Aus dem
CharitéCentrum für Therapieforschung (CC 4)
Institut für Theoretische Biologie - ITB
Direktor: Prof. Dr. Hanspeter Herzel

Habilitationsschrift

A Systems Biology View on the Role of Alternative Splicing and the Circadian Clock in Tumour Progression

zur Erlangung der Lehrbefähigung
für das Fach Molekularbiologie und Bioinformatik

vorgelegt dem Fakultätsrat der Medizinischen Fakultät
Charité-Universitätsmedizin Berlin

von

Dr. Angela Moreira Borralho Relógio

geboren in Beja, Portugal

Eingereicht:	**April/2015**
Dekan:	**Prof. Dr. med. Axel R. Pries**
1. Gutachter/in:	**Prof. Francis Levi, Warwick/United Kingdom**
2. Gutachter/in:	**Prof. Steven Brown, Zürich/ Switzerland**

Abbreviations

AS	alternative splicing
BMAL 1,2	brain and muscle ARNT like 1,2
CCGs	clock-controlled genes
CELF/CUGBP proteins	CUG triplet repeat, RNA binding proteins
CKIε	casein kinase epsilon
CLOCK	circadian locomotor output cycles kaput
CRY 1,2	cryptochrome 1,2
EPO	erythropoietin
FASPS	familial sleep disorders
HIF-1alpha	hypoxia inducible factor 1, alpha
hnRNP	heterogeneous nuclear ribonucleoproteins
JNK2	c-Jun kinase 2
KH proteins	K Homology (KH) domain proteins
MBNL	muscle bind like proteins
MDM2	mouse double minute 2 homolog
MYC	v-myc avian myelocytomatosis viral oncogene homolog
NPAS	neuronal PAS domain-containing protein
PER 1,2,3	period 1,2,3
RBM	RNA binding motif proteins
ROR	retinoic acid receptor related orphan receptor
SAD	seasonal affective disorders
SCN	suprachiasmatic nucleus
SIRT1	sirtuin 1
snRNPs	small nuclear ribonucleoproteins
SR proteins	serine /arginine rich proteins
VEGF	vascular endothelial growth factor
WEE1	WEE1 G2 checkpoint kinase

Index

1. Introduction

1.1. Topic of research

Cancer is a long known disease affecting, alone in Germany, many thousands of persons per year. The associated physical and emotional stress is very high, for patients and their families. Any progress in its diagnostic and treatment, including the reduction of secondary toxic effects, would represent a significant improvement in the life quality of patients.

There are several difficulties concerning the treatment of malignant tumours among which the complexity and genetic heterogeneity of the disease, and the secondary toxic effects caused by radiotherapy and/or chemotherapy procedures. Genetic stage-specific diagnostics and a comprehensive knowledge of the circadian pattern of relevant genes may allow for the day-time optimization of treatment (chronotherapy) [1-3]. This approach diminishes the toxicity to healthy tissues and shows positive effects on the survival of patients, pointing to a role of the circadian system in cancer. Additionally, data published by me and others, points to a stage-specific genetic signature of tumour progression [4]. Such specificity is partially due to the diverse gene products generated via alternative splicing from a single gene. Alternative splicing is regulated by splicing factors and recent correlations between these and the circadian system have been described [5-8].

This work describes a comprehensive systems biology approach to couple circadian molecular processes and alternative splicing with tumourigenesis, using different cancer cell lines as a model system. To mechanistically understand the connection of the circadian system to tumour progression is of major relevance. We use a systems biology approach involving wet-lab experiments, including genome wide screening of gene expression of human cell lines and tissues bioinformatics, and computational models, to try to understand the dynamic interplay between cancer and the clock. The comprehensive analysis of the circadian clock will benefit from mathematical modelling, and oncogenic pathways can be investigated and modelled in a way that discloses tumour evasion strategies.

With such a methodology we address the following questions:

-How are the pathways, which connect the circadian clock to cancer, regulated?

-Is this regulation specific for different stages of tumour progression?

-Can a circadian signature for tumour progression be defined and is there a potential role for splicing within such a regulation?

1.2. Background knowledge

Living organisms have developed an internal time-generating system which robustly produces rhythms in the expression of genes and proteins. Such rhythms are endogenous and in mammals are present in 10% of all genes [1]. This time-system is known as the circadian clock and allows for the adaptation of cellular processes to specific times of the day preparing organisms to react to external cues.

Failures of the circadian clock may occur and have been object of study related to different pathological conditions such as: predisposition to cancer [9,10], familial sleep disorders (FASPS) [11], bipolar disorder, seasonal affective disorders (SAD), sleep problems in the elderly, diabetes [12] and obesity [13]. Hence, it is of no surprise that the correct regulation of cellular processes may contribute to adjustment of therapy including the readjustment of drug intake-times. Recent reports suggest that chronomodulation of drug intake-times in cancer patients may diminish toxicity, as well as increase treatment effectiveness [1].

Several core-clock genes are directly associated to cancer stimulating mechanisms. PER1 and Timeless bind to DNA damage response proteins and *Per1* overexpression can suppress the growth of human cancer cell lines. *Per2* mRNA levels are down-regulated in several human lymphoma cell lines and in non-small-cell lung cancer tissues [3]. *Per* genes are deregulated in human breast cancer cells [4]. Casein kinase epsilon (CKIε), which phosphorylates PER proteins, shows increased expression levels in various human cancers such as leukaemia and prostate cancer and its knock-down induces growth inhibition in cancer cells [5]. Mutations in *Npas,* a paralog of *Clock*, are linked to augmented incidence of breast cancer and non-Hodgkin's lymphoma [6, 7]. Furthermore, recent findings described clock elements as potential cell cycle regulators underlining a potential path via which the circadian clock might influence senescence as well. Sirtuin 1 (SIRT1), was recently identified as a core-clock component [8]. This histone deacetylase can also both promote and suppress tumour growth, depending on its interacting partners [9]. Additionally, *Wee1*, *Myc*, and *cyclin D1,* all regulators of the cell cycle, are direct output genes of the core-clock [10, 11]. Clock-controlled genes (CCGs) are regulated at the transcriptional level by CLOCK-BMAL and SIRT1, and are involved in the cell cycle, in proliferation and metabolism. Examples of such CCGs are MDM2 and VEGF, involved in proliferation and the glycolytic enzyme pyruvate kinase involved in metabolism. Altogether, these results suggest the existence of a direct link between transformation and perturbations of core-clock genes. Moreover, several of the genes mentioned are directly involved in the cell cycle and proliferation and as a consequence might play an important role in tumorigenesis. This interplay between transformation, cell cycle and the circadian clock is an extremely interesting topic not yet explored and one of the research lines being developed in my group.

Cancer is also directly linked to deregulations of the circadian system, even though the mechanisms involved are not understood [10]. Current studies point to a connection between the circadian clock and splicing [5,7,14]. It has been shown that alternative patterns of splicing are associated in a specific manner to different tumour types and can be responsible for the malignant transformation of tumours [4,15,16]. Interestingly, some of the alternatively spliced genes were independently described as clock-controlled genes (CCGs) and are involved in many molecular processes essential for malignant transformation of tumour cells [9,10]. The coupling of the circadian system to the regulation of splicing and its implications in tumour progression is still an unexploited field of research. A systems biology approach seems appropriate to understand this topic. The relationship between splicing and cancer and the connection to the clock can be investigated using an

interdisciplinary and iterative approach, involving a combined theoretical/experimental methodology, including genome wide screening of gene expression, bioinformatics, and computational models.

2. Results

The systems biology work described in the next pages allows for the study of cancer cell lines, as a model system, nested in a network of gene regulation at the circadian and splicing levels. With our research, we expect to be able to provide valuable insights into the mechanism of circadian regulation of tumourigenesis per se and to contribute to a better understanding of the cancer-clock system.

