A High–Throughput Analysis of Circadian Protein Stability

Inaugural–Dissertation

to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the
Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by

M.Sc. KATJA SCHELLENBERG

born in Hohenmölsen

2013

This thesis was conducted from 2009 to 2013,
supervised by Prof. Dr. rer. nat. Achim Kramer,
at the Laboratory of Chronobiology,
Institute for Medical Immunology, Charité Berlin

1st Reviewer: Prof. Dr. rer. nat. Achim Kramer, Charité Berlin
2nd Reviewer: Prof. Dr. rer. nat. Markus Wahl, Freie Universität Berlin

date of defence: 24.02.2014

Success is not final,
failure is not fatal:
it is the courage to continue that counts.

–Sir Winston Churchill–

Abstract

Circadian clocks are endogenous oscillations that drive 24–hour rhythms of physiology and behavior. Circadian rhythms exist in nearly all cells of the body with a gene–regulatory network as fundamental clockwork. These molecular clocks not only drive the rhythmic transcription of about ten percent of all genes but also seem to regulate rhythmic protein abundances by mechanisms beyond transcriptional control. Indeed, up to twenty percent of the mammalian proteome are circadian. The extent of post–transcriptional processes controlling circadian protein abundance, however, is hardly studied so far.

Here, the contribution of *timed degradation* for the regulation of circadian protein rhythms is investigated. To this end, a human protein library was analyzed using a fluorescence based reporter system that measures alterations in protein abundance as readout of altered protein stability. In order to determine time–of–day dependent protein stability, the cellular clock was 'clamped' at a specific circadian phase by the ectopic overexpression of CRY1, a strong negative inhibitor of molecular clock dynamics. Applying the fluorescence based method in cells with a 'clamped' clock in comparison to unsynchronized cells enabled to perform a proteome–wide analysis of circadian stability.

Revealed screen results represent a snapshot of the circadian proteome 'clamped' at a phase of increased CRY1 level. About nine percent of analyzed proteins were identified as **CAAPs** (**C**RY1 mediated **a**ltered **a**bundant **p**roteins), representing proteins with potential circadian abundance as result of timed degradation. Indeed, rhythmically abundant proteins are enriched among CAAPs. Furthermore, in accordance with circadian proteome studies, CAAPs are overrepresented in processes related to vesicular trafficking and mitosis. Interestingly, CAAPs are underrepresented among modifiers of circadian dynamics. This postulates a role of circadian protein stability for the rhythmic fine-tuning of clock output functions rather than for regulatory mechanisms of molecular core clock dynamics.

Together, the results of this study indicate first of all, an unexpected role of rhythmic protein degradation for the control of the circadian proteome, and secondly, suggest that rhythmic protein stability is essential as a timing signal for circadian clock output processes.

Zusammenfassung

Circadiane Uhren sind endogen getriebene Oszillationen, welche 24–Stunden Rhythmen der Physiologie und des Verhaltens steuern. Circadiane Rhythmen existieren in nahezu allen Zellen des Körpers und basieren auf einem genregulatorischen 'Uhrwerk'. Diese molekularen Uhren steuern die rhythmische Transkription von etwa zehn Prozent aller Gene. Darüber hinaus regulieren molekulare Uhren Proteinrhythmen, vermutlich durch zusätzliche Mechanismen neben der transkriptionellen Kontrolle. In der Tat liegen bis zu zwanzig Prozent des Säugetierproteoms circadian abundant vor. In welchem Umfang circadiane Proteinrhythmen durch post–transkriptionelle Prozesse kontrolliert werden, ist jedoch bisher wenig untersucht.

In dieser Studie wurde der Beitrag des *tageszeitspezifischen Abbaus* zur Regulation circadianer Proteinrhythmen untersucht. Dazu wurde eine humane Proteinbibliothek unter Anwendung eines fluoreszenzbasierten Reportersystems untersucht. Das Reportersystem misst Unterschiede in der Proteinmenge als Kennzeichen verschiedener Proteinstabilitäten. Um tageszeitabhängige Proteinstabilität bestimmen zu können, wurde die zellulare Uhr in einer spezifischen Phase 'arretiert'. Dies erfolgte durch die ektopische Überexpression von CRY1, einem stark negativen Inhibitor der molekularen Uhr–Dynamik. Unter Anwendung der fluoreszenzbasierten Methode in Uhr–'arretierten' Zellen im Vergleich zu unsynchronisierten Zellen, konnten circadiane Stabilitäten proteomweit analysiert werden.

Die Ergebnisse der Untersuchung repräsentieren eine Momentaufnahme des circadianen Proteoms, 'arretiert' in einer Phase erhöhter CRY1 Proteinmengen. Etwa neun Prozent aller untersuchten Proteine wurden als **CAAPs** (**C**RY1 mediated **a**ltered **a**bundant **p**roteins; *deutsch*: Proteine mit CRY1 induzierter, veränderter Abundanz) identifiziert. Diese stellen eine Gruppe von Proteinen mit vermutlich circadianer Abundanz als Ergebnis des tageszeitspezifischen Abbaus, dar. In der Tat sind rhythmisch abundante Proteine unter den CAAPs angereichert. In Übereinstimmung mit circadianen Proteomstudien, sind CAAPs in Prozessen des vesikulären Transports und der Mitose überrepräsentiert. Interessanterweise sind CAAPs unter molekularen Komponenten, welche die circadiane Uhr beeinflussen, unterrepräsentiert. Das deutet auf eine Funktion der circadianen Proteinstabilität in der rhythmischen Feinabstimmung Uhr–getriebener Prozesse hin, welche die molekulare Uhr selbst, nicht regulieren.

