circadian phase advance (sleep promotion), which is much more difficult to accommodate compared with a circadian phase delay (sleep deterrence) required for westward travel (16,52). Subsequently, the effect of jet lag remains longer with eastbound compared with westbound travel” (Williams et al. 2017).

To reduce jet lag symptoms in the travelling athlete, Williams et al. (2017) give recommendations based on direction of travel as well as number of time zones. They further divide their recommendations in the phases “pre-travel”, “during travel” and “post travel”. Please consult the following figures (Figure 27a and Figure 27b) for a summary or consult the publication of Williams et al. (2017) directly. There, they outline a set of recommendations. Pre travel recommendations are regarding sleep and exercise as well as bedtime and mealtime shifting before travel. During travel, their recommendations focus on nutrition, sleep and methods to avoid joint stiffness. Recommendations for after travel differ between number of time zones travel, but are generally concerned with the timing of naps, light exposure and exercise.



(a) Coping Strategies for the Eastward Traveller (Williams et al. 2017)



(b) Coping Strategies for the Westward Traveller (Williams et al. 2017)

Figure 27: Coping Strategies for the West- and Eastwad Traveller (Williams et al. 2017)

In summary, we must differentiate between competitions within the same time zone and within another time zone. We also must differentiate whether we know the exact time of the competition or not. If the competition takes place within the same time zone, we only need to differentiate between known and unknown time of competition. If we know the time when we will compete, Chtourou et al. (2014) recommend to time workouts at roughly the same time of day to entrain anaerobic and aerobic systems. If the time of the competition is unknown, it is recommended to use exercise in the morning to reduce circadian rhythm amplitude. If the competition takes place within another time zone, travel by flight is almost always necessary. For this case, Williams et al. (2017) recommend different methods before, during and after travel. These depend on the direction of travel, i.e. eastward or westward.

5 Bioinformatic Methods

5.1 Gene Lists

This thesis investigates the circadian genome in relation to exercise. For this reason, I selected genes from current research for further bioinformatics analysis.

I used two gene lists from a study by McCarthy et al. (2007): Firstly, the list of guide genes used to train the COSOPT algorithm, which is used to identify genes with a circadian pattern of expression (Table 33). According to McCarthy et al. (2007), COSOPT "employs a MMC-β parameter to assess the statistical probability a transcript has a circadian pattern of expression displaying a cosine wave form of period length 20–28 h". Guide genes (see Table 33) are known genes that COSOPT was able to identify correctly as circadian, allowing the researchers to set MMC-β "by providing a conservative balance of the false positive and false negative identifications by COSOPT" (McCarthy et al. 2007). Those results were quantitatively confirmed by Real-Time PCR.

Table 33: Guide Genes according to McCarthy et al. (2007)

Gene Symbol	**MMC-β[a]**
Bmal1	0.013
Dbp	0.013
Rev-erbα	0.017
Nfil3	0.02
Ccrn4l	0.027
Per2	0.04
Rev-erbβ	0.04
Tef	0.064
Cry2	0.098
Rora	0.101
Hlf	0.111
Bhlhb2	0.113
Hnrpu	0.137
Wee1	0.163
Per3	0.197
Clock	0.2

[a] MMC, multiple measures corrected.

McCarthy et al. (2007) then used those guide genes to find 215 genes in skeletal muscle that have a circadian pattern of expression. Among those of interest were “circadian genes involved in lipid metabolism [that] are enzymes that catalyze the synthesis or hydrolysis of fatty acids” (see Table 34).

Table 34: Guide Genes involved with lipid metabolism

Gene Symbol	**Description**
Rora	transcription factor: control of lipid homeostasis in skeletal muscle
Rev-erbβ	transcription factor: lipid and energy homeostasis in skeletal muscle
PGC-1β	transcription factor; regulates expression MCAD, key enzyme in FA oxidation
Myod1	transcription factor; regulates expression of *Ucp3*
Ucp3	transporter: facilitates FA[a] oxidation
Pank1	enzyme: catalyze rate-limiting step in CoA synthesis
Dbt	enzyme: catalyze conversion of α-keto acids into acyl-CoA
Dgat2	enzyme: catalyze final step of triglyceride synthesis
Acat2	enzyme: involved in cholesterol absorption and synthesis
Idh1	enzyme: provide NADPH for cholesterol and FA synthesis
Ces3	enzyme: hydrolysis of triglycerides
Pnpla3	enzyme: triacylglycerol lipase
S3-12	coat protein: involved in packaging of triglycerides
Fabp9	predicted FA binding protein

[a]FA, fatty acid

Since there is a lot of evidence in the literature pointing to a connection between the circadian genome and important genes related to metabolism (Liu et al. 2007; Cho et al. 2012; Sen et al. 2018; Duez and Staels 2009), I will use the genes from this list to further explore the links between the circadian genome in muscle and metabolism. I combined these lists to create the gene corpus used in this thesis:

Table 35: Gene corpus used in this thesis

Gene Symbol	Synonyms
ACAT2	
ARNTL	*BMAL1*
BHLHE40	*BHLHB2*
CCRN4L	*NOCT*
CES3	
CLOCK	
CRY1	
CRY2	
CSNK1D	
CSNK1E	
DBP	*DABP*
DBT	
DGAT2	
FABP9	
HLF	
HNRNPU	*HNRPU*
IDH1	
MYOD1	
NFIL3	
NR1D1	*REV-ERBa*, *EAR-1*
NR1D2	*REV-ERBb*, *EAR-1R*
PANK1	
PER1	
PER2	
PER3	
PGC-1b	*PPARGC1B*
PLIN4	*S3-12*
PNPLA3	
RORA	
TEF	
UCP3	
WEE1	

5.2 String Database

I employed STRING database (see string-db.org (2018)) to find out whether the circadian genes I study are functionally linked and whether they belong to the same signaling pathways. This database includes known and predicted protein-protein interactions, including direct physical as well as indirect functional associations stemming from computational prediction. It also employs knowledge transfers from other organisms and searches for interactions aggregated from other databases.

STRING Database describes the flow of information through the interface of protein interactions, which are biochemically diverse and information-rich. Proteins can influence each other's production, transcriptional

and post transcriptional half-life, exchange reaction products, participate in signal relay mechanisms and jointly contribute toward specific organismal functions (string-db.org 2018).
The protein networks that form from these interactions are well suited for data integration, visualization and molecular discovery. STRING is a comprehensive database which "covers the largest number of organisms and uses the widest breath of input sources, including automated text-mining and computational predictions". The interactions are annotated with benchmarked confidence scores separated per evidence type. Sources can be accessed freely (string-db.org 2018).
These confidence scores are scaled between zero and one, estimating the likelihood of it being biologically meaningful, specific and reproducible. Evidence is divided into seven channels, depending on origin and type of evidence (string-db.org 2018):

1. Experimental: evidence from biochemical, biophysical and genetic experiments
2. Database: evidence asserted by human expert curator
3. Text mining: evidence of co-mentionings of proteins
4. Coexpression: evidence of matched expression patterns
5. Neighborhood: observation of genomic neighborhood
6. Fusion: evidence of organisms where protein orthologs have fused
7. Co-occurrence: evidence of similar phylogenetic distribution of orthologs in a given organism

For my own analysis, I entered the genes of my investigated corpus into StringDB and then extracted the output data into my own excel datasheet. There, I annotated the results and also added additional information from StringDB when necessary.

5.3 ToppGene

I employed ToppGene Suite (toppgene.cchmc.org) to find out whether the genes I study have common features. ToppGene Suite is an online portal for gene list functional enrichment and candidate gene prioritization of novel disease candidates. It uses functional annotations and network analysis to achieve this and strives to identify and prioritize novel disease candidates in the interactome (toppgene.cchmc.org 2018). Their functional annotation based disease candidate gene prioritization

> "uses fuzzy-based similarity measure to compute the similarity between any two genes based on semantic annotations. The similarity scores from individual features are combined into an overall score using statistical meta-analysis. A p-value of each annotation of a test gene is derived by random sampling from the whole genome. The protein-protein interaction network (PPIN) based disease candidate gene prioritization uses social and Web networks analysis algorithms (extended versions of the PageRank and HITS algorithms, and the K-Step Markov method)" (toppgene.cchmc.org 2018).

For further explanation of the work flow and methodology of ToppGene, please consult Figure 28.

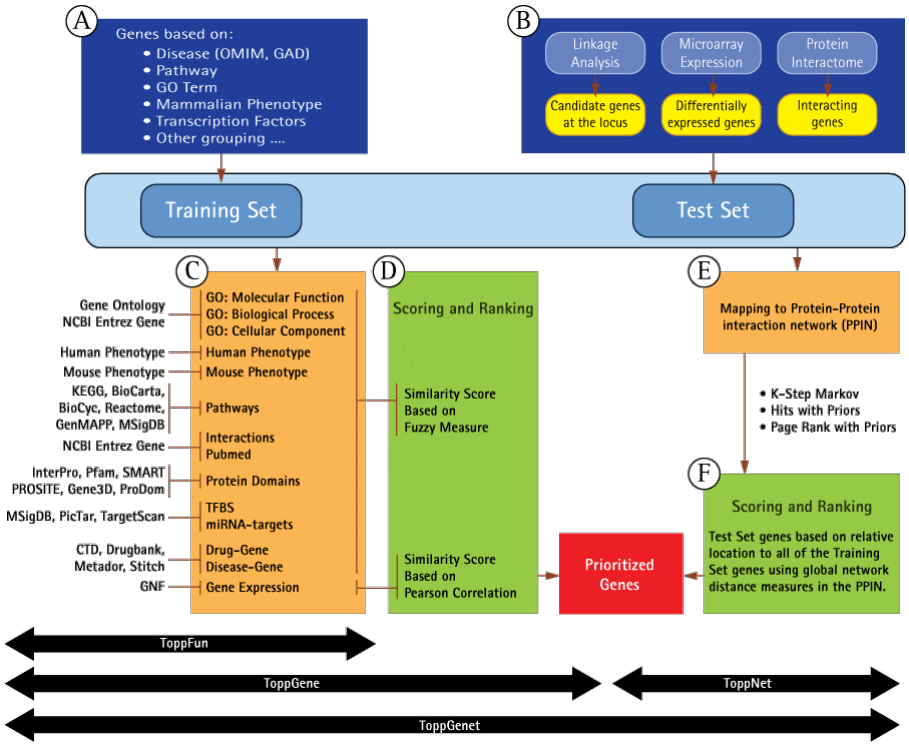


Figure 28: Schematic representation of work flow and methodology in ToppGene Suite applications

"Schematic representation of work flow and methodology in ToppGene Suite applications (A) Genes in the training set are selected based on their attributes or current gene annotations (genes associated with a disease, phenotype, pathway or a GO term). (B) Test gene source can be candidate genes from linkage analysis studies or genes differentially expressed in a particular disease or phenotype or genes from the interactome. (C) ToppFun - Enriched terms of the gene annotations and sequence features, namely, GO: Molecular Function, GO: Biological Process, Mouse Phenotype, Pathways, Protein Interactions, Protein Domains, transcription factor binding sites, miRNA-target genes, disease-gene associations, drug-gene interactions, and Gene Expression, compiled from various data sources, and also used to build training set gene profile. (C and D) ToppGene - A similarity score is generated for each annotation of each test gene by comparing to the enriched terms in the training set of genes. The final prioritized gene list is then computed based on the aggregated values of the fourteen similarity scores. (E and F) ToppNet – Training and test set genes are mapped to protein-protein interaction network. Scoring and ranking of test set genes based on relative location to all of the training set genes using global network-distance measures in the PPIN" (toppgene.cchmc.org 2018).