Altogether, this data contributes to better understand the interactions of splicing and cancer on the one hand and the circadian clock and cancer on the other, by connecting mathematical modelling of the mammalian circadian clock to microarray studies of human cell lines and tissues. A systems biology approach is being applied to investigate the following questions: How are the circadian pathways regulated? Is this regulation specific for different tumour types? Is it altered in proliferating cells? Is there a circadian transformed fingerprint for tumour progression? Such an approach is expected to provide novel insights into the mechanism of circadian regulation of splicing and tumourigenesis *per se* contributing to its better understanding. Furthermore, we aim with our work to be able to convey valuable information in the improvement of diagnosis and prognosis in cancer, with the prospective of highlighting novel drug-targets and optimizing therapy.

2.1. Splicing

Microarray technology and deep-sequencing data has provided important insights in cancer research. Mostly, the diagnostic of tumours depended on indirect indicators based on which the disease was categorized into histological and morphologic subtypes. Currently, the ability to scan, at the same time, the entire genome of a patient and with the same platform compare several patient samples, thereby establishing molecular signatures of disease state, offers major advantages. A detailed genetic diagnostic of the disease and also a possible prognostic can be attained. Also for studies of splicing and overall gene expression changes can such high-throughput systems be a major advantage.

Eukaryotic cells are capable of synthesizing different messenger RNAs and proteins from the same set of genes, depending on the developmental stage, tissue function of the cell, or in response to intracellular, or extracellular signals. This ability to generate diversity is due to regulation of gene expression at several levels: transcription, processing of the messenger RNA precursor (pre-mRNA), transport of the mRNA from the nucleus to the cytoplasm, degradation of the mRNA, translational regulation and posttranslational modifications that affect the activity of the protein.

A major step in the regulation of gene expression is the mechanism by which the non-translated (intronic) sequences of the pre-mRNA are excised and the exonic (translated) sequences are fused together to produce a continuous protein-coding mRNA [17]. This process is named splicing. Splicing is catalysed by the spliceosome which is a large RNA-

protein complex composed of five small nuclear ribonucleoproteins (snRNPs). Additionally, a large number of proteins, named splicing factors, participate in the recognition of splice sites and are able to activate or repress the use of a splice site [18].
Splicing factors belong to various protein families. These comprise the SR (serine /arginine - rich proteins), hnRNP (heterogeneous nuclear ribonucleoproteins), RBM (RNA-binding motif proteins), MBNL (Muscle bind-like proteins), CELF/CUGBP (CUG triplet repeat, RNA binding proteins) and KH (K Homology (KH) domain proteins) families. The ratios between these proteins play a role in tissue-specific splicing. [19]. To be able to investigate the relative ratios of splicing factors among the tumour cell lines, we developed specific microarrays which were optimized for studying both fold changes in splicing factors, as well as the corresponding resulting splicing events [20].

Relógio A., Schwager C., Richeter, A., Ansorge W., Valcárcel J.

Optimization of oligonucleotide-based DNA microarrays, Nucleic Acids Research, Jun/2002, e51, 30(11).

Number of citations: 186
IF: 8.81

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

doi: 10.1093/nar/30.11.e51

2.2 Alternative splicing and cancer progression

In higher eukaryotes, the number of introns in pre-mRNAs can vary from 0 to more than 50 and can extend to more than 200,000 nucleotides. The presence of more than one intron in a pre-mRNA allows for a combinatorial fusion of exons, resulting in different protein isoforms, or in an on-off switch of gene expression. This process is known as alternative splicing (AS). It is estimated that at least 70% of human genes undergo alternative splicing [21].

AS decisions are controlled by the protein families of splicing factors previously mentioned (hnRNPs and SR proteins). Their cooperative and antagonistic effects combined at regulatory elements are important for specificity in cell- and tissue-specific alternative splicing patterns [22]. Variations in alternative splicing can originate or alter the progression of a substantial number of pathologies [23]. Evidence accumulates that AS is important in tumour progression. We used our developed microarrays to screen cancer lines for their relative differences in terms of splicing factors and alternative splicing events.

Microarray technology has provided important insights in cancer research. Global expression analysis using microarray technology permits the screening of entire genomes in a high-throughput manner and allows for the generation of molecular signatures of a particular tumour sample. We can assume that the particular group of genetic defects of a tumour determines its clinical behaviour. These genetic lesions modify the expression pattern of the mRNA in cancer, this is viewed as a "molecular signature" or "fingerprint" of the tumour. As a consequence, the tumours which share genetic lesions have comparable "signatures" and as such also related clinical phenotypes. Therefore, the following hypothesis are valid: a) gene expression profiling will contribute to create a clinically and biologically significant tumour grouping; b) gene expression data may be used for in prognostic and even for the establishment of therapy procedures; and c) a careful analysis of high-throughput expression data may permit the identification of marker-genes for a certain disease phenotype. In cancer research, it can provide high-throughput information and valuable understanding into differences in an individual's tumour. Across individuals, such comparisons can provide tissue-specific disease signatures that generate diagnosis based on hundreds of informative genes.

With our methodology, we were able to identify AS changes in Hodgkin lymphoma which may be used as potent markers of tumour progression using high-throughput screenings of mammalian tissues and cell lines [4,24]. In summary, our analysis suggests the existence of a variety of mechanisms by which the splicing regulatory program of Hodgkin's lymphoma cells can be altered. These include changes in the expression of snRNP components or sub-complexes within them, changes in the expression of hnRNP and SR proteins or of enzymes that modulate them, and expression of splicing regulators normally expressed in other cell types.

In particular, we found that differential expression of the splicing regulators Nova1 and Nova2 can be detected in different Hodgkin lymphoma cell lines and that they correlate with

the alternative splicing observed for JUNK2 kinase and most interesting, that this correlation seems to be dependent of the tumour grade.

Relógio A., Ben-Dov C., Baum M., Ruggiu M., Gemund C., Benes V., Darnell RB., Valcarcel J. Alternative splicing microarrays reveal functional expression of neuron-specific regulators in Hodgkin lymphoma cells, Journal of Biological Chemistry, Feb/2005, 4779-84, 280(6).
Number of citations: 54
IF: 4.60

doi: 10.1074/jbc.M411976200

doi: 10.1074/jbc.M411976200

doi: 10.1074/jbc.M411976200

doi: 10.1074/jbc.M411976200

doi: 10.1074/jbc.M411976200

2.3 The circadian clock in tumour progression

2.3.1 A network of circadian regulation

Circa 24 hours rhythms in the expression of genes and proteins have been described in most organisms, from cyanobacteria to mammals and appear to represent an evolutionary advantage. This time constrain is generated by molecular clocks [25,26] and assumed to allow for an ideal adaptation to dark/light cycles. The molecular clock is formed by a network of interconnected genes and their protein products structured into two main feedback loops: a positive ROR (retinoic acid receptor-related orphan receptor)/*Bmal (Brain and Muscle ARNT-like)*/REV (Rev-Erb nuclear orphan receptor) loop; and a negative PER (Period)/CRY (Cryptochrome) loop. These are able to generate *circa* 24h fluctuations in the expression of genes, even in the absence of external clues, and can be found in 10% of all genes. In mammals, the main pacemaker is found in the suprachiasmatic nucleus (SCN). The light signals, via the retinohypothalamic tract, can directly entrain the SCN clock [27]. This central clock is formed by a group of ~200.000 neurons in humans and generates oscillations which are synchronized both in the brain, as well as in the remaining peripheral organs. A circadian tissue specificity is accountable for the fine-tuning of local transcriptional activity for *eg*. in the liver, heart, kidney and skin. The circadian system regulates the time of numerous processes at the cellular and system level, such as sleep-awake cycles, memory consolidation [28,29], metabolism of glucose, lipids and drugs, bone formation [30], hormone regulation, immunity, the cell cycle time, heart rate, blood pressure and body temperature [31,32]. In this work, we could provide an extension of circadian regulation by a high-throughput analysis of tissue-specific patterns of oscillations for several CCGs. Among the CCGs, we investigated the promoter regions of 2065 genes and searched for highly overrepresented transcription factors binding sites. We found many known regulators which points to the robustness of our methodology, but we also found several new promising candidates for circadian regulation. In particular, we showed that one of our predicted regulators HIF-1alpha binds to the EPO promoter region and activates its circadian expression.