I entered the genes of my investigated corpus into ToppGene and then extracted the output data into my own excel datasheet for analysis. There, I annotated the results and also added additional information when necessary.

5.4 Gtex Portal

To find out in which tissues the genes I study are expressed, I employed the Genotype-Tissue Expression (GTEx) programs Online Data Resource Gtex Portal (gtexportal.org). It is an online portal for "storing, cataloging, searching, and sharing aggregated level data" about the "mechanisms of gene regulation by studying human gene expression and regulation in multiple tissues", which includes healthy individuals as well as unhealthy ones to gain knowledge about "disease-related perturbations in a variety of human diseases" (gtexportal.org 2018). They are also "examining sexual dimorphisms in gene expression and regulation in multiple tissues" (Lonsdale et al. 2013). The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health,

and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 23.05.2018.
For my research, I entered the genes of my investigated corpus into Gtex Portal and then extracted the output data in the form of data figures for use within my own excel datasheet. There, I annotated the figures and also created tables for data overview.
I focused on the different expression levels of circadian genes across different tissue, with a special focus on how much it is expressed in skeletal muscle relatively to other tissues. This was done to determine how important the genes could be in regard to exercise.

5.5 Cell Secretion

I employed the Human Protein Atlas (proteinatlas.org 2018c) and data from Naba et al. (2016) to find out whether the circadian genes I investigate are predicted to be secreted proteins or extracellular matrix genes. I did so by comparing their gene lists for overlap with mine. For this, I used an online overlapping tool (BaRC 2018).
The Human Protein Atlas is “a Swedish-based program initiated in 2003 with the aim to map all the human proteins in cells, tissues and organs using integration of various omics technologies, including antibody-based imaging, mass spectrometry-based proteomics, transcriptomics and systems biology” (proteinatlas.org 2018c). Their data is provided freely for all researchers to use and is divided into three categories:

1. The tissue atlas, which shows the distribution of the proteins across all major tissues and organs in the human body
2. The cell atlas, which shows the subcellular localization of proteins in single cells
3. The pathology atlas, which shows the impact of protein levels for survival of patients with cancer.

I used the cell atlas, to find out more about the subcellular localization of proteins produced by circadian genes in single cells. It contains mRNA expression profiles from 64 human-derived cell lines that provide insights into different germ layers and tissues. Their data is the basis for my gene list comparison to search for genes that are predicted to be secreted from the cell.
Naba et al. (2016) looked at the extracellular matrix (ECM), “a fundamental component of multicellular organisms that provides mechanical and chemical cues that orchestrate cellular and tissue organization and functions” (Naba et al. 2016). They investigated the ECM by integrating a diverse array of modern proteomics data that was publicly available and reanalyzing it. I used their data to compare it to my gene list to find out whether any of the genes in the corpus are part of the extracellular matrix.

5.6 Online Mendelian Inheritance in Man (OMIM)

Online Mendelian Inheritance in Man (OMIM) (see omim.org (2018a)) is a freely available online compendium of human genes and genetic phenotypes. I used it to find out the current state of research regarding circadian genes. It was started in the early 1960s by Dr. Victor A. McKusick as a catalog of

mendelian traits and disorders and there were twelve book editions published between 1966 and 1998. The online version was created in 1985 and made generally available on the internet in 1987.
It features full-text referenced overviews and contains information on all known mendelian disorders and over 15,000 genes (Hamosh et al. 2005). I entered the genes of my investigated corpus into OMIM and then extracted the text output about current research topics to use it within my own excel datasheet. There, I also created a table for data overview concerning current research topics.

5.7 Exercise and Gene Expression

To find out whether and how the circadian genes I study change their expression after endurance and strength exercise, I consulted the paper “Simplified data access on human skeletal muscle transcriptome responses to differentiated exercise” (Vissing and Schjerling 2014) and its respective simplified access excel data sheet. . They “investigated the human skeletal muscle transcriptome responses to differentiated exercise and non-exercise control intervention” and developed a “straightforward search tool to allow for easy access and interpretation” (Vissing and Schjerling 2014).
The provided spread sheet contains transcriptome data of mRNA expression showing changes in skeletal muscle before and after exercise. The data was obtained with fourteen young untrained and healthy male subjects, who were randomly divided into an endurance and a resistance training group. Six additional subjects served as a non-exercising control group.
All of them were biopsied and then the training groups commenced a 10-week training phase. After this training phase they completed a single-bout trial, while the control group did nothing. Muscle biopsies were taken from all subjects at 0, 2.5, 5 and 22 h after exercise at the same time point of the day.
The study identified several hundred transcripts as specifically responsive to a single exercise type, especially resistance training. At the same time, the findings within the control group show a general time dependent effect from circadian rhythm or dietary status. Here, 192 time-regulated transcripts were found, 131 at 2.5 h and 103 at 5 h. Around 1000 mRNA were significantly changed in both intervention groups and comparison between all groups show great differences between the effects of endurance and resistance exercise on the transcriptome.
Using the datasheet, I searched for data regarding the 32 genes in my corpus and found data for 31 of them (no data for *FABP9*). I produced figures and tables containing the individual data and performed a descriptive overview and extracted the p-values for each condition. If a condition significantly changed the genes expression ($p < 0{,}05$), I extracted the corresponding fold change and added it to my own data sheet. After this, I applied a colored formatting to the data: yellow for significance, green for upregulation and red for downregulation (see supplemental data).

5.8 Exercise and Phosphoproteomics

I consulted two studies: “Global Phosphoproteomic Analysis of Human Skeletal Muscle Reveals a Network of Exercise-Regulated Kinases and AMPK Substrates” (Hoffman et al. 2015) and “A map of the phosphoproteomic alterations that occur after a bout of maximal-intensity contractions” (Potts et al. 2017) to find

out whether the proteins encoded by the genes in my corpus become phosphorylated in skeletal muscle in response to exercise.

Hoffman et al. (2015) found a global approach to phosphoproteomic analysis of human skeletal muscle and uncovered a network of exercise-regulated kinases and AMPK substrates using this approach. This was done via global analysis of protein phosphorylation of skeletal muscle biopsies.

The study used a group of four healthy male subjects and let them perform a single bout of high-intensity cycle exercise and performed muscle biopsies pre- and post-exercise using multiplexed isobaric labelling and phosphopeptide enrichment coupled to tandem mass spectrometry (MS).

They were able to quantify 8511 unique phosphorylation sites, of which around 12% were regulated with exercise. This included more than 900 sites not previously annotated as exercise responsive.

1322 phosphopeptides were significantly regulated with acute exercise, while “only five proteins were quantified as having altered abundance following acute exercise, indicating that changes in phosphopeptide abundance are a direct result of phosphorylation.” Hoffman et al. (2015) provided simplified data access via an excel spreadsheet, which I used for this thesis to examine which of the proteins encoded by the genes investigated here are phosphorylated by high intensity endurance exercise.

Potts et al. (2017) took the approach of global phosphoproteomic analysis developed by Hoffman et al. (2015) and applied it to resistance training, developing a phosphoproteomics workflow to examine the effects of electrically evoked maximal-intensity contractions (MICs) on protein phosphorylation in mouse skeletal muscle. This was done because by using aerobic exercise in their study, Hoffman et al. (2015) did most likely not trigger the signaling events that potentially drive increases in muscle mass, as it can be observed because of resistance training. It is further supported by the fact that the results provided by Hoffmann et al. show that “mTOR signaling was actually inhibited by the aerobic exercise stimulus that was employed in their study” (Potts et al. 2017).

In response to this, Potts et al. (2017) used “two bottom-up proteomics workflows for examining the effect of electrically evoked maximal-intensity contractions (MIC) on protein phosphorylation in mouse skeletal muscle”. These workflows used the traditional proteomics method as well as a centrifugation-based fractionation method to uncover a total of 663 different phosphorylation sites that experienced major MIC-induced alterations.

Potts et al. (2017) provided simplified data access via an excel spreadsheet, which I utilized for this thesis to examine which of the proteins encoded by the genes are phosphorylated by high intensity resistance exercise. I applied sorting and heatmapping functions to further visualize differences and annotate the data. Finally, I created my own tables for the sake of overview.

5.9 Synergist Ablation

To see whether the circadian genes belonging to my gene corpus change their expression because of mechanical overload, I investigated the study *Time course of gene expression during mouse skeletal muscle hypertrophy* by Chaillou et al. (2013), where they subjected young mice to synergist ablation to induce muscle hypertrophy. This was done to identify signaling pathways that are active during hypertrophic response in young muscle.

Chaillou et al. (2013) used the bilateral synergist ablation model to induce hypertrophy of the plantaris muscle in mice by removing the entire soleus muscle together with most of the soleus muscle. Later, plantaris muscles were excised, weighed and in placed in RNAlater. The samples were collected at 1, 3, 5, 7, 10, and 14 days after synergist ablation surgery. They also collected control plantaris muscle, which was collected from mice subjected to sham ablation surgery.
To determine global gene expression patterns, Chaillou et al. (2013) used microarray analysis at each sampling day to analyze the effects of functional overload in mouse plantaris muscle. They used two parameters: Principal component analysis (PCA) of gene expression and the number of differentially expressed genes. They defined three expression patterns of the hypertrophic response: early (1 day), intermediate (3, 5, and 7 days), and late (10 and 14 days) patterns.
This allowed them to identify

> "1) several pathways associated with the immune response and injury/disease during the early gene expression pattern; 2) a number of pathways related to metabolism, mechanotransduction, and mitochondrial function during the intermediate gene expression pattern; and 3) one pathway related to blood coagulation during the late gene expression pattern" (Chaillou et al. 2013).

I consulted the data provided by Chaillou et al. (2013) to conduct my own analysis of the expression of circadian genes after synergist ablation. For this, I extracted the data that was concerned with circadian genes into my own excel datasheet. There, I applied sorting and heatmapping functions to further visualize differences and annotate the data. Finally, I created my own tables for the sake of overview.

5.10 Cardiotoxin Injury

Another factor to be considered regarding exercise and circadian genes is the response to muscle injury. To investigate this, I looked at the study "Genomic Profiling Reveals That Transient Adipogenic Activation Is a Hallmark of Mouse Models of Skeletal Muscle Regeneration" Lukjanenko et al. (2013). There, Lukjanenko and colleagues induced muscle injury either by cardiotoxin or glycerol injection and mapped gene expression changes.
One of the hallmark changes after muscle injury is intramuscular adipogenesis. There is a correlation between ectopic fat in skeletal muscle and the reduction of muscle function and metabolic homeostasis, however the data of Lukjanenko et al. also demonstrate that "transient adipogenic activation is an integral response of skeletal muscle regeneration which is differentially modulated according to physio-pathological conditions through inflammatory and metabolic cues."
For their experiment, adult male C57BL/6J mice were purchased and kept in a 12 hour light 12 hour dark cycle. They hat unrestricted access to regular diet and water and were injured at 12 to 14 weeks old. They were injected with either 25 ml of 50% v/v glycerol or 25 ml of 10 mM cardiotoxin (CTX) into their tibialis anterior muscle. After excision, muscles were cut into two parts: one part was used for total RNA extraction and the other part for histological analysis.

Comparing both methods of triggering regeneration to injury, Lukjanenko et al. (2013) found that both induced In vivo ectopic adipogenesis, but the reaction induced by glycerol had a stronger amplitude and persistence after injury. Also, it triggers a stronger expression of anti-inflammatory cytokines and activates adipogenic networks while repressing fatty acid oxidation. In line with these results, it regulates more genes than CTX at the genome wide level.

Lukjanenko et al. (2013) provide a comprehensive transcriptomic resource of the genes commonly or differentially regulated during muscle regeneration and ectopic muscle adipogenesis, which I used to further investigate the role of circadian genes in this process. For this I extracted the data that was concerned with circadian genes from Lukjanenko et al. (2013) into excel to perform my own analysis. There, I employed sorting and heatmapping functions to further visualize differences. I also annotated the data and created my own tables to provide an overview.

5.11 Berlin High vs Berlin Low

I studied the data collected by Lionikas et al. in their paper “Genetic and genomic analyses of musculoskeletal differences between BEH” (Lionikas et al. 2013) to further illuminate the impact of the circadian genes within my gene corpus. There, they analyzed the transcriptome of two distinct breeds of mice: The Berlin high (BEH) and Berlin low (BEL), who were “selected for divergent growth and differ threefold in body weight” (Lionikas et al. 2013). The Berlin high strain is bred for high body size and muscle mass and is homozygous for the murine myostatin mutation known as compact allele. The Berlin low stain was selected for small body size and muscle mass, allowing me to see whether the genes under investigation are differently expressed between those two.

Lionikas et al. (2013) explored morphological characteristics of the soleus muscle, analyzed the transcriptome of the soleus muscle and initiated quantitative trait locus (QTL) mapping to further illuminate the muscle traits and genomic and gene expression differences between those strains to provide a new model “for the search of genes involved in muscle growth and musculoskeletal morphogenesis” (Lionikas et al. 2013). The study was performed on male animals originating from BEH (n = 16) and BEL (n= 21) strains as well as various other types created by crossing and backcrossing. An assortment of hindlimb muscles, including the gastrocnemius muscle, were dissected at the age of 70 days for histological analysis and analysis of gene expression during the preweaning phase.

During their analysis, Lionikas et al. (2013) were able to map between 38 and 54 million RNA-Sequences to the genome. The transcriptome of gastrocnemius muscle consisted of 13793 expressed genes and the 2148 significantly expressed genes were further analyzed for functional annotation clustering. This revealed a very diverse association with Gene Oncology terms: extracellular matrix, ion binding, regulation of transcription, vascular development, cell adhesion, bone development, transcription repressor activity, muscle development, regulation of striated muscle differentiation and, most importantly, circadian rhythm.

I hope by analyzing the data Lionikas et al. (2013) collected about circadian rhythm genes, I was able to further examine how differently they are expressed between the two strains. For this I extracted the data that was concerned with circadian genes from Lionikas et al. (2013) into excel to perform my own analysis. There, I made use of sorting and heatmapping functions to highlight differences and annotated the data. I

also created my own tables to further outline the data.

5.12 Type 1 vs Type 2b

Chemello and colleagues approached the topic of differential gene expression in single muscle fibers in their publication “Microgenomic Analysis in Skeletal Muscle: Expression Signatures of Individual Fast and Slow Myofibers” (Chemello et al. 2011). They investigated the differences in gene expression between type 1 and type 2b mouse muscle on the single fiber level. The method of scaling down the phenotypic analysis to the single fiber level has the considerable advantage of being able to see gene expression within the muscle fiber, undisturbed by influences of surrounding tissue. Fibers were obtained by excising mouse soleus mucle, a muscle with a balanced composition of type 1 and type 2b fibers. An artificial control was also created by combining and treating soleus and EDL muscles. All muscle samples were obtained from wild-type CD1 mice, housed in a normal environment provided with food and water. After analysis of microarray data, Chemello et al. (2011) found that “a great number of genes are indeed [differentially expressed] between fibers within a muscle”. From their data, Chemello et al. (2011) hypothesize that there are three distinct groups of genes that may share a common regulatory pattern:

1. Genes of fatty acid metabolism regulated by PPAR-α
2. Slow isoforms of contractile proteins controlled by NFATs
3. Genes of oxidative metabolism promoted by PGC-1α

They also generated “the first wide catalogue of gene expression in type 1 and type 2B fibers”. I consulted this catalogue to investigate how circadian genes are expressed between those fiber types. For this I extracted the data that was concerned with circadian genes from Chemello et al. (2011) into excel to perform my own analysis. There, after sorting and heatmapping functions were applied to further visualize differences, I annotated the data. I subsequently generated my own tables for synopsis.

5.13 Muscle young and old proteome

Investigating the role of circadian genes in muscle, I looked at two studies: one that analyzed the differences between gene expression between whole muscle and C2C12 myotubes (Deshmukh et al. 2015) and another that compared concentrations between young and old type 1 and 2A muscle fibers (Murgia et al. 2017).

The first study by Deshmukh et al. (2015) is termed “Deep Proteomics of Mouse Skeletal Muscle Enables Quantitation of Protein Isoforms, Metabolic Pathways, and Transcription Factors”. The authors used a mass spectrometry (MS) workflow to compare the C2C12 cell line model to muscle tissue, identifying 10218 proteins. These included skeletal muscle specific transcription factors like *Myod1* and *Myogenin* as well as circadian clock proteins.

The reasoning behind using the C2C12 model is that this immortalized mouse myoblast cell line can easily be differentiated into myotubes, which allows the user to more easily compare “molecular, biochemical, or pathological changes of skeletal muscle” because of their 2D structure “outside of specialized muscle

functions characteristic of the tissue context" (Deshmukh et al. 2015). C2C12 myotubes were grown in Eagle's minimum essential medium and were harvested after 8 days for proteomic analysis.
The animals providing the tissue samples for this study were fifteen-week-old female C57BL/6 mice who were maintained on a 12:12h light-dark cycle with access to standard chow diet. Their triceps muscles were surgically removed for analysis. After going through the mass spectrometry workflow, Deshmukh et al. conducted a bioinformatic analysis with the Perseus software. The UniProt database provided categorical annotation in the form of "Gene Ontology (GO), biological process (BP), molecular function (MF), and cellular component (CC), as well as participation in a KEGG pathway" (Deshmukh et al. 2015).
I used the data collected by Deshmukh et al. (2015) to further analyze the gene corpus regarding differences between expression levels in C2C12 myotubes and muscle tissue.
Another analysis was conducted by Murgia et al. (2017) in their study "Single Muscle Fiber Proteomics Reveals Fiber-Type-Specific Features of Human Muscle Aging". They compared single young and old muscle slow and fast fibers, also using mass spectrometry. This fiber-type-specific analysis offers "an unprecedented proteomic view into the muscle tissue, which is also devoid of confounding contributions of connective tissue, blood vessels, and nerves, which are present in whole muscle lysates" (Murgia et al. 2017).
The authors obtained muscle biopsies of vastus lateralis from eight subjects either younger (aged 22 to 27) or older (aged 65 to 75), who were healthy and physically active. These needle biopsies consisted of "multiple fiber fascicles obtained from all donors, allowing [them] to perform parallel analyses with different techniques" (Murgia et al. 2017). After the first analysis of the biopsies, where they observed atrophy of fast, but not slow, fibers in the older donors, without specific tension being affected, Murgia et al. (2017) conducted a proteome analysis of the differing proteomic features of young and old slow and fast muscle fibers. I used this data to search for age differences regarding the circadian genes of my corpus. To do so, I extracted the data that was concerned with circadian genes from Murgia et al. (2017) into excel to perform my own analysis. After sorting and heatmapping functions were applied to further visualize differences, I annotated the data. Additionally, I compiled my own tables to represent the data.

5.14 CircaDB

CircaDB is a database of circadian transcriptional profiles from time course expression experiments from mice and humans. It was created in 2012 by Angel Pizarro, Katharina Hayer, Nicholas F. Lahens and John B. Hogenesch to "systematically collect, analyse and visualize circadian expression profiles for bench researchers in a simple and straightforward fashion" (Pizarro et al. 2012).
It consists of multiple datasets from publicly available microarray time course studies. Data from each study were re-analyzed using three circadian rhythm detection algorithms:

- JTK_CYCLE, an algorithm "designed to efficiently identify and characterize cycling variables in large datasets" (Hughes, Hogenesch, and Kornacker 2010).
- Lombe Scargle, "a direct method to treat missing values and unevenly spaced time points" (Glynn, Chen, and Mushegian 2005).

- De Lichtenberg et al. (2004), “a simple permutation-based method that performs better than most existing methods” (De Lichtenberg et al. 2004), i.e. without loss of accuracy by “only model[ing] the shape of the expression profile without taking into account the magnitude of regulation”.

By using these methods, CircaDB could identify “more than 3000 cycling genes in the liver and smaller subsets of rhythmic genes in the pituitary and fibroblasts. [They] are currently expanding our initial studies to identify rhythmic transcripts in over 15 different tissue types including the heart, kidneys and brain” (circadb.org 2018b).
I used CircaDB to find out how and if the genes I investigated are expressed in a circadian fashion within skeletal muscle. To do so, I entered the genes of my investigated corpus into CircaDB and then extracted the output data in the form of data figures for use within my own excel datasheet. There, I annotated the figures by way of qualitative analysis and also created tables for data overview.