Bozek K.*, Relógio A.*, Kielbasa S.M., Heine M., Dame C., Kramer A., Herzel H.
Regulation of clock-controlled genes in mammals, PLoS One, Mar/2009, e4882, 4(3).
***equally contributed**
Number of citations: 82
IF: 3.53

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

doi: 10.1371/journal.pone.0004882

2.3.2. A mathematical model for the mammalian circadian clock

In order to better understand the molecular interactions within the circadian clock and their implications in disease, we constructed a single cell mathematical model for the mouse circadian core-clock. Mathematical models permit the study of all variables and parameters of a regulatory network separately which is important in the identification of threshold points of activity or analysis of perturbation properties on the network system. The model developed contains the following clock genes/gene families and corresponding proteins: *Per*, *Cry*, *Rev-Erb*, *Clock* and *Bmal* interconnected via positive and negative feed-back loops. Relevant biological reactions/regulations such as, transcription, translation, degradation, phosphorylation and import/export are also represented. It permits an independent analysis of the feed-back loops, it is biologically broad, highlights a parameterization based on biochemical measurements, and reproduces the present knowledge of research on the field.

Although much is described about the circadian system, the kinetics of the single reactions is unknown which makes the parameterization procedure difficult. We developed a method to exploit phases and amplitudes within the variables of the model and used the linear-time-invariant (LTI) system theory [33] to attain approximations for most of the unknown parameters. The resulting model was tested, using genome-wide RNAi experiments [34,35], transcriptional inhibition data [36] and new experimental data. One of our predictions related to the role of the positive feedback loop showed a loss of oscillations of *Bmal* reporter upon overexpression of *Rev-Erb* and *Ror*.

This model has been used to generate hypothesis which were successfully tested experimentally. One open question which the model could answer is related to the kinetics of degradation and how it affects the period of the system. A faster degradation of a clock protein, such as CRY1, might lead to a shorter or longer period [37]. Since there is a correlation of degradation rates and delay [38], it could be hypothesised that a faster degradation would lead to a shorter delay and as such to a shorter period. This effect was measured in a model for FASPS pathology [11]. However, a faster degradation can as well decrease the nuclear accumulation of PER/CRY and lead to a longer period. Both assumptions could be valid and consequently, a detailed quantitative analysis is necessary to determine the correct one. In such a system, the intrinsic period will vary according to the time of gene expression, but also with the degradation and nuclear/cytoplasmic accumulation. Therefore, predictions are difficult and modelling can indeed bring valuable insights into such topics. Our *in silico* data show that short and long periods can be attained. These model predictions are consistent with recent experimental data. Our model brings novel explanations for circadian results and helps to understand seemingly incompatible phenotypes.

This model is being extended in order to further analyse the circadian clock and its molecular implications in both cancer and splicing.

Relógio A*., Westermark P., Wallach T., Schellenberg K., Kramer A., Herzel H.
Tuning the Mammalian Circadian Clock: Robust Synergy of Two loops, PLoS Computational Biology, Dec/2011, e1002309, 7(12).
***corresponding author**
Number of citations: 24
IF: 4.83

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

doi: 10.1371/journal.pcbi.1002309

2.3.3. Ras oncogene can lead to the deregulation of the circadian clock in cell line models of colon and skin cancer.

The mammalian clock network is constituted by two interconnected feedback loops which are capable of generating ~ 24 hour rhythms [39]. Due to its complexity and biological relevance, the circadian clock has been studied both experimentally and theoretically.

We developed a circadian core-clock network using text mining tools [40,41] and manual curation. A set of 'signature' genes involved in tumour progression was assembled. We searched for oscillatory patterns in published microarrays studies regarding mammalian cells and tissues [42] to identify clock-controlled genes and extract the phases and amplitudes from all genes present in the network previously designed. This information was integrated in the network.

In parallel, we extended a recent mathematical model of the circadian clock, developed by me and colleagues to CCGs related to cell cycle, metabolism and proliferation. Early mathematical models of the clock omit indispensable elements such as the nuclear receptor ROR [43-48] or posttranslational modifications [49]. Other models are fairly large and as a result the determination of kinetic parameters is extremely complex [45,46].

The model develop contains 20 species, including RNAs and proteins and allows the study of the two feedback loops: PER/CRY loop (PC) and ROR/*Bmal*/REV (RBR); as well as other elementary mechanisms essential for the regulation of the circadian system. Such mechanisms include transcription, translation, import/export, degradation and phosphorylation and by what means these processes disturb the period and the strength of the system. It is concentrated on the core-clock in the SCN which is presumed to be accountable for the management of the entire circadian network-structure and it might be responsible for overall failures and consequent disruption of peripheral clocks, leading to the disturbance of normal oscillation-times [1,10,11].

We used our model to generate hypothesis regarding the role of RAS as a deregulator of the circadian system which we tested experimentally, using single cell bioluminescence assays.

We could show that RAS influences the mammalian circadian clock by interfering with BMAL1-mediated transcription. As such an overexpression of RAS, which we found to occur in our models for tumours in advanced progression stages, lead to a decrease of BMAL1 - mediated transcription and to an increase in the period of the system. Using an inhibitor of RAS we could recover the original clock phenotype with the corresponding shorter period, of about 24h, as observed in normal cells.

Relógio A*., Medina-Pérez P., Thomas P., Gloc E., Bervoets S., Maier B., Schaefer R., Leser U., Herzel H., Kramer A., Sers C.

Ras – mediated deregulation of the circadian clock in cancer, PLoS Genetics, May/2014, e1004338, 10(5). *corresponding author

Number of citations: 5

IF: 8.52

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

doi: 10.1371/journal.pgen.1004338

3. Discussion

3.1. The circadian clock, splicing and tumour progression

Almost all organisms evolved an endogenous circadian clock which regulates the timing of central biological processes and provides a way to adapt physiology and behaviour to daily dark/light rhythms [50-52]. In mammals, malfunctions of the circadian system are associated to known pathologies ranging from sleep to metabolic disorders, to cancer [53-55]. Hence, a detailed overview of the underlying genetic network that shapes the mammalian circadian system is of major interest to the circadian and medical field.

With the work here presented an attempt was made to apply stage-of-art technology to in-depth bioinformatics methodologies, which complemented with molecular biology and biochemical methods aimed to couple splicing and cancer and to show a correlation between cancer and the circadian clock. This work paves the way for what is currently the main topic of research in my group: a full circadian analysis of the splicing machinery and its putative consequences on alternative splice choices in tumourigenesis.

3.2. How can a living system tell time?

The definition of the circadian system is by itself a controversial topic. Much has been gathered in the last decades regarding the regulation and generation of time. It is largely accepted that the mammalian circadian system is hierarchically organized. A main pacemaker formed by two clusters of ~100,000 neurons (in humans) is located in the suprachiasmatic nucleus (SCN), but peripheral oscillators exist in virtually every of our 3.5×10^{13} body cells [56,57]. Within a cell, a reduced set of 14 genes has been reported to form the so called core-clock network (CCN). These genes encode for members of the PER (period), CRY (cryptochrome), BMAL (brain and muscle ARNT-like protein), CLOCK (circadian locomotor output cycles kaput), NPAS2 (neuronal PAS domain-containing protein 2) in neuronal tissue), ROR (retinoic acid receptor-related orphan receptor) and REV-ERB (nuclear receptor, reverse strand of ERBA) families. The CCN is arranged into two main interconnected feed-back loops: a) the RORs/*Bmal*/REV-ERBs (RBR) loop and b) the PERs/CRYs (PC) loop [58]. Both loops are able to produce rhythms in gene expression, independently, but need to be interconnected to robustly generate oscillations with a period of circa 24 hours [59,60]. In the centre of the core-clock network lays the heterodimer complex CLOCK/BMAL1. This complex regulates the transcription of elements of both the RBR and PC loop by binding to E-Box sequences in the promoter region of the target genes. In the RBR-loop, *Rev-Erbα,β* and *Rorα,β,γ* are transcribed. After translation, the resulting proteins compete for RORE elements within the *Bmal1* promoter region and hold antagonistic effects, thereby fine-tuning *Bmal1* expression. In the PC loop, following transcription and translation, PER1,2,3 and CRY1,2 form complexes and inhibit CLOCK/BMAL mediated-transcription, thus regulating the expression of all target genes mentioned above.