5.15 IMPC

The International Mouse Phenotyping Consortium (IMPC) is an organization composed of 19 research institutions and 5 national funders from 11 countries. The following are the objectives of IMPC according to their website (mousephenotype.org 2018):

- “Maintain and expand a world-wide consortium of institutions with capacity and expertise to produce germ line transmission of targeted knockout mutations in embryonic stem cells for 20,000 known and predicted mouse genes.
- Test each mutant mouse line through a broad based primary phenotyping pipeline in all the major adult organ systems and most areas of major human disease.
- Through this activity and employing data annotation tools, systematically aim to discover and ascribe biological function to each gene, driving new ideas and underpinning future research into biological systems.
- Maintain and expand collaborative “networks” with specialist phenotyping consortia or laboratories, providing standardized secondary level phenotyping that enriches the primary dataset, and end-user, project specific tertiary level phenotyping that adds value to the mammalian gene functional annotation and fosters hypothesis driven research; and
- Provide a centralized data centre and portal for free, unrestricted access to primary and secondary data by the scientific community, promoting sharing of data, genotype-phenotype annotation, standard operating protocols, and the development of open source data analysis tools.
- Members of the IMPC may include research centers, funding organizations and corporations”.

“The IMPC is a mass scale mouse phenotyping project, that plans to “expand the ongoing phenotyping effort from hundreds of mouse lines to thousands, with the ultimate aim of tackling the whole genome” (Brown and Moore 2012). The project was launched in September 2011 and “envisages a 10-year project with the aim of completing the generation and phenotyping of 20,000 mouse mutant lines by 2021” (Brown

and Moore 2012). Each member is concerned with the production and phenotyping of mouse mutant lines and deposits data into a Data Centre for analysis and dissemination. After analysis, the data is distributed to the end user (Figure 29).

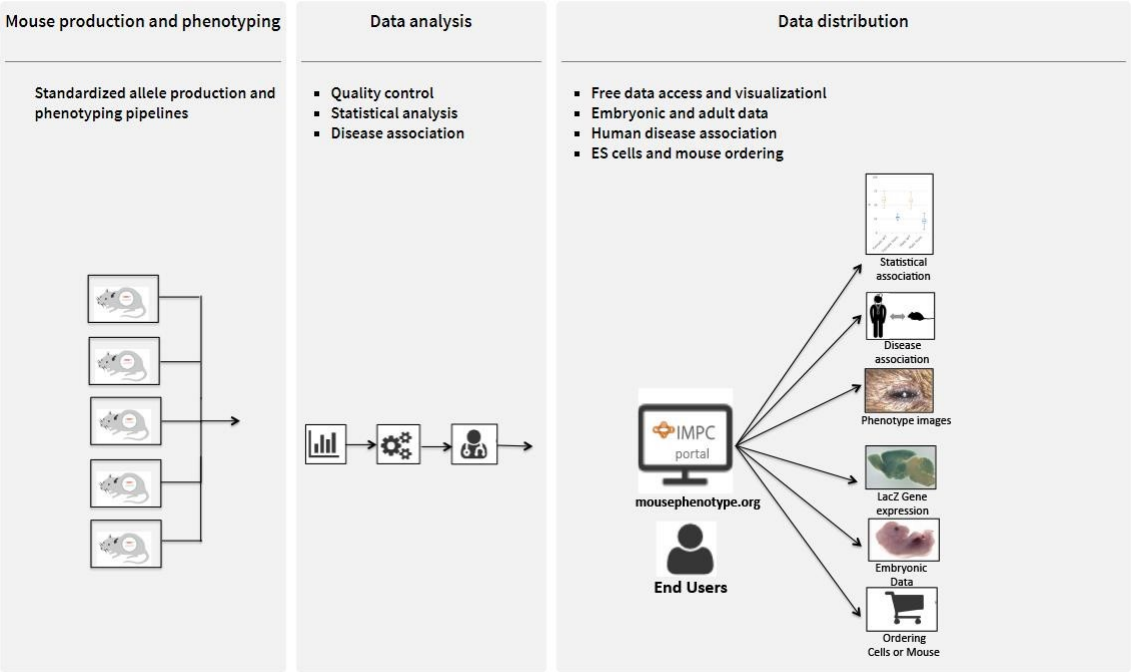


Figure 29: What does the IMPC do? (mousephenotype.org 2018)

I applied the data of the IMPC project to further analyze the genes of my corpus in relation to possible mouse mutant phenotypes. To do so, I entered the genes of my investigated corpus into IMPC and then extracted the output data in the form of data figures for use within my own excel datasheet. There, I annotated the figures and also created tables to further illustrate the data.

5.16 GWAS

The GWAS Catalog is a project by the National Human Genome Research Institute (NHGRI) (see ebi.ac.uk (2018b)), which was created in 2008 to catalogue and summarize published genome-wide association studies. The catalog is freely available and provides data access to published SNP-trait associations.
Eligible GWA studies are identified via literature search and then assessed by experienced molecular biologists. These curators extract reported traits, significant SNP-trait associations and sample metadata (ebi.ac.uk 2018c).
The GWAS Catalog extracts its data from the current scientific literature through weekly PubMed searches. These searches are followed by two levels of curation (ebi.ac.uk 2018a):

1. Extracting of data including SNP, trait and general information by the first curator
2. Double checking the data for accuracy by a secondary curator and validating the data by an automated pipeline.

“Studies are only included into the GWAS catalog if they meet the following criteria:

- A primary GWAS analysis, defined as array-based genotyping and analysis of 100,000+ pre-QC SNPs selected to tag variation across the genome and without regard to gene content must be included.
- GWAS data from published studies which are incorporated into new GWAS analyses are eligible if they meet the other criteria.
- Studies imputing sequencing data to genotyping arrays are eligible if the arrays include sufficient genome-wide coverage so that the post-imputation analysis meets the definition of a GWAS analysis" (ebi.ac.uk 2018a).

I extracted the data that was concerned with circadian genes from GWAS into excel to perform my own analysis. There, I annotated and categorized the it before compiling my own tables for the sake of comprehensibility.

5.17 EXAC

The Exome Aggregation Consortium (ExAC) is a coalition of investigators who combine exome sequencing data from a wide variety of large-scale sequencing projects and make summary data available for the wider scientific community (broadinstitute.org 2018b).
The data set provided on their website spans 60,706 unrelated human individuals, who were sequenced as part of various disease-specific and population genetic studies. It was generated after sequencing data processing, variant calling, quality control and filtering of over 91,000 exomes. After they applied sample filtering, a final dataset was produced. Therein, they identified 10,195,872 candidate sequence variants and applied stringent depth and site/genotype quality filters to define a subset of 7,404,909 high-quality variants, including 317,381 insertions or deletions (indels) (Lek et al. 2016).
With this extensive dataset, ExAC allows researchers to "investigate genic intolerance to PTVs, highlighting 3,230 highly LoF-intolerant genes, 72% of which have no established human disease phenotype in the OMIM or ClinVar databases of observed human genetic mutations" (Lek et al. 2016). I used the data provided by ExAC to further research how the circadian genes of my corpus behave related to loss of function mutations. To do so, I extracted the data that was concerned with circadian genes from EXAC into excel to perform my own analysis. There, I employed sorting and heatmapping functions to further visualize differences before annotating the data. I also created my own tables to make the data more tangible.

5.18 Cancer Mutations

To find out whether the circadian genes of my corpus are frequently mutated in human cancer, I consulted two online databases: TumorPortal (tumorportal.org 2018) and cBio Portal (cbioportal.org 2018).
TumorPortal is a comprehensive database consisting of data by Lawrence et al. (2014), who "analyzed somatic point mutations in exome sequences from 4,742 human cancers and their matched normal-tissue samples across 21 cancer types". They found that "large-scale genomic analysis can identify nearly all known cancer genes in these tumour types" (Lawrence et al. 2014).

> “The samples span 21 tumour types, which include 12 from The Cancer Genome Atlas (TCGA) and 14 from non-TCGA projects at the Broad Institute, with some overlapping tumour types [...]. The number of samples per tumour type varied between 35 and 892” (Lawrence et al. 2014).

The 21 tumor types analyzed were: Acute myeloid leukaemia, bladder, breast, carcinoid, chronic lymphocytic leukaemia, colorectal, diffuse large B-cell lymphoma, endometrial, oesophagal adenocarcinoma, glioblastoma multiforme, head and neck, kidney clear cell, lung adenocarcinoma, lung squamous cell carcinoma, medulloblastoma, melanoma, multiple myeloma, neuroblastoma, ovarian, prostate and rhabdoid tumor.
They analyzed the data

> “through the Broad’s stringent filtering and annotation pipeline to obtain a uniform set of mutation calls (Methods). The data set consists of 3,078,483 somatic single nucleotide variations (SSNVs), 77,270 small insertions and deletions (SINDELs) and 29,837 somatic di-, tri- or oligonucleotide variations (DNVs, TNVs and ONVs, respectively), with an average of 672 per tumour–normal pair. The mutations include 540,831 missense, 207,144 synonymous, 46,264 nonsense, 33,637 splice-site, and 2,294,935 non-coding mutations (used to improve our background model). The analysis has sensitivity of 90% based on the sequencing depth and tumour purity and ploidy” (Lawrence et al. 2014).

cBio Portal is an open-access portal, where cancer genomics data sets can be explored freely and interactively. Its data sets include “data from more than 5,000 tumor samples from 20 cancer studies” (Cerami et al. 2012).

> “[The dataset was] specifically designed to address the unique data integration issues posed by large-scale cancer genomics projects and to make the raw data generated by largescale cancer genomic projects more easily and directly available to the entire cancer research community” (Cerami et al. 2012).

cBio portal is designed to make analyses easy and provides a streamlined 4-step web interface (Cerami et al. 2012):

1. Select a cancer study of interest
2. Select one or more genomic profiles
3. Select a patient case set
4. Select a gene set of interest

The portal currently contains 5 published data sets and 15 provisional The Cancer Genome Atlas (TCGA) data sets. These also contain data from the International Cancer Genome Consortium (ICGC). The data of these projects has not directly been available to researchers before cBio portal.

I used both these research portals to find out more about cancer mutations in circadian genes of my corpus. I did so by entering the genes of my investigated corpus into them and then extracting the output data in the form of data figures for use within my own excel datasheet. Furthermore, I annotated the figures and also created tables as a breakdown.

5.19 Cancer Expression & Survival

I analyzed the dataset provided by Gentles et al. (2015) to find out more about how the expression of the genes of my corpus is associated with survival in cancer. They performed a meta-analysis of expression signatures from 18,000 human tumors with overall survival outcomes across 39 malignancies. This was done by integrating tumor gene expression profiles (GEPs) and overall survival data within a meta-analytical framework that enhances statistical power and improves reproducibility. They also applied the CIBERSORT method to analyze associations between clinical outcomes and abundance of diverse tumor-associated leukocyte (TAL) subsets. The product of their efforts:

> "[A] global pan-cancer map of the landscape of both genes and TALs predicting clinical outcomes, integrating with existing resources such as ENCODE20. Our findings reveal genome-wide molecular portraits of human tumors and identify candidate genes and TALs for prognostic stratification and targeted therapy" (Gentles et al. 2015).

Data about cancer gene expression was obtained by searching NCBI Gene Expression Omnibus (GEO), EBI ArrayExpress, NCI caArray, and Stanford Microarray Database for the terms, survival, prognosis, prognostic, or outcome. After applying statistical analyses, a z-score was obtained. The key indicator in their analysis is the "Unweighted meta-z (all cancers)" value. Positive values indicate that high expression is associated with poor survival. Negative values suggest that high expression is associated with good survival.
I used the various z-scores generated by Gentles et al. (2015) to find out in which way the expression of circadian genes can be beneficial or detrimental to survival after cancer outbreak in human individuals. To do so, I extracted the data that was concerned with circadian genes from their datasheet into excel to perform my own analysis. Here I utilized sorting and heatmapping functions to further visualize differences and annotate the data. In addition, I designed my own tables for the sake of clarity.

5.20 Cancer Cell Line Encyclopedia

To find out in what cell lines the genes of my corpus are expressed, I consulted the Cancer Cell Line Encyclopedia (CCLE). It is a compilation of gene expression, chromosomal copy number and parallel sequencing data from 947 human cancer cell lines. If combined with pharmacological profiles for 24 anticancer drugs across 479 of the cell lines, this collection allows identification of genetic, lineage, and gene-expression-based predictors of drug sensitivity (Barretina et al. 2012). To create the Cancer Cell Line Encyclopedia, Barretina et al. (2012) proceeded as follows:

> “A total of 947 independent cancer cell lines were profiled at the genomic level (data available at http://www.broadinstitute.org/ccle and Gene Expression Omnibus (GEO) using accession number GSE36139) and compound sensitivity data were obtained for 479 lines (SupplementaryTable11). Mutation information was obtained both by using massively parallel sequencing of .1,600 genes (Supplementary Table 12) and by mass spectrometric genotyping (OncoMap), which interrogated 492 mutations in 33 known oncogenes and tumour suppressors. Genotyping/copy number analysis was performed using Affymetrix Genome-Wide Human SNP Array 6.0 and expression analysis using the GeneChip Human Genome U133 Plus 2.0 Array” (Barretina et al. 2012).

I reanalyzed the data created this way in order to find out in what cell lines my researched circadian genes are expressed. To do so, I entered the genes of my investigated corpus into CCLE and then extracted the output data in the form of data figures for use within my own excel datasheet. There, I annotated the figures and also produced tables to further outline the data.

5.21 YAP-ERMS

To find out whether the circadian genes of my corpus are differently expressed between embryonal rhabdomyosarcomas caused by *YAP1 S127A* expression in activated satellite cells compared to control skeletal muscle, I turned to the data created by Tremblay et al. (2014).
They investigated the role of the Hippo pathway effector *YAP1* in soft tissue sarcomas and were able to identify *YAP1* as a potent *ERMS* oncogenic driver and a promising target for differentiation therapy. Thus, they assessed the abundance and cellular localization of *YAP1* protein in 78 ARMS and 196 ERMS samples.
To test if high *YAP1* activity in RMS is causal, they generated doxycycline-inducible (DOX-inducible) *Myf5-hYAP1 S127A* mice, who were then treated with DOX to “induce the overexpression of the TetO-hYAP1 S127A transgene in the mature skeletal musculature, as well as in adult quiescent and activated satellite cells and their myoblast progeny” and were then further analyzed.
I adopted their collected data and reanalyzed it to find out more about how my selection of circadian genes is expressed under the influence of *YAP*-induced rhabdomyosarcoma. Accordingly, I extracted the data that was concerned with circadian genes from Tremblay et al. (2014) into excel to perform my own analysis. There, I employed sorting and heatmapping functions to further contrast differences and annotated the data. Lastly, I devised my own tables to provide an outline.
This final phase concludes the last experimental step of this thesis. I will therefore move on to my closing comments.

6 Closing

The objective of this thesis has been to establish the current state of scientific knowledge about a set of circadian rhythm genes, to reanalyze pre-existing studies about them by way of bioinformatic analysis and to give recommendations for further studies to close gaps in our knowledge about circadian genetics.
Chapter 2 provided an overview of the theoretical basics of circadian genetics, including basic mechanisms of transcription and translation as well as the genetic loops that form the circadian rhythm.
Chapter 3 included the results of all bioinformatic analyses performed on the topics of:

1. the general structure and localization of circadian genes
2. the effect of exercise on circadian genes
3. the role of circadian genes in muscle fiber
4. the phenotypes and diseases associated with circadian genes
5. the role of circadian genes in cancer

My analyses could recreate the well established structural characteristics of circadian gene loops as well as highlight lesser known interactions between genes of the forward arm and myogenic and metabolic genes. The secretion of two genes, *ARNTL* and *CRY1*, was uncovered. With regard to the influence of both resistance and endurance exercise, I could detect changes in expression of both forward and negative arm as well as key metabolic genes. A set of metabolic genes was influenced both in their expression as well as phosphorylation. Looking at myogenic processes due to mechanical overload or injury, I could uncover overlapping patterns of upregulation for the modulator gene *BHLHE40* as well as different metabolic genes. Foward arm and myogenic gene activity differed between conditions. Muscle properties varied with increased foward arm activity in hypertrophied mice and increased negative arm and metabolic gene activity in less hypertrophied mice as well as increased metabolic gene activity in slower and older fibers. Circadian expression patterns were also analysed and a general pattern could be detected, including interesting interactions between *MYOD1*, *UCP3*, *RORA* and *NR1D1*. The analysis of associated phenotypes and diseases revealed influences of circadian genes on survivability when it comes to core oscillator genes and associations with metabolic diseases. Furthermore, phase-shift based psychological disorders could be uncovered. Lastly, in regard to cancer, the frequently mutated circadian gene *IDH1* came into focus but the activity of different genes of the core oscillator, cell cycle and metabolism proved highly relevant in different contexts for cancer treatment and survival.
Chapter 4 also included a section on practical recommendations which could be given based on our current state of research. It focused on the areas of general performance rhythms for regular exercise as well as countermeasures against performance decrease when attending a competition that is outside of ones normal rhythm, be it because of scheduling or international travel. Generally, training at times close to competition time were found to be beneficial. In the absence of knowledge of these times, training in the morning or afternoon (depending on anaerobic or aerobic type of performance) was recommended. For competitions outside of one's own time zone requiring international travel, a set of evidence-based

recommendations depending on the direction of travel created by another author (Williams et al. 2017) was presented.

Chapter 5 dealed with the different methodologies of the studies used for bioinformatic reanalysis as well as my own method of performing said analyses. Here I elaborated on the many different methodologies of the studies and databases that form the foundation of this thesis.

After conducting my research on the genetic machinery of circadian rhythms within the context of exercise, I come to the conclusion that big strides have been made by the scientific community since its beginnings in the last century. Yet, practical applications are still scarce. As soon as we can apply our findings to the timing of drug administration or exercise interventions, chrono-therapy and chrono-exercise will finally become reality.

References

Abbas, Tarek and Anindya Dutta (2009). “p21 in cancer: intricate networks and multiple activities”. In: *Nature Reviews Cancer* 9.6, p. 400.

Adams, Gregory R (2006). “Satellite cell proliferation and skeletal muscle hypertrophy”. In: *Applied Physiology, Nutrition, and Metabolism* 31.6, pp. 782–790.

Adams, Gregory R, Fadia Haddad, and Kenneth M Baldwin (1999). “Time course of changes in markers of myogenesis in overloaded rat skeletal muscles”. In: *Journal of Applied Physiology* 87.5, pp. 1705–1712.

Al Fayi, Majed Saad et al. (2016). “The increased expression of fatty acid-binding protein 9 in prostate cancer and its prognostic significance”. In: *Oncotarget* 7.50, p. 82783.

Andreani, Tomas S et al. (2015). “Genetics of circadian rhythms”. In: *Sleep medicine clinics* 10.4, pp. 413–421.

Andrews, Jessica L et al. (2010). “CLOCK and BMAL1 regulate MyoD and are necessary for maintenance of skeletal muscle phenotype and function”. In: *Proceedings of the National Academy of Sciences*, p. 201014523.

Arendt, Josephine (2000). *Melatonin, circadian rhythms, and sleep*.

Argilés, Josep M et al. (2014). “Cancer cachexia: understanding the molecular basis”. In: *Nature reviews Cancer* 14.11, p. 754.

Bae, Kiho et al. (2001). “Differential functions of mPer1, mPer2, and mPer3 in the SCN circadian clock”. In: *Neuron* 30.2, pp. 525–536.

Baehr, Erin K, Louis F Fogg, and Charmane I Eastman (1999). “Intermittent bright light and exercise to entrain human circadian rhythms to night work”. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 277.6, R1598–R1604.

BaRC (2018). *Bioinformatics & Research Computing*. URL: `http://barc.wi.mit.edu/tools/` (visited on 09/23/2018).

Barretina, Jordi et al. (2012). “The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity”. In: *Nature* 483.7391, p. 603.

Bass, Joseph and Joseph S Takahashi (2010). “Circadian integration of metabolism and energetics”. In: *Science* 330.6009, pp. 1349–1354.

Bent, Martin J Van den et al. (2010). “IDH1 and IDH2 mutations are prognostic but not predictive for outcome in anaplastic oligodendroglial tumors: a report of the European Organization for Research and Treatment of Cancer Brain Tumor Group”. In: *Clinical Cancer Research*, pp. 1078–0432.

Bjarnason, Georg A et al. (2001). “Circadian expression of clock genes in human oral mucosa and skin: association with specific cell-cycle phases”. In: *The American journal of pathology* 158.5, pp. 1793–1801.

Brand, Martin D and Telma C Esteves (2005). “Physiological functions of the mitochondrial uncoupling proteins UCP2 and UCP3”. In: *Cell metabolism* 2.2, pp. 85–93.

broadinstitute.org (2018a). *Exoma Aggregation Consortium*. URL: `http://exac.broadinstitute.org/faq` (visited on 09/23/2018).

– (2018b). *Exome Aggregation Consortium*. URL: `http://exac.broadinstitute.org/` (visited on 09/23/2018).

Brown, Steve DM and Mark W Moore (2012). *Towards an encyclopaedia of mammalian gene function: the International Mouse Phenotyping Consortium*.

Buxton, Orfeu M et al. (2003). “Exercise elicits phase shifts and acute alterations of melatonin that vary with circadian phase”. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 284.3, R714–R724.

cancer.org (2018). *American Cancer Society*. URL: https://www.cancer.org/cancer/rhabdomyosarcoma/about/what-is-rhabdomyosarcoma.html (visited on 09/23/2018).