With our work we have greatly contributed to this concept of a molecular clock and to a deeper understanding of how the core-clock genes and proteins interact within the positive and negative feed-back loops and are able to generate oscillations. The simple idea of a core-clock network leads to further questions related to the mechanism of how regulation via this CCN can be achieved in order to propagate time cues throughout the organism.
Therefore, the CCN has been studied, on a fine scale, at the transcriptional, translational and post-translational level both experimentally and with mathematical models [58,61-67]. Various efforts have been made to describe the exact mechanisms through which the mammalian CCN may regulate its target genes, the clock-controlled genes (CCG), as well as to identify new CCGs and to define new network topologies [42,68,69].
We have addressed the network generation, curation and validation topic for the clock system and could also identify several network modules which are particularly relevant for cancer and even for splicing studies. Yet, a more detailed knowledge on the full range of genes and subsequent biological processes that are regulated by the core of the circadian clock is still missing. Therefore, a comprehensive analysis of the relevance of such connections, as well as on the putative effects of deregulations on circadian output and resulting pathological phenotypes, is needed.

3.3. Time-regulation as a hall mark of cancer

Aberrant function of the circadian clock has been shown to be correlated with generation and/or progression of cancer [3,70,71]. Clinical studies indicate that, on the one hand cancer patients with disturbed circadian rhythms have a worse prognosis [2], and on the other hand the administration of chemotherapy at specific times of the day (Chronotherapy) can improve treatment efficacy [1]. Moreover, circadian rhythms control several molecular processes such as cell growth and survival, DNA damage responses and energy metabolism. Disruption of such processes can lead to aberrant cellular proliferation and is linked to cancer [72,73].
From a molecular point of view several core-clock genes are directly associated to tumorigenesis. As already mentioned, *Per* levels are deregulated in several cancers [74] *Per1* overexpression has been shown to suppress the growth of human cancer cell lines. *Per2* mRNA levels are down regulated in several human lymphoma cell lines and in non-small-cell lung cancer tissues [70]. Casein kinase epsilon shows increased expression levels in leukaemia and prostate cancer [75]. Mutations in *Npas,* are linked with higher risk of breast cancer and non-Hodgkin lymphoma [3,76]. SIRT1 can also both promote and suppress tumour growth depending on its interacting partners [72,77]. The clock also controls the cell cycle check point regulators *Wee1*, *Myc*, and *cyclin D1* [71,73] and as such cell division. Also cellular metabolism is under circadian control. Altogether these results suggest the existence of a direct link between cancer and perturbations of core-clock genes which was further confirmed by our findings. We could show that RAS overexpression is associated with a

deregulated clock which is found in metastasis in a skin cancer model system and in colon cancer cell line models. Many questions regarding the exact mechanisms via which this regulation is achieved remain open though. Ongoing targeted experiments on my group aim to dissect such mechanisms.

3.4. A time for splicing

A coupling between core-clock elements and splicing factors has also been recently described. The splicing factors hnrnpQ, and PTB can modulate oscillations in the clock gene *Rev-Erb* [7]; hnrnpD has been shown to control stability of *Cry1* [5] and circadian binding of splicing factor SAFA to chromatin has been described as being required for the circadian transcriptional regulation of *Bmal1* [6].

Clock-controlled genes (CCGs) have also been shown by me and others to be involved in many cellular processes [42]. CCGs are regulated at the transcriptional level by CLOCK-BMAL and SIRT1, and are involved in the cell cycle, in proliferation and metabolism. Interestingly, many of these CCGs such as MDM2, VEGF and glycolitic enzyme pyruvate kinase show alternative patterns of splicing which correlate with cancer progression. We are still in the beginning of investigating how time may play a role in the splicing process and whether this may indeed act on cancer progression. The data produced so far is intriguing and the usage of time-course transcriptome arrays complemented with deep-sequencing techniques is ongoing and will eventually allow us to better understand and eventually unravel this process.

3.5. Regulation of alternative splicing

Alternative splicing of pre-mRNAs is a potent and adaptable controlling apparatus that can disturb the regulation of gene expression and diversification of proteins. It influences main developmental choices and also refinement of gene function. Genetic and biochemical methodologies identified cis-acting regulatory elements and trans-acting factors that regulate alternative splicing of specific pre-mRNAs. Some alternative splicing choices are regulated by specific factors. Splicing factors have different roles in both constitutive splicing and fine-tuning of alternative splicing. The combinatorial binding of splicing factors at regulatory elements of genes is crucial for generating specificity and also variation in cell-type-specific alternative splicing patterns.

In principle, negative regulation of alternative splicing could be achieved by a factor that binds to a certain sequence of the pre-mRNA and blocks access of the spliceosome to a 5’ or 3’ splice site, but in most cases, the regulated splice sites must be used in at least some cell types or stages. This may constrain the degree to which the sequence and organization of splice sites can vary without requiring compensatory regulatory mechanisms to ensure

activation when appropriate. Binding to non-target splice sites must be avoided, as well as binding to authentic targets in the wrong context. This may require cooperative or antagonistic interactions, particularly if the repressors are not highly sequence-specific. Many examples of inhibitory *cis*-acting elements were described which show that splice sites can be obstructed by secondary structure by elements binding to regulatory regions within introns or exons, or by the splicing machinery itself when it binds to regulatory elements near the repressed sites.

Also positive regulators have been described, one group of which consists of SR proteins. They have been shown to influence splicing decisions in a dose-dependent manner, through stabilization of U1 snRNP binding or through binding to exonic enhancers.

In many cases, the key determinant of splice site choice is the ratio between the concentrations of factors with opposing effects. The ability of SR proteins to promote the use of proximal 5' splice sites is antagonized by hnRNP A1 and related proteins, in vitro and in transfected cells. Over the last years further examples have been found in which cell type specific ratios between SR proteins and hnRNP proteins determine the level of a given mRNA. The molar ratio of hnRNP A1 to SF2/ASF varies over a range of at least 100-fold among different rat tissues, well beyond what suffices to induce a switch between some alternative 5' splice sites in vitro [78].

hnRNPs can function as positive or negative regulators, depending on what type of hnRNP protein and on where, on the pre-mRNA, the binding site is located, for the interaction with the RNA.

Differences in the ratios of splicing proteins may alter the splice-site decision. The relative concentration levels of the different splicing proteins (splicing factors) are thought to define a molecular code for splicing decisions, this fine-tuning can easily be disrupted in pathological conditions [19]. Some examples of this mechanism are described below and constitute one of the motivations for our future work.

Serine-Arginine-rich proteins show different concentrations, depending on the progression stage of tumours and these result on the inclusion/exclusion of exons in CD44 RNA, known as a tumour metastasis-related protein [79]. A choice between mutually exclusive exons in the FGF Receptor 2 alters its specificity to bind growth-stimulating proteins in prostate cancer [80]. The vascular endothelial growth factor (VEGF) can act as pro-angiogenic, stimulating cancer progression or anti-angiogenic protein, depending on the splice site choice. These alternative splice forms of VEGF are regulated by the SR protein SF2/ASF. This protein is phosphorylated by a SR protein kinase (SRPK1/2) and the inhibition of the kinase was shown to reduce angiogenesis which is a critical process for cancer progression and metastasis [81]. The protein SF2/ASF, has also been reported as a potent oncogene able to induce changes in AS of tumour suppressors that lead to the activation of the mTORC1 pathway [15]. The p16/p14ARF locus (also known as INK4a/ARF) encodes two protein products [82]. The transcript encoding p16 is directed by an upstream promoter and is composed of three exons (1α, 2 and 3). The mRNA encoding p14ARF is transcribed from a promoter downstream of exon 1α, and fuses an alternative exon 1 (1β) to a 3' splice site

within exon 2. Exon 1β is ignored in the longer transcripts because it lacks a 3' splice site at its 5' end and because the distance between the 5' end of the transcript and the exon 1β 5' splice site is too long for the stimulatory effect of the cap structure to be noticeable [83]. p16/p14ARF exemplifies a case of alternative splicing directed by different promoters. P16 is a cyclin-dependent kinase inhibitor, while p14ARF has a totally different amino acid sequence and binds and sequesters the MDM2 protein, which targets the tumour suppressor p53 for degradation [84]. The genetic activity of this locus is at the centre of two critical cell cycle regulatory pathways, those of p53 and Retinoblastoma (Rb). Extensive analysis have shown that the products of the p16/p14ARF locus are affected in a wide variety of cancers [84].