Cases, Sylvaine et al. (2001). "Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members". In: *Journal of Biological Chemistry*.

cbioportal.org (2018). *cBioPortal*. URL: http://www.cbioportal.org/ (visited on 09/23/2018).

Cerami, Ethan et al. (2012). *The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data*.

Chaillou, Thomas et al. (2013). "Time course of gene expression during mouse skeletal muscle hypertrophy". In: *Journal of applied physiology* 115.7, pp. 1065–1074.

Chawla, AJAY and Mitchell A Lazar (1993). "Induction of Rev-ErbA alpha, an orphan receptor encoded on the opposite strand of the alpha-thyroid hormone receptor gene, during adipocyte differentiation." In: *Journal of Biological Chemistry* 268.22, pp. 16265–16269.

Chemello, Francesco et al. (2011). "Microgenomic analysis in skeletal muscle: expression signatures of individual fast and slow myofibers". In: *PLoS One* 6.2, e16807.

Chen, Shou-Tung et al. (2005). "Deregulated expression of the PER1, PER2 and PER3 genes in breast cancers". In: *Carcinogenesis* 26.7, pp. 1241–1246.

Chen, Weiqin et al. (2013). "Inactivation of Plin4 downregulates Plin5 and reduces cardiac lipid accumulation in mice". In: *American Journal of Physiology-Endocrinology and Metabolism* 304.7, E770–E779.

Cho, Han et al. (2012). "Regulation of circadian behaviour and metabolism by REV-ERB-α and REV-ERB-β". In: *Nature* 485.7396, p. 123.

Chtourou, H et al. (2014). "Temporal Specificity of Training: An Update. J Athl Enhancement 3: 3". In: *The Journal of Strength & Conditioning Research* 4, pp. 90–100.

Chtourou, Hamdi and Nizar Souissi (2012). "The effect of training at a specific time of day: a review". In: *The Journal of Strength & Conditioning Research* 26.7, pp. 1984–2005.

Chtourou, Hamdi et al. (2011). "Diurnal variation in Wingate-test performance and associated electromyographic parameters". In: *Chronobiology international* 28.8, pp. 706–713.

Chtourou, Hamdi et al. (2012). "The effect of strength training at the same time of the day on the diurnal fluctuations of muscular anaerobic performances". In: *The Journal of Strength & Conditioning Research* 26.1, pp. 217–225.

Chung, Shih Ying et al. (2015). "Bhlhe40 represses PGC-1α activity on metabolic gene promoters in myogenic cells". In: *Molecular and cellular biology*, MCB–00387.

circadb.org (2018a). *Circadian Expression Profiles Data Base*. URL: http://circadb.hogeneschlab.org/ (visited on 09/23/2018).

– (2018b). *Circadian Expression Profiles Data Base*. URL: http://circadb.hogeneschlab.org/about (visited on 09/23/2018).

Clancy, Suzanne and William Brown (2008). "Translation: DNA to mRNA to protein". In: *Nature Education* 1.1, p. 101.

Collins, P et al. (2002). "Muscle UCP-3 mRNA levels are elevated in weight loss associated with gastrointestinal adenocarcinoma in humans". In: *British journal of cancer* 86.3, p. 372.

Colwell, Christopher S (2011). "Linking neural activity and molecular oscillations in the SCN". In: *Nature Reviews Neuroscience* 12.10, p. 553.

Conte, Maria et al. (2013). "Increased Plin2 expression in human skeletal muscle is associated with sarcopenia and muscle weakness". In: *PLoS One* 8.8, e73709.

Currie, Erin et al. (2013). "Cellular fatty acid metabolism and cancer". In: *Cell metabolism* 18.2, pp. 153–161.

Dalen, Knut Tomas et al. (2004). “Adipose Tissue Expression of the Lipid Droplet–Associating Proteins S3-12 and Perilipin Is Controlled by Peroxisome Proliferator–Activated Receptor-γ”. In: *Diabetes* 53.5, pp. 1243–1252.

Damiola, Francesca et al. (2000). “Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus”. In: *Genes & development* 14.23, pp. 2950–2961.

De Lichtenberg, Ulrik et al. (2004). “Comparison of computational methods for the identification of cell cycle-regulated genes”. In: *Bioinformatics* 21.7, pp. 1164–1171.

DeBruyne, Jason P et al. (2006). “A clock shock: mouse CLOCK is not required for circadian oscillator function”. In: *Neuron* 50.3, pp. 465–477.

Delezie, Julien et al. (2016). “Rev-erbα in the brain is essential for circadian food entrainment”. In: *Scientific reports* 6, p. 29386.

Deshmukh, Atul S et al. (2015). “Deep proteomics of mouse skeletal muscle enables quantitation of protein isoforms, metabolic pathways, and transcription factors”. In: *Molecular & Cellular Proteomics* 14.4, pp. 841–853.

Desotelle, Joshua A, Melissa J Wilking, and Nihal Ahmad (2012). “The circadian control of skin and cutaneous photodamage”. In: *Photochemistry and photobiology* 88.5, pp. 1037–1047.

Dimitrov, Lilia et al. (2015). “New developments in the pathogenesis and therapeutic targeting of the IDH1 mutation in glioma”. In: *International journal of medical sciences* 12.3, p. 201.

Dityatev, Alexander, Melitta Schachner, and Peter Sonderegger (2010). “The dual role of the extracellular matrix in synaptic plasticity and homeostasis”. In: *Nature Reviews Neuroscience* 11.11, p. 735.

Downes, Michael, Amanda J Carozzi, and GE Muscat (1995). “Constitutive expression of the orphan receptor, Rev-erbA alpha, inhibits muscle differentiation and abrogates the expression of the myoD gene family.” In: *Molecular Endocrinology* 9.12, pp. 1666–1678.

Drolet, DW et al. (1991). “TEF, a transcription factor expressed specifically in the anterior pituitary during embryogenesis, defines a new class of leucine zipper proteins.” In: *Genes & Development* 5.10, pp. 1739–1753.

Dudek, Michal and Qing-Jun Meng (2014). “Running on time: the role of circadian clocks in the musculoskeletal system”. In: *Biochemical journal* 463.1, pp. 1–8.

Duez, Hélène and Bart Staels (2009). “Rev-erb-α: an integrator of circadian rhythms and metabolism”. In: *Journal of applied physiology* 107.6, pp. 1972–1980.

ebi.ac.uk (2018a). *GWAS*. URL: `https://www.ebi.ac.uk/gwas/docs/methods` (visited on 09/23/2018).

– (2018b). *GWAS Catalog*. URL: `https://www.ebi.ac.uk/gwas/home` (visited on 09/23/2018).

– (2018c). *GWAS Catalog*. URL: `https://www.ebi.ac.uk/gwas/docs/about` (visited on 09/23/2018).

Etchegaray, Jean-Pierre et al. (2009). “Casein kinase 1 delta regulates the pace of the mammalian circadian clock”. In: *Molecular and cellular biology* 29.14, pp. 3853–3866.

Facer-Childs, Elise and Roland Brandstaetter (2015). “The impact of circadian phenotype and time since awakening on diurnal performance in athletes”. In: *Current Biology* 25.4, pp. 518–522.

Forman, Barry Marc et al. (1994). “Cross-talk among ROR alpha 1 and the Rev-erb family of orphan nuclear receptors.” In: *Molecular endocrinology* 8.9, pp. 1253–1261.

Friedrich, Timo et al. (2011). “Mutation of zebrafish dihydrolipoamide branched-chain transacylase E2 results in motor dysfunction and models maple syrup urine disease”. In: *Disease models & mechanisms*, pp. dmm–008383.

Froy, Oren (2009). “Metabolism and circadian rhythms—implications for obesity”. In: *Endocrine reviews* 31.1, pp. 1–24.

– (2011). “Circadian rhythms, aging, and life span in mammals”. In: *Physiology* 26.4, pp. 225–235.

Fu, Loning et al. (2002). “The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo”. In: *cell* 111.1, pp. 41–50.

Gaddameedhi, Shobhan et al. (2011). “Control of skin cancer by the circadian rhythm”. In: *Proceedings of the National Academy of Sciences* 108.46, pp. 18790–18795.

genecards.org (2018a). URL: https://www.genecards.org/cgi-bin/carddisp.pl?gene=DEC1&keywords=DEC1 (visited on 09/23/2018).

– (2018b). URL: https://www.genecards.org/cgi-bin/carddisp.pl?gene=DBP&keywords=dbp (visited on 09/23/2018).

– (2018c). URL: https://www.genecards.org/cgi-bin/carddisp.pl?gene=HLF&keywords=hlf (visited on 09/23/2018).

– (2018d). URL: https://www.genecards.org/cgi-bin/carddisp.pl?gene=RORA&keywords=rora (visited on 09/23/2018).

Gentles, Andrew J et al. (2015). “The prognostic landscape of genes and infiltrating immune cells across human cancers”. In: *Nature medicine* 21.8, p. 938.

Gery, Sigal et al. (2006). “The circadian gene per1 plays an important role in cell growth and DNA damage control in human cancer cells”. In: *Molecular cell* 22.3, pp. 375–382.

Glynn, Earl F, Jie Chen, and Arcady R Mushegian (2005). “Detecting periodic patterns in unevenly spaced gene expression time series using Lomb–Scargle periodograms”. In: *Bioinformatics* 22.3, pp. 310–316.

Gorbacheva, Victoria Y et al. (2005). “Circadian sensitivity to the chemotherapeutic agent cyclophosphamide depends on the functional status of the CLOCK/BMAL1 transactivation complex”. In: *Proceedings of the National Academy of Sciences* 102.9, pp. 3407–3412.

Goto, Masahide et al. (2017). “Role of a circadian-relevant gene NR1D1 in brain development: possible involvement in the pathophysiology of autism spectrum disorders”. In: *Scientific Reports* 7, p. 43945.

Groot, CJ et al. (1977). “A patient with severe neurologic symptoms and acetoacetyl-CoA thiolase deficiency.” In: *Pediatric research* 11.10 Pt 2, pp. 1112–1116.

gtexportal.org (2018). URL: https://gtexportal.org/home/ (visited on 09/23/2018).

Hamosh, Ada et al. (2005). “Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders”. In: *Nucleic acids research* 33.suppl_1, pp. D514–D517.

Harrington, Mary (2010). “Location, location, location: important for jet-lagged circadian loops”. In: *The Journal of clinical investigation* 120.7, pp. 2265–2267.

Hayes, Lawrence D, Gordon F Bickerstaff, and Julien S Baker (2010). “Interactions of cortisol, testosterone, and resistance training: influence of circadian rhythms”. In: *Chronobiology international* 27.4, pp. 675–705.

Hill, William David et al. (2017). “Genetic contribution to two factors of neuroticism is associated with affluence, better health, and longer life”. In: *bioRxiv*, p. 146787.