Our results registered a reduction in p16 mRNA levels in at least in some of the Hodgkin's lymphoma cell lines. Intriguingly, signals associated with the corresponding splice junctions did not vary, at least to the same extent. A possible explanation for the discrepancy is that while probes designed to detect exons typically cover 15 non-overlapping 25 nucleotide sequences, those detecting splice junctions correspond to shifts of a few nucleotides around the 25 nucleotides of the junction and are therefore less versatile to provide significantly different sequence information. In contrast to the decrease in p16 levels, measures specific for transcripts encoding p14ARF are apparently unchanged. This is in contrast to previous reports indicating that p14ARF levels are decreased in other Hodgkin cell lines [85]. Although the significance of these observations for the genetic activity of the locus are not clear, the results at least suggest that the microarray results have the potential to detect differences in the behaviour of closely related transcripts.

The CD44 gene encodes for a family of cell adhesion molecules acting in a variety of cell types, involved both in homotypic and in heterotypic interactions with extracellular matrix components [86]. It has been proposed to serve as a lymph node homing receptor on circulating lymphocytes. The genomic structure of the human gene consists of 20 exons, ten of which are included or skipped in a variable fashion [87]. The precise combination of these exons (v1 to v10) included in the mature transcript defines the attachment properties of the molecule. Multiple isoforms have been detected in different tissues, during development and in disease [88]. Different CD44 isoforms have been implicated in cancer metastasis. For example, [89] demonstrated that inclusion of exon v6 was sufficient to confer metastatic potential to tumour cells injected in rats. Exon v5 has been shown to be included in response to Ras and MAP kinase signalling [90] and this to trigger phosphorylation of the splicing adaptor protein Sam 68 [91].

MDM2 binds and enables degradation of the tumour suppressor p53. Over 40 splice variants of MDM2 transcripts were reported in tumours and normal tissues. Over-expression of MDM2 in acute lymphoblastic leukaemia is associated with an unfavourable prognosis [84]. In contrast, in tumours of the head and neck region, the loss of MDM2 expression is associated with a poor prognosis [92]. A possible explanation of these findings may be that alternatively spliced MDM2 variants are expressed in certain tumours and, when present, influence prognosis. It seems that some splice variants of MDM2 exhibit an antiapoptotic

function. This would explain why some patients whose tumours contain p53 mutations and over-express MDM2 have a worse prognosis than do those whose tumours have only one of these modifications [84]. In ''normal'' cells, MDM2 inhibits the function of wild-type p53, stimulating cell growth. If the MDM2 splice variants, mentioned above, are present, they bind to full-length MDM2 protein, releasing p53, causing growth inhibition. In cancer cells, when p53 may be mutated or not expressed, MDM2 can still lead to cell growth. One *in vivo* demonstration of how AS can contribute to the proliferative capacity of tumours concerns a switch between mutually exclusive exons in the gene encoding the glycolitic enzyme pyruvate kinase. This switch enables tumour cells to maintain elevated levels of glucose uptake and lactate production under aerobic conditions, which facilitates tumour growth [16].

A systems biology approach, including high-throughput sequencing of exons and introns, combined with proteomics data to comprehensively analyse the different splicing factors and their relative levels, will be essential. Such data can be brought together using powerful bioinformatics and statistical analysis. This will provide the basis for hypothesis generation and identification of interesting splicing decisions which then need to be verified at the cellular level, for particular genes and splicing events. With such an approach, we might get closer to understand the complexity and the power of alternative splicing to generate a multitude of functional proteins from a template RNA.

3.6. Clock and cancer progression: a role for splicing?

Several reports have been published describing cycling of alternative transcripts in different species and also indicating known cancer genes, with alternative splice pattern, to be circadian regulated [35,93,94]. Taken together these results point to a coupling between the circadian clock and splicing and its effect in cancer progression.

The existence of a strong cross-regulation between the circadian system and splicing is very likely, and its consequent effect on cancer regulation has noteworthy suggestions for future work. To understand the links between the circadian clock and splicing requires a systems analysis at the tissue and cell levels of hundreds of proteins and their relative concentrations, as well as of their target RNA sequences, which together define the cellular codes that rule splicing. Such an analysis will require the integration of bioinformatics, mathematical modelling and genome wide screening of gene expression.

In my group we are investigating the hypothesis that oscillations of splicing factors are necessary for the correct production of a specific protein at a specific day-time; that a disruption of such regulation interferes in several molecular processes, among which, metabolism, cell cycle and proliferation and is directly associated with tumour progression. In normal healthy cells, energy is generated via oxidative phosphorylation in the mitochondria with release of CO_2. Oxidative phosphorylation is the major producer of ROS (reactive oxygen species), known to be toxic to the cells and induce DNA damage, leading to

apoptosis. In yeast for instance the production of ROS is stopped during S phase [95], which is the stage of the cell cycle where DNA is more exposed, and maximized afterwards. In proliferating cells and tumour cells there is a switch in the energy production mechanism and aerobic glycolysis is used instead (Warburg effect). One advantage is that this process does not lead to high levels of ROS which in cells undergoing division could lead to DNA damage and consequently apoptosis. This switch is generated by an alternative splice form of the glycolitic enzyme pyruvate kinase which allows cells to maintain high levels of glucose uptake under aerobic conditions, as already described [96]. This process leads to production of NADH and is needed for fatty acid synthesis and nucleotide synthesis, crucial for proliferating cells.

The production of NADH also has consequences at the circadian clock level; NADH has been proposed to affect circadian regulation. *In vitro* NADH can enhance the binding of CLOCK:BMAL to its target sequences. Also SITR1 a core clock protein is dependent on NAD+ and influences the stability of PER2 and activity of BMAL1 [77,97]. Additionally, SIRT1, has also been reported as a regulator of PPARγ [13] which in turn can associated with SF2/ASF [8], a slicing factor implicated in many alternative splicing proliferative events, and promote the transcription of the selected splice variant, as previously described.

The above described processes give insights regarding the coupling of splicing and of the circadian clock and its contribution for cell proliferation and transformation. We use stage-specific cell-lines to analyse progression-specific differences in oscillations and in splicing patterns. A genome wide screening using exon-arrays can allow the identification of specific splice forms of genes, at a specific time point. It would be conceivable that a certain splice factor could have a different oscillatory pattern in different stages of progression. This effect could be an outcome of a malfunctioning clock. Ultimately, the consequence could be the selection of a certain alternative splice form of a gene, leading to a protein able to promote malignant transformation. This succession of events would lead the tumour to evolve to more aggressive stages. The work here described concerns a comprehensive system biology approach, to couple circadian molecular processes with alternative splicing and tumourigenesis. It has the potential to contribute to the improvement of diagnostic and therapy optimization by investigating the potential role of the circadian system which may lead to future applications of chronoterapy to cancer treatment

4. Conclusions and perspectives

This work represents an interdisciplinary approach to the coupling of the circadian clock, splicing and tumour progression. Results obtained are and will continue to be submitted to scientific journals and will be of use to both the academic scientific community and the pharmaceutical industry.

We are working on the further development of the circadian core-clock network [98] into a splicing-circadian-network using text mining and manual curation. Ultimately, we will

generate a large network of circadian regulation of splicing in tumour progression. We are also designing a gene expression data base with all information contained in the network [99,100]. Oscillatory patterns in published microarrays studies regarding mammalian cells and tissues [42] will be integrated. We are further developing methods to identify clock controlled genes and extract the phases and amplitudes from all genes present in the splicing network previously designed. Results will be integrated and used to improve the model. In parallel a Boolean model [101-103] of the SPLICE-network will be developed, tested and optimized with published microarray data and later will be updated and validated with our own data. These results will be integrated with a semi-quantitative model to study circadian regulation of splicing in tumour progression, containing the most important coupling elements between the circadian clock, the splicing mechanism and tumourigenesis. An effort will be made to identify circadian tumour progression gene-signatures. The model resultant from this iterative approach - experiments/mathematical/bioinformatics methods, should be able to simulate the levels of different genes and circadian behaviour in a tumour-stage specific way. It should also allow the comparison of healthy *vs*. tumour system and to establish time-intervals when perturbations, due to an external agent (drugs), are high in tumours and minimal in the healthy system.

New hypotheses generated from a further analysis of the group's own array data and from both mathematical models will be tested experimentally. To look for a circadian regulation of splicing-signature we use exon-microarrays. We screen RNA extracted from cancer cell lines, with 12 time points per cell line (synchronization of cells and RNA collection at the different time points). The oscillatory capability of all cell lines, in culture, will be compared and the ones with better defined oscillatory pattern will be used. The results will be taken to refine the model. The new hypothesis which might be inferred from results of the model will be experimentally validated at the RNA and at the protein levels using bioluminescence assays and western blots.

All information will be used in the praxis to help in providing a more precise and individual diagnostic, based on a genetic signature of stage-specific tumorigenesis. Results involving simulations of the effects of different sorts of perturbations will be applicable to chronotherapy. This would offer the possibility to dynamically predict the effect of intake-time of different drugs. Therefore, results from our study have the potential of being relevant to clinicians. As a perspective, our group hopes to contribute with valuable information to the improvement of diagnosis and prognosis in cancer, potentially leading to the identification of novel drug-targets and optimization of therapy.

Furthermore, insights regarding the cellular mechanisms, which regulate circadian decisions, in tumour progression could be attained. Such results are also of great interest to the pharmaceutical industry: 1) our published models could be applied to pharmacokinetics and pharmacodynamics studies; 2) a circadian stage specific signature defined by the group could be further developed as a precise diagnostic tool; 3) relevant genes found in our high throughput screenings could be further exploited as drug targets and 4) results from our chronobiology studies could be used to optimize the intake time of drugs in order to reduce

toxicity. Insights regarding the cellular mechanisms, which regulate circadian decisions of alternative splicing, in tumour progression could be attained.

I expect that results of the work in my group will be able to provide answers to many scientific questions and to produce meaningful data applicable to health research. However, I also expect to find many perspectives of future research work such as: new questions related to the coupling of the circadian clock and splicing; new relevant genes in tumourigenesis and associations between circadian dynamics of gene expression involved in malignancy and the cellular codes of alternative splicing.

5. References

1. Levi F, Okyar A, Dulong S, Innominato PF, Clairambault J Circadian timing in cancer treatments. Annu Rev Pharmacol Toxicol 50: 377-421.
2. Lis CG, Grutsch JF, Wood P, You M, Rich I, et al. (2003) Circadian timing in cancer treatment: the biological foundation for an integrative approach. Integr Cancer Ther 2: 105-111.
3. Hoffman AE, Zheng T, Stevens RG, Ba Y, Zhang Y, et al. (2009) Clock-cancer connection in non-Hodgkin's lymphoma: a genetic association study and pathway analysis of the circadian gene cryptochrome 2. Cancer Res 69: 3605-3613.
4. Relogio A, Ben-Dov C, Baum M, Ruggiu M, Gemund C, et al. (2005) Alternative splicing microarrays reveal functional expression of neuron-specific regulators in Hodgkin lymphoma cells. J Biol Chem 280: 4779-4784.
5. Woo KC, Ha DC, Lee KH, Kim DY, Kim TD, et al. Circadian amplitude of cryptochrome 1 is modulated by mRNA stability regulation via cytoplasmic hnRNP D oscillation. Mol Cell Biol 30: 197-205.
6. Onishi Y, Hanai S, Ohno T, Hara Y, Ishida N (2008) Rhythmic SAF-A binding underlies circadian transcription of the Bmal1 gene. Mol Cell Biol 28: 3477-3488.
7. Kim DY, Woo KC, Lee KH, Kim TD, Kim KT hnRNP Q and PTB modulate the circadian oscillation of mouse Rev-erb alpha via IRES-mediated translation. Nucleic Acids Res 38: 7068-7078.
8. Kim DJ, Oh B, Kim YY (2009) Splicing factor ASF/SF2 and transcription factor PPAR-gamma cooperate to directly regulate transcription of uncoupling protein-3. Biochem Biophys Res Commun 378: 877-882.
9. Takahashi JS, Hong HK, Ko CH, McDearmon EL (2008) The genetics of mammalian circadian order and disorder: implications for physiology and disease. Nat Rev Genet 9: 764-775.
10. Sahar S, Sassone-Corsi P (2009) Metabolism and cancer: the circadian clock connection. Nat Rev Cancer 9: 886-896.
11. Vanselow K, Vanselow JT, Westermark PO, Reischl S, Maier B, et al. (2006) Differential effects of PER2 phosphorylation: molecular basis for the human familial advanced sleep phase syndrome (FASPS). Genes Dev 20: 2660-2672.
12. Marcheva B, Ramsey KM, Buhr ED, Kobayashi Y, Su H, et al. Disruption of the clock components CLOCK and BMAL1 leads to hypoinsulinaemia and diabetes. Nature 466: 627-631.
13. Picard F, Kurtev M, Chung N, Topark-Ngarm A, Senawong T, et al. (2004) Sirt1 promotes fat mobilization in white adipocytes by repressing PPAR-gamma. Nature 429: 771-776.
14. Staiger D, Koster T Spotlight on post-transcriptional control in the circadian system. Cell Mol Life Sci 68: 71-83.
15. Karni R, de Stanchina E, Lowe SW, Sinha R, Mu D, et al. (2007) The gene encoding the splicing factor SF2/ASF is a proto-oncogene. Nat Struct Mol Biol 14: 185-193.
16. Christofk HR, Vander Heiden MG, Harris MH, Ramanathan A, Gerszten RE, et al. (2008) The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. Nature 452: 230-233.
17. Maniatis T, Reed R (2002) An extensive network of coupling among gene expression machines. Nature 416: 499-506.
18. Caceres JF, Stamm S, Helfman DM, Krainer AR (1994) Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. Science 265: 1706-1709.
19. Mendes Soares LM, Valcarcel J (2006) The expanding transcriptome: the genome as the 'Book of Sand'. EMBO J 25: 923-931.
20. Relogio A, Schwager C, Richter A, Ansorge W, Valcarcel J (2002) Optimization of oligonucleotide-based DNA microarrays. Nucleic Acids Res 30: e51.
21. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860-921.