Hoffman, Nolan J et al. (2015). “Global phosphoproteomic analysis of human skeletal muscle reveals a network of exercise-regulated kinases and AMPK substrates”. In: *Cell metabolism* 22.5, pp. 922–935.

Hoopes, L (2008). “Introduction to the gene expression and regulation topic room”. In: *Nature Education* 1.1, p. 160.

Horne, Jim A and Olov Östberg (1976). “A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms.” In: *International journal of chronobiology*.

Hsiao, Sheng P et al. (2009). “P/CAF rescues the Bhlhe40-mediated repression of MyoD transactivation”. In: *Biochemical Journal* 422.2, pp. 343–352.

Hughes, Michael E, John B Hogenesch, and Karl Kornacker (2010). “JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets”. In: *Journal of biological rhythms* 25.5, pp. 372–380.

Ikonen, Elina (2006). “Mechanisms for cellular cholesterol transport: defects and human disease”. In: *Physiological reviews* 86.4, pp. 1237–1261.

Jarjour, Nicholas N et al. (2018). *The transcription factor Bhlhe40 is a novel regulator of large peritoneal macrophages and type 2 immunity*.

Jordan, Sabine D et al. (2017). “CRY1/2 selectively repress PPARδ and limit exercise capacity”. In: *Cell metabolism* 26.1, pp. 243–255.

Juda, Myriam, Céline Vetter, and Till Roenneberg (2013). “The Munich chronotype questionnaire for shift-workers (MCTQShift)”. In: *Journal of biological rhythms* 28.2, pp. 130–140.

Kahen, Elliot et al. (2016). “Identification of clinically achievable combination therapies in childhood rhabdomyosarcoma”. In: *Cancer chemotherapy and pharmacology* 78.2, pp. 313–323.

Kamei, Yasutomi et al. (2003). “PPARγ coactivator 1β/ERR ligand 1 is an ERR protein ligand, whose expression induces a high-energy expenditure and antagonizes obesity”. In: *Proceedings of the National Academy of Sciences* 100.21, pp. 12378–12383.

Kang, Mi Ran et al. (2009). “Mutational analysis of IDH1 codon 132 in glioblastomas and other common cancers”. In: *International journal of cancer* 125.2, pp. 353–355.

Kashiwada, Masaki et al. (2011). “NFIL3/E4BP4 controls type 2 T helper cell cytokine expression”. In: *The EMBO journal* 30.10, pp. 2071–2082.

Kawai, Masanobu et al. (2010). “A circadian-regulated gene, Nocturnin, promotes adipogenesis by stimulating PPAR-γ nuclear translocation”. In: *Proceedings of the National Academy of Sciences* 107.23, pp. 10508–10513.

Kellogg, Douglas R (2003). “Wee1-dependent mechanisms required for coordination of cell growth and cell division”. In: *Journal of cell science* 116.24, pp. 4883–4890.

Kim, Yong Hoon et al. (2018). “Rev-erbα dynamically modulates chromatin looping to control circadian gene transcription”. In: *Science* 359.6381, pp. 1274–1277.

Kusuda, Jun et al. (1996). “Sequence analysis of the cDNA for the human casein kinase I δ (CSNK1D) gene and its chromosomal localization”. In: *Genomics* 32.1, pp. 140–143.

Lack, Leon et al. (2009). “Chronotype differences in circadian rhythms of temperature, melatonin, and sleepiness as measured in a modified constant routine protocol”. In: *Nature and science of sleep* 1, p. 1.

Lau, Patrick et al. (2004). “ROR α regulates the expression of genes involved in lipid homeostasis in skeletal muscle cells: Caveolin-3 and CPT-1 are direct targets of ROR”. In: *Journal of Biological Chemistry*.

Lawrence, Michael S et al. (2014). “Discovery and saturation analysis of cancer genes across 21 tumour types”. In: *Nature* 505.7484, p. 495.

Lek, Monkol et al. (2016). “Analysis of protein-coding genetic variation in 60,706 humans”. In: *Nature* 536.7616, p. 285.

Lin, Jiandie, Christoph Handschin, and Bruce M Spiegelman (2005). “Metabolic control through the PGC-1 family of transcription coactivators”. In: *Cell metabolism* 1.6, pp. 361–370.

Lionikas, Arimantas et al. (2013). “Genetic and genomic analyses of musculoskeletal differences between BEH and BEL strains”. In: *Physiological genomics* 45.20, pp. 940–947.

Liu, Chang et al. (2007). “Transcriptional coactivator PGC-1α integrates the mammalian clock and energy metabolism”. In: *Nature* 447.7143, p. 477.

Lonsdale, John et al. (2013). “The genotype-tissue expression (GTEx) project”. In: *Nature genetics* 45.6, p. 580.

Lowrey, Phillip L and Joseph S Takahashi (2011). “Genetics of circadian rhythms in Mammalian model organisms”. In: *Advances in genetics*. Vol. 74. Elsevier, pp. 175–230.

Lukjanenko, Laura et al. (2013). “Genomic profiling reveals that transient adipogenic activation is a hallmark of mouse models of skeletal muscle regeneration”. In: *PloS one* 8.8, e71084.

Lundberg, Tommy R et al. (2016). “Aerobic exercise augments muscle transcriptome profile of resistance exercise”. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 310.11, R1279–R1287.

Mansour, HA et al. (2006). “Association study of eight circadian genes with bipolar I disorder, schizoaffective disorder and schizophrenia”. In: *Genes, brain and behavior* 5.2, pp. 150–157.

Mao, Yingying et al. (2015). “The circadian gene CRY2 is associated with breast cancer aggressiveness possibly via epigenomic modifications”. In: *Tumor Biology* 36.5, pp. 3533–3539.

Mazzoccoli, Gianluigi et al. (2012). “Association study of a polymorphism in clock gene PERIOD3 and risk of inflammatory bowel disease”. In: *Chronobiology international* 29.8, pp. 994–1003.

string-db.org (2018). *String Consortium*. URL: https://string-db.org/ (visited on 09/23/2018).

McCarthy, John J et al. (2007). “Identification of the circadian transcriptome in adult mouse skeletal muscle”. In: *Physiological genomics* 31.1, pp. 86–95.

McClung, Colleen A (2007). “Circadian genes, rhythms and the biology of mood disorders”. In: *Pharmacology & therapeutics* 114.2, pp. 222–232.

McGowan, Clare H and Paul Russell (1995). “Cell cycle regulation of human WEE1.” In: *The EMBO journal* 14.10, pp. 2166–2175.

Mengatto, Cristiane Machado et al. (2011). “Circadian rhythm and cartilage extracellular matrix genes in osseointegration: a genome-wide screening of implant failure by vitamin D deficiency”. In: *PloS one* 6.1, e15848.

Metallo, Christian M et al. (2012). “Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia”. In: *Nature* 481.7381, p. 380.

Miller, Brooke H et al. (2007). “Circadian and CLOCK-controlled regulation of the mouse transcriptome and cell proliferation”. In: *Proceedings of the National Academy of Sciences* 104.9, pp. 3342–3347.

Miyazaki, Toshihiko et al. (2001). “Phase-advance shifts of human circadian pacemaker are accelerated by daytime physical exercise”. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 281.1, R197–R205.

mousephenotype.org (2018). *International Mouse Phenotyping Consortium (IMPC)*. URL: http://www.mousephenotype.org/objectives-and-background (visited on 09/23/2018).

Mullenders, Jasper et al. (2009). “A large scale shRNA barcode screen identifies the circadian clock component ARNTL as putative regulator of the p53 tumor suppressor pathway”. In: *PloS one* 4.3, e4798.

Murgia, Marta et al. (2017). “Single muscle fiber proteomics reveals fiber-type-specific features of human muscle aging”. In: *Cell reports* 19.11, pp. 2396–2409.

Naba, Alexandra et al. (2016). “The extracellular matrix: Tools and insights for the “omics” era”. In: *Matrix Biology* 49, pp. 10–24.

Nievergelt, Caroline M et al. (2006). “Suggestive evidence for association of the circadian genes PERIOD3 and ARNTL with bipolar disorder”. In: *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 141.3, pp. 234–241.

omim.org (2018a). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/about (visited on 09/23/2018).

– (2018b). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/602550?search=arntl&highlight=arntl (visited on 09/23/2018).

– (2018c). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/604256?search=bhlhe40&highlight=bhlhe40 (visited on 09/23/2018).

– (2018d). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/601851?search=clock&highlight=clock (visited on 09/23/2018).

– (2018e). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/601933?search=cry1&highlight=cry1 (visited on 09/23/2018).

omim.org (2018f). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/600864?search=csnk1d&highlight=csnk1d (visited on 09/23/2018).
– (2018g). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/602408?search=nr1d1&highlight=nr1d1 (visited on 09/23/2018).
– (2018h). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/605327?search=nfil3&highlight=nfil3 (visited on 09/23/2018).
– (2018i). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/603427?search=per3&highlight=per3 (visited on 09/23/2018).
– (2018j). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/608886?search=ppargc1b&highlight=ppargc1b (visited on 09/23/2018).
– (2018k). *Online Mendelian Inheritance in Man*. URL: https://www.omim.org/entry/116899?search=CDKN1A&highlight=cdkn1a (visited on 09/23/2018).
Ozturk, Nuri et al. (2009). "Loss of cryptochrome reduces cancer risk in p53 mutant mice". In: *Proceedings of the National Academy of Sciences* 106.8, pp. 2841–2846.
Patke, Alina et al. (2017). "Mutation of the human circadian clock gene CRY1 in familial delayed sleep phase disorder". In: *Cell* 169.2, pp. 203–215.
Pittendrigh, Colin S (1960). "Circadian rhythms and the circadian organization of living systems". In: *Cold Spring Harbor symposia on quantitative biology*. Vol. 25. Cold Spring Harbor Laboratory Press, pp. 159–184.
Pizarro, Angel et al. (2012). "CircaDB: a database of mammalian circadian gene expression profiles". In: *Nucleic acids research* 41.D1, pp. D1009–D1013.
Potts, Gregory K et al. (2017). "A map of the phosphoproteomic alterations that occur after a bout of maximal-intensity contractions". In: *The Journal of physiology*.
Preitner, Nicolas et al. (2002). "The orphan nuclear receptor REV-ERBα controls circadian transcription within the positive limb of the mammalian circadian oscillator". In: *Cell* 110.2, pp. 251–260.
proteinatlas.org (2018a). URL: https://www.proteinatlas.org/ENSG00000133794-ARNTL/cell#human (visited on 09/23/2018).
– (2018b). URL: https://www.proteinatlas.org/ENSG00000121671-CRY2/cell#human (visited on 09/23/2018).
– (2018c). *The Human Protein Atlas*. URL: https://www.proteinatlas.org/ (visited on 09/23/2018).
Reppert, Steven M and David R Weaver (2001). "Molecular analysis of mammalian circadian rhythms". In: *Annual review of physiology* 63.1, pp. 647–676.
Richardson, Cele E et al. (2017). "Can exercise regulate the circadian system of adolescents? Novel implications for the treatment of delayed sleep-wake phase disorder". In: *Sleep medicine reviews* 34, pp. 122–129.
Richter, Curt Paul (1922). "A behavioristic study of the activity of the rat." In: *Comparative Psychology Monographs*.
Roenneberg, Till, C Jairaj Kumar, and Martha Merrow (2007). "The human circadian clock entrains to sun time". In: *Current Biology* 17.2, R44–R45.
Roenneberg, Till, Anna Wirz-Justice, and Martha Merrow (2003). "Life between clocks: daily temporal patterns of human chronotypes". In: *Journal of biological rhythms* 18.1, pp. 80–90.
Rosenberg, Laura H et al. (2015). "Therapeutic targeting of casein kinase 1δ in breast cancer". In: *Science translational medicine* 7.318, 318ra202–318ra202.
Rudic, R Daniel et al. (2004). "BMAL1 and CLOCK, two essential components of the circadian clock, are involved in glucose homeostasis". In: *PLoS biology* 2.11, e377.
Rudnicki, Michael A et al. (1992). "Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development". In: *Cell* 71.3, pp. 383–390.