22. de la Grange P, Gratadou L, Delord M, Dutertre M, Auboeuf D Splicing factor and exon profiling across human tissues. Nucleic Acids Res 38: 2825-2838.
23. Hartmann B, Valcarcel J (2009) Decrypting the genome's alternative messages. Curr Opin Cell Biol 21: 377-386.
24. Grosso AR, Martins S, Carmo-Fonseca M (2008) The emerging role of splicing factors in cancer. EMBO Rep 9: 1087-1093.
25. Young MW, Kay SA (2001) Time zones: a comparative genetics of circadian clocks. Nat Rev Genet 2: 702-715.
26. Panda S, Hogenesch JB, Kay SA (2002) Circadian rhythms from flies to human. Nature 417: 329-335.
27. Hankins MW, Peirson SN, Foster RG (2008) Melanopsin: an exciting photopigment. Trends Neurosci 31: 27-36.
28. Borrelli E, Nestler EJ, Allis CD, Sassone-Corsi P (2008) Decoding the epigenetic language of neuronal plasticity. Neuron 60: 961-974.
29. Chaudhury D, Wang LM, Colwell CS (2005) Circadian regulation of hippocampal long-term potentiation. J Biol Rhythms 20: 225-236.
30. Maronde E, Schilling AF, Seitz S, Schinke T, Schmutz I, et al. The clock genes Period 2 and Cryptochrome 2 differentially balance bone formation. PLoS One 5: e11527.
31. Young MW (2000) The tick-tock of the biological clock. Sci Am 282: 64-71.
32. Doi M, Hirayama J, Sassone-Corsi P (2006) Circadian regulator CLOCK is a histone acetyltransferase. Cell 125: 497-508.
33. Lestas I, Vinnicombe G, Paulsson J Fundamental limits on the suppression of molecular fluctuations. Nature 467: 174-178.
34. Baggs JE, Price TS, DiTacchio L, Panda S, Fitzgerald GA, et al. (2009) Network features of the mammalian circadian clock. PLoS Biol 7: e52.
35. Zhang EE, Liu AC, Hirota T, Miraglia LJ, Welch G, et al. (2009) A genome-wide RNAi screen for modifiers of the circadian clock in human cells. Cell 139: 199-210.
36. Dibner C, Sage D, Unser M, Bauer C, d'Eysmond T, et al. (2009) Circadian gene expression is resilient to large fluctuations in overall transcription rates. EMBO J 28: 123-134.
37. Leloup JC, Goldbeter A Modelling the dual role of Per phosphorylation and its effect on the period and phase of the mammalian circadian clock. IET Syst Biol 5: 44.
38. Bernard S, Cajavec B, Pujo-Menjouet L, Mackey MC, Herzel H (2006) Modelling transcriptional feedback loops: the role of Gro/TLE1 in Hes1 oscillations. Philos Transact A Math Phys Eng Sci 364: 1155-1170.
39. Ko CH, Takahashi JS (2006) Molecular components of the mammalian circadian clock. Hum Mol Genet 15 Spec No 2: R271-277.
40. Nguyen QL, Tikk D, Leser U Simple tricks for improving pattern-based information extraction from the biomedical literature. J Biomed Semantics 1: 9.
41. Tikk D, Thomas P, Palaga P, Hakenberg J, Leser U A comprehensive benchmark of kernel methods to extract protein-protein interactions from literature. PLoS Comput Biol 6: e1000837.
42. Bozek K, Relogio A, Kielbasa SM, Heine M, Dame C, et al. (2009) Regulation of clock-controlled genes in mammals. PLoS One 4: e4882.
43. Becker-Weimann S, Wolf J, Herzel H, Kramer A (2004) Modeling feedback loops of the Mammalian circadian oscillator. Biophys J 87: 3023-3034.
44. Becker-Weimann S, Wolf J, Kramer A, Herzel H (2004) A model of the mammalian circadian oscillator including the REV-ERBalpha module. Genome Inform 15: 3-12.
45. Forger DB, Peskin CS (2003) A detailed predictive model of the mammalian circadian clock. Proc Natl Acad Sci U S A 100: 14806-14811.
46. Forger DB, Peskin CS (2005) Stochastic simulation of the mammalian circadian clock. Proc Natl Acad Sci U S A 102: 321-324.

47. Leloup JC, Goldbeter A (2003) Toward a detailed computational model for the mammalian circadian clock. Proc Natl Acad Sci U S A 100: 7051-7056.
48. Leloup JC, Goldbeter A (2004) Modeling the mammalian circadian clock: sensitivity analysis and multiplicity of oscillatory mechanisms. J Theor Biol 230: 541-562.
49. Mirsky HP, Liu AC, Welsh DK, Kay SA, Doyle FJ, 3rd (2009) A model of the cell-autonomous mammalian circadian clock. Proc Natl Acad Sci U S A 106: 11107-11112.
50. Lowrey PL, Takahashi JS (2011) Genetics of circadian rhythms in Mammalian model organisms. Advances in genetics 74: 175-230.
51. Albrecht U (2012) Timing to perfection: the biology of central and peripheral circadian clocks. Neuron 74: 246-260.
52. Bass J (2012) Circadian topology of metabolism. Nature 491: 348-356.
53. Kondratova AA, Kondratov RV (2012) The circadian clock and pathology of the ageing brain. Nature reviews Neuroscience 13: 325-335.
54. Takahashi JS, Hong H-K, Ko CH, McDearmon EL (2008) The genetics of mammalian circadian order and disorder: implications for physiology and disease. Nature reviews Genetics 9: 764-775.
55. Levi F, Schibler U (2007) Circadian rhythms: mechanisms and therapeutic implications. Annual review of pharmacology and toxicology 47: 593-628.
56. Saini C, Suter DM, Liani A, Gos P, Schibler U (2011) The mammalian circadian timing system: synchronization of peripheral clocks. Cold Spring Harbor symposia on quantitative biology 76: 39-47.
57. Bollinger T, Schibler U (2014) Circadian rhythms - from genes to physiology and disease. Swiss Med Wkly 144: w13984.
58. Relogio A, Westermark PO, Wallach T, Schellenberg K, Kramer A, et al. (2011) Tuning the mammalian circadian clock: robust synergy of two loops. PLoS computational biology 7: e1002309.
59. Ueda HR, Hayashi S, Chen W, Sano M, Machida M, et al. (2005) System-level identification of transcriptional circuits underlying mammalian circadian clocks. Nature genetics 37: 187-192.
60. Zhang EE, Kay SA (2010) Clocks not winding down: unravelling circadian networks. Nature reviews Molecular cell biology 11: 764-776.
61. Rey G, Cesbron Fc, Rougemont J, Reinke H, Brunner M, et al. (2011) Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. PLoS biology 9: e1000595.
62. Koike N, Yoo S-H, Huang H-C, Kumar V, Lee C, et al. (2012) Transcriptional Architecture and Chromatin Landscape of the Core Circadian Clock in Mammals. Science (New York, NY).
63. Hughes ME, DiTacchio L, Hayes KR, Vollmers C, Pulivarthy S, et al. (2009) Harmonics of circadian gene transcription in mammals. PLoS genetics 5: e1000442.
64. Feng D, Liu T, Sun Z, Bugge A, Mullican SE, et al. (2011) A circadian rhythm orchestrated by histone deacetylase 3 controls hepatic lipid metabolism. Science (New York, NY) 331: 1315-1319.
65. Cho H, Zhao X, Hatori M, Yu RT, Barish GD, et al. (2012) Regulation of circadian behaviour and metabolism by Rev-Erb-α and Rev-Erb-β. Nature 485: 123-127.
66. Bugge A, Feng D, Everett LJ, Briggs ER, Mullican SE, et al. (2012) Rev-erb α and Rev-erb β coordinately protect the circadian clock and normal metabolic function. Genes & development 26: 657-667.
67. Korencic A, Kosir R, Bordyugov G, Lehmann R, Rozman D, et al. (2014) Timing of circadian genes in mammalian tissues. Sci Rep 4: 5782.
68. Wallach T, Schellenberg K, Maier B, Kalathur RKR, Porras P, et al. (2013) Dynamic circadian protein-protein interaction networks predict temporal organization of cellular functions. PLoS genetics 9: e1003398.