Sahar, Saurabh and Paolo Sassone-Corsi (2009). “Metabolism and cancer: the circadian clock connection”. In: *Nature Reviews Cancer* 9.12, p. 886.

Sato, Trey K et al. (2004). “A functional genomics strategy reveals Rora as a component of the mammalian circadian clock”. In: *Neuron* 43.4, pp. 527–537.

Schiaffino, Stefano, Bert Blaauw, and Kenneth A Dyar (2016). “The functional significance of the skeletal muscle clock: lessons from Bmal1 knockout models”. In: *Skeletal muscle* 6.1, p. 33.

Schittek, Birgit and Tobias Sinnberg (2014). “Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis”. In: *Molecular cancer* 13.1, p. 231.

Schroder, Elizabeth A and Karyn A Esser (2013). “Circadian rhythms, skeletal muscle molecular clocks and exercise”. In: *Exercise and sport sciences reviews* 41.4.

Sen, Satish et al. (2018). “Expression of the clock gene Rev-erbα in the brain controls the circadian organisation of food intake and locomotor activity, but not daily variations of energy metabolism”. In: *Journal of neuroendocrinology* 30.1, e12557.

Seuter, Sabine et al. (2013). “The gene for the transcription factor BHLHE40/DEC1/stra13 is a dynamically regulated primary target of the vitamin D receptor”. In: *The Journal of steroid biochemistry and molecular biology* 136, pp. 62–67.

Shearman, Lauren P et al. (2000). “Interacting molecular loops in the mammalian circadian clock”. In: *Science* 288.5468, pp. 1013–1019.

Shen, Qiwei et al. (2017). “Organ-specific alterations in circadian genes by vertical sleeve gastrectomy in an obese diabetic mouse model”. In: *Science bulletin* 62.7, p. 467.

Shukla, Neerav et al. (2011). “Oncogene mutation profiling of pediatric solid tumors reveals significant subsets of embryonal rhabdomyosarcoma and neuroblastoma with mutated genes in growth signaling pathways”. In: *Clinical Cancer Research*.

Solanes, Gemma et al. (2003). “Functional relationship between MyoD and PPAR-Dependent regulatory pathways in the control of the human uncoupling protein-3 gene transcription”. In: *Mol. Endocrinol* 17, pp. 1944–1958.

Soria, Virginia et al. (2010). “Differential association of circadian genes with mood disorders: CRY1 and NPAS2 are associated with unipolar major depression and CLOCK and VIP with bipolar disorder”. In: *Neuropsychopharmacology* 35.6, p. 1279.

Spitzer, RL et al. (1997). “Columbia jet lag scale”. In: *Biometrics research. New York: NY State Psychiatric Institute*.

Starkie, Rebecca L et al. (1999). “Effect of temperature on muscle metabolism during submaximal exercise in humans”. In: *Experimental physiology* 84.4, pp. 775–784.

Stephan, FK (1984). “Phase shifts of circadian rhythms in activity entrained to food access”. In: *Physiology & behavior* 32.4, pp. 663–671.

Stokkan, Karl-Arne et al. (2001). “Entrainment of the circadian clock in the liver by feeding”. In: *Science* 291.5503, pp. 490–493.

Stubblefield, Jeremy J, Jérémy Terrien, and Carla B Green (2012). “Nocturnin: at the crossroads of clocks and metabolism”. In: *Trends in Endocrinology & Metabolism* 23.7, pp. 326–333.

Swanson, Garth R, Helen J Burgess, and Ali Keshavarzian (2011). “Sleep disturbances and inflammatory bowel disease: a potential trigger for disease flare?” In: *Expert review of clinical immunology* 7.1, pp. 29–36.

Tahara, Yu, Shinya Aoyama, and Shigenobu Shibata (2017). “The mammalian circadian clock and its entrainment by stress and exercise”. In: *The Journal of Physiological Sciences* 67.1, pp. 1–10.

Tahara, Yu et al. (2015). “Entrainment of the mouse circadian clock by sub-acute physical and psychological stress”. In: *Scientific reports* 5, p. 11417.

Taillard, Jacques et al. (2003). “The circadian and homeostatic modulation of sleep pressure during wakefulness differs between morning and evening chronotypes”. In: *Journal of sleep research* 12.4, pp. 275–282.

Takahashi, Joseph S et al. (2008). “The genetics of mammalian circadian order and disorder: implications for physiology and disease”. In: *Nature Reviews Genetics* 9.10, p. 764.

Tang, Zhaohua, Thomas R Coleman, and William G Dunphy (1993). “Two distinct mechanisms for negative regulation of the Wee1 protein kinase.” In: *The EMBO Journal* 12.9, pp. 3427–3436.

Teo, Weipeng, Michael J Newton, and Michael R McGuigan (2011). “Circadian rhythms in exercise performance: implications for hormonal and muscular adaptation”. In: *Journal of sports science & medicine* 10.4, p. 600.

Tirinato, Luca et al. (2017). “An overview of lipid droplets in cancer and cancer stem cells”. In: *Stem cells international* 2017.

toppgene.cchmc.org (2018). *ToppGene Suite*. URL: `https://toppgene.cchmc.org/help/help.jsp` (visited on 09/23/2018).

Tremblay, Annie M et al. (2014). “The Hippo transducer YAP1 transforms activated satellite cells and is a potent effector of embryonal rhabdomyosarcoma formation”. In: *Cancer cell* 26.2, pp. 273–287.

tumorportal.org (2018). *TumorPortal*. URL: `http://www.tumorportal.org/` (visited on 09/23/2018).

Van Der Horst, Gijsbertus TJ et al. (1999). “Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms”. In: *Nature* 398.6728, p. 627.

Vidal-Puig, Antonio et al. (1997). “UCP3: an uncoupling protein homologue expressed preferentially and abundantly in skeletal muscle and brown adipose tissue”. In: *Biochemical and biophysical research communications* 235.1, pp. 79–82.

Vissing, Kristian and Peter Schjerling (2014). “Simplified data access on human skeletal muscle transcriptome responses to differentiated exercise”. In: *Scientific data* 1, p. 140041.

Vitaterna, Martha Hotz et al. (1999). “Differential regulation of mammalian period genes and circadian rhythmicity by cryptochromes 1 and 2”. In: *Proceedings of the National Academy of Sciences* 96.21, pp. 12114–12119.

Wang, Qing et al. (2010). “Cardiac hypertrophy, low blood pressure, and low aldosterone levels in mice devoid of the three circadian PAR bZip transcription factors DBP, HLF, and TEF”. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 299.4, R1013–R1019.

Ward, Patrick S et al. (2010). “The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting α-ketoglutarate to 2-hydroxyglutarate”. In: *Cancer cell* 17.3, pp. 225–234.

Waterhouse, Jim, Thomas Reilly, and Greg Atkinson (1997). “Jet-lag”. In: *The Lancet* 350.9091, pp. 1611–1616.

Weintraub, Harold (1993). “The MyoD family and myogenesis: redundancy, networks, and thresholds”. In: *Cell* 75.7, pp. 1241–1244.

Williams, Biju et al. (2017). “Managing Performance Throughout Periods of Travel”. In: *Strength & Conditioning Journal* 39.4, pp. 22–29.

Wolff, Gretchen and Karyn A Esser (2012). “Scheduled exercise phase shifts the circadian clock in skeletal muscle”. In: *Medicine and science in sports and exercise* 44.9, p. 1663.

Wolins, Nathan E et al. (2003). “Adipocyte protein S3-12 coats nascent lipid droplets”. In: *Journal of Biological Chemistry* 278.39, pp. 37713–37721.

Yamazaki, Shin et al. (2000). “Resetting central and peripheral circadian oscillators in transgenic rats”. In: *Science* 288.5466, pp. 682–685.

Yang, Zhihong et al. (2009). “MyoD and E-protein heterodimers switch rhabdomyosarcoma cells from an arrested myoblast phase to a differentiated state”. In: *Genes & development* 23.6, pp. 694–707.

yeastgenome.org (2018). URL: https://www.yeastgenome.org/go/GO:0035770 (visited on 09/23/2018).

Zaidi, Nousheen, Johannes V Swinnen, and Karine Smans (2012). “ATP-citrate lyase: a key player in cancer metabolism”. In: *Cancer research*.

Zambon, Alexander C et al. (2003). “Time-and exercise-dependent gene regulation in human skeletal muscle”. In: *Genome biology* 4.10, R61.

Zanello, Susana B, David M Jackson, and Michael F Holick (2000). “Expression of the circadian clock genes clock and period1 in human skin”. In: *Journal of investigative dermatology* 115.4, pp. 757–760.

Zhou, Bing et al. (2001). “A novel pantothenate kinase gene (PANK2) is defective in Hallervorden-Spatz syndrome”. In: *Nature genetics* 28.4, p. 345.

Zhu, Y et al. (2006). “RORA, a large common fragile site gene, is involved in cellular stress response”. In: *Oncogene* 25.20, p. 2901.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit mit dem Titel

Circadian Rhythm Genes & Exercise: A Bioinformatic Analysis

an der Technischen Universität München im Sommersemester 2018 selbstständig verfasst und keine anderen als die von mir im Literaturverzeichnis angegebenen Werke/Hilfsmittel benutzt habe. Die Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinne nach entnommen sind, wurden in allen Fällen unter Angabe der Quellen (einschließlich des World Wide Web und anderer elektronischer Text- und Datensammlungen) kenntlich gemacht. Dies gilt auch für beigegebene Zeichnungen, bildliche Darstellungen und Skizzen.

München, 25. September 2018

.....

Christian Soetebier