69. Relogio A, Thomas P, Medina-Perez P, Reischl S, Bervoets S, et al. (2014) Ras-mediated deregulation of the circadian clock in cancer. PLoS genetics 10: e1004338.
70. Gery S, Komatsu N, Baldjyan L, Yu A, Koo D, et al. (2006) The circadian gene per1 plays an important role in cell growth and DNA damage control in human cancer cells. Mol Cell 22: 375-382.
71. Fu L, Pelicano H, Liu J, Huang P, Lee C (2002) The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo. Cell 111: 41-50.
72. Brooks CL, Gu W (2009) How does SIRT1 affect metabolism, senescence and cancer? Nat Rev Cancer 9: 123-128.
73. Matsuo T, Yamaguchi S, Mitsui S, Emi A, Shimoda F, et al. (2003) Control mechanism of the circadian clock for timing of cell division in vivo. Science 302: 255-259.
74. Chen ST, Choo KB, Hou MF, Yeh KT, Kuo SJ, et al. (2005) Deregulated expression of the PER1, PER2 and PER3 genes in breast cancers. Carcinogenesis 26: 1241-1246.
75. Yang WS, Stockwell BR (2008) Inhibition of casein kinase 1-epsilon induces cancer-cell-selective, PERIOD2-dependent growth arrest. Genome Biol 9: R92.
76. Yi CH, Zheng T, Leaderer D, Hoffman A, Zhu Y (2009) Cancer-related transcriptional targets of the circadian gene NPAS2 identified by genome-wide ChIP-on-chip analysis. Cancer Lett 284: 149-156.
77. Nakahata Y, Sahar S, Astarita G, Kaluzova M, Sassone-Corsi P (2009) Circadian control of the NAD+ salvage pathway by CLOCK-SIRT1. Science 324: 654-657.
78. Hanamura A, Caceres JF, Mayeda A, Franza BR, Jr., Krainer AR (1998) Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. RNA 4: 430-444.
79. Bajorath J (2000) Molecular organization, structural features, and ligand binding characteristics of CD44, a highly variable cell surface glycoprotein with multiple functions. Proteins 39: 103-111.
80. Carstens RP, Eaton JV, Krigman HR, Walther PJ, Garcia-Blanco MA (1997) Alternative splicing of fibroblast growth factor receptor 2 (FGF-R2) in human prostate cancer. Oncogene 15: 3059-3065.
81. Nowak DG, Amin EM, Rennel ES, Hoareau-Aveilla C, Gammons M, et al. Regulation of vascular endothelial growth factor (VEGF) splicing from pro-angiogenic to anti-angiogenic isoforms: a novel therapeutic strategy for angiogenesis. J Biol Chem 285: 5532-5540.
82. Haber DA (1997) Splicing into senescence: the curious case of p16 and p19ARF. Cell 91: 555-558.
83. Lewis JD, Izaurralde E, Jarmolowski A, McGuigan C, Mattaj IW (1996) A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. Genes Dev 10: 1683-1698.
84. Wurl P, Meye A, Schmidt H, Lautenschlager C, Kalthoff H, et al. (1998) High prognostic significance of Mdm2/p53 co-overexpression in soft tissue sarcomas of the extremities. Oncogene 16: 1183-1185.
85. Garcia JF, Villuendas R, Sanchez-Beato M, Sanchez-Aguilera A, Sanchez L, et al. (2002) Nucleolar p14(ARF) overexpression in Reed-Sternberg cells in Hodgkin's lymphoma: absence of p14(ARF)/Hdm2 complexes is associated with expression of alternatively spliced Hdm2 transcripts. Am J Pathol 160: 569-578.
86. Isacke CM, Yarwood H (2002) The hyaluronan receptor, CD44. Int J Biochem Cell Biol 34: 718-721.
87. Cooper DL, Dougherty GJ (1995) To metastasize or not? Selection of CD44 splice sites. Nat Med 1: 635-637.
88. Screaton GR, Bell MV, Jackson DG, Cornelis FB, Gerth U, et al. (1992) Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. Proc Natl Acad Sci U S A 89: 12160-12164.
89. Gunthert U, Hofmann M, Rudy W, Reber S, Zoller M, et al. (1991) A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells. Cell 65: 13-24.

90. Weg-Remers S, Ponta H, Herrlich P, Konig H (2001) Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway. EMBO J 20: 4194-4203.
91. Matter N, Herrlich P, Konig H (2002) Signal-dependent regulation of splicing via phosphorylation of Sam68. Nature 420: 691-695.
92. Millon R, Muller D, Schultz I, Salvi R, Ghnassia JP, et al. (2001) Loss of MDM2 expression in human head and neck squamous cell carcinomas and clinical significance. Oral Oncol 37: 620-631.
93. Miller BH, McDearmon EL, Panda S, Hayes KR, Zhang J, et al. (2007) Circadian and CLOCK-controlled regulation of the mouse transcriptome and cell proliferation. Proc Natl Acad Sci U S A 104: 3342-3347.
94. Grundschober C, Delaunay F, Puhlhofer A, Triqueneaux G, Laudet V, et al. (2001) Circadian regulation of diverse gene products revealed by mRNA expression profiling of synchronized fibroblasts. J Biol Chem 276: 46751-46758.
95. Chen Z, Odstrcil EA, Tu BP, McKnight SL (2007) Restriction of DNA replication to the reductive phase of the metabolic cycle protects genome integrity. Science 316: 1916-1919.
96. Vander Heiden MG, Christofk HR, Schuman E, Subtelny AO, Sharfi H, et al. Identification of small molecule inhibitors of pyruvate kinase M2. Biochem Pharmacol 79: 1118-1124.
97. Asher G, Gatfield D, Stratmann M, Reinke H, Dibner C, et al. (2008) SIRT1 regulates circadian clock gene expression through PER2 deacetylation. Cell 134: 317-328.
98. Layek RK, Datta A, Dougherty ER From biological pathways to regulatory networks. Mol Biosyst 7: 843-851.
99. Mani KM, Lefebvre C, Wang K, Lim WK, Basso K, et al. (2008) A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas. Mol Syst Biol 4: 169.
100. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, et al. (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403: 503-511.
101. Teschendorff AE, Gomez S, Arenas A, El-Ashry D, Schmidt M, et al. Improved prognostic classification of breast cancer defined by antagonistic activation patterns of immune response pathway modules. BMC Cancer 10: 604.
102. Eduati F, Corradin A, Di Camillo B, Toffolo G A Boolean approach to linear prediction for signaling network modeling. PLoS One 5.
103. Park I, Lee KH, Lee D Inference of combinatorial Boolean rules of synergistic gene sets from cancer microarray datasets. Bioinformatics 26: 1506-1512.

Acknowledgments

I am deeply grateful to the many brilliant scientists with whom I had and have the pleasure to work with: Prof. Dr. Christine Sers, Prof. Dr. Achim Kramer, Prof Dr. Ulf Leser, Prof. Dr. Hanspeter Herzel, Prof. Dr. Juan Valcarcel and members of their groups. With their knowledge and support many fascinating themes could be investigated and new biological mechanisms revealed. I would also like to acknowledge Prof. Dr. Clemens Schmitt for his support and mentoring as part of the mentoring program of the Charité and as director of the Berlin School of Integrative Oncology (BSIO). I am also grateful to members of the Molecular Cancer Research Centre of the Charité and to members of the BSIO for technical help.

To my many students, the future medical doctors, the students of the molecular medicine masters program of the Charité and many others who I thought along the way, goes a warm thank you for the challenge of teaching, for the questions and for the willing of learning which made teaching them a pleasure and an honour. To the members of my group who face every day together with me the scientific challenges the ups and downs of experiments, the development of computational models and new techniques, for their fantastic work I am as well most thankful.

I was fortunate to have financial support throughout my still young career. As a PhD and Post-doc from the Portuguese Ministry for Science and Technology and from the European Molecular Biology Laboratory. As an independent scientist I was supported by the Rahel-Hirsch grant of the Charité and currently my group is financed by a large grant from the BMBF, with additional financial support from the BSIO and from the Institute for Theoretical Biology. To all the funding parts without whom research would not be possible goes my deep gratitude.

I am extremely grateful to the Charité for the opportunity to work, teach and learn in what is not only the largest medical university in Europe, but also an internationally renowned research institution. I have been fortunate to develop here my work and I could as well contribute to the scientific and academic environment with my teaching and research and for these reasons I also feel honoured to carry out my habilitation in this remarkable university.

Finally, to my family, Anna, Philipp and Tobias for their continuous support and understanding and for making everything possible, goes my most deeply THANK YOU!

Erklärung

§ 4 Abs. 3 (k) der HabOMed der Charité

Hiermit erkläre ich, dass

- weder früher noch gleichzeitig ein Habilitationsverfahren durchgeführt oder angemeldet wurde,
- die vorgelegte Habilitationsschrift ohne fremde Hilfe verfasst, die beschriebenen Ergebnisse selbst gewonnen sowie die verwendeten Hilfsmittel, die Zusammenarbeit mit anderen Wissenschaftlern/Wissenschaftlerinnen und mit technischen Hilfskräften sowie die verwendete Literatur vollständig in der Habilitationsschrift angegeben wurden,
- mir die geltende Habilitationsordnung bekannt ist.

Ich erkläre ferner, dass mir die Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis bekannt ist und ich mich zur Einhaltung dieser Satzung verpflichte.

…… ……..

Datum Unterschrift