Zusammenfassend weisen die Ergebnisse dieser Studie dem rhythmischen Proteinabbau eine unerwartete Rolle in der Regulation des circadianen Proteoms zu und deuten des Weiteren daraufhin, dass rhythmische Proteinstabilität als Übermittler von Zeitinformationen in Uhr–gesteuerten Prozessen essenziell ist.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Biological 24–hour Rhythms

Evolution has forced organisms on earth to precisely adapt to environmental conditions. To anticipate daily changes e.g. of light and temperature resulting from the earth's rotation around its axis, internal timing mechanisms have evolved. These so called biological clocks are found in prokaryotic and eukaryotic organisms. They enable a timed implementation of endogenous processes to external conditions, including for example nitrogen fixation, leaf movement, asexual conidiation, eclosion and activity pattern of cyanobacteria, plants, fungi, insects and mammals, respectively[1–5]. Thus, self–sustained circadian (Latin *circa* = about; *dies* = day,[6]) rhythms are efficient mechanisms that increase organism fitness and reduce energy costs[5].

1.1.1 A Brief Insight on Chronohistory

The existence of internal timekeepers was first described in the year 1729. De Mairan described that the heliotrope plant *Mimosa pudica* continued its daily leaf movement in constant darkness. This observation opened the question on the existence of internal timepieces[7]. Following studies on circadian activity rhythms in insects, fungi and rodents either deprived or kept in different light environments, strengthened the idea of internal timing mechanisms[8–12]. Evidence for a genetic basis of an endogenous clock was provided by Bünning, showing that the circadian rhythm length (= period, see subsection 1.1.3) is heritable in bean plants[13]. However, a genetic component of the molecular clock was first identified in the late 20^{th} century, the time of forward genetics and large scale approaches. Here, Konopka and Benzer identified in a chemical mutant screen in *Drosophila melanogaster* various mutant alleles of the same genome locus resulting in either longer or shorter circadian periods or arhythmicity of pupal eclosion and locomotor activity[14]. The encoded gene *Period*

(*Per*) was identified about ten years later[15]. Likewise in mammals, the *Circadian Locomotor Output Cycles Kaput* (*Clock*) gene, important to sustain circadian rhythms, was discovered in a large mutant screen in mice. Here, a semi–dominant mutation in the *Clock* gene locus caused a period lengthening of mouse locomotor activity[16].

1.1.2 Properties of Circadian Rhythms

Investigations across species defined three essential features that describe circadian rhythms. First of all, circadian rhythms are self–sustained, cell–autonomous, endogenous cycles of approximately 24–hours that persist in constant conditions. Second, although being self–sustained, circadian rhythms are responsive towards external time cues. Such so called 'Zeitgebers' are able to synchronize (thus entrain, see subsection 1.1.3) endogenous rhythms to environmental stimuli like the light–dark cycle or temperature changes. As a third characteristic, circadian rhythms are temperature compensated, meaning that their 24–hour period remains relatively constant over a broad range of ambient temperatures[17,18,2].

1.1.3 Circadian Terminology

The following box explains and visualizes main terms of circadian biology.

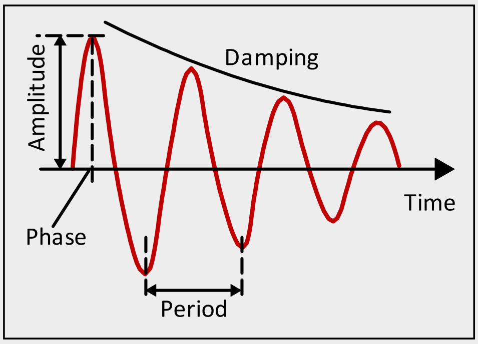


Figure 1.1: Visualization of circadian terms.

Amplitude — The difference between the peak (or trough) and the mean value of a wave (see Fig.1.1).

Damping — The process over time in which the amplitude of an oscillation flattens. Contrary to a self–sustained oscillation (see Fig.1.1).

Entrainment — The process of an oscillator to adapt to the time (period) of a rhythmic synchronizing cue (e.g. the light–dark cycle).

Period — The cycle length of an oscillation. For circadian rhythms about 24–hours (see Fig.1.1).

Phase — The peak (or trough) of an oscillation with reference to an external point such as time (see Fig.1.1).

Synchronization — A progress that resets two or more oscillators to the same phase. This can be achieved by entrainment or a single pulse e.g. the administration of a single glucocorticoid pulse to cultured cells.

Zeitgeber — External cue that entrains or synchronizes internal (circadian) rhythms. Modified from[19].

1.2 Mammalian Circadian Clocks

Mammalian physiology follows 24–hour rhythms. Processes like the sleep–wake cycle or the regulation of blood pressure, core body temperature, hormone secretion and metabolic functions are orchestrated by biological clocks. They adjust molecular pathways resulting in a precise regulation in accordance with the environment[20–23]. Consequently, circadian rhythms are essential to health and their disruption is associated with various diseases including sleep disorders, cancer, depression, metabolic syndrome or inflammation[24–26]. Accumulating evidence reveals that especially the asynchrony of our internal molecular clocks with the environment results in adverse health effects, making it of special interest for researchers and clinicians alike to explore basic mechanisms of the mammalian molecular oscillator.

1.2.1 Anatomy and Physiology of Mammalian Circadian Clocks

The mammalian timing system frames three major levels: (i) the input of an external time cue, (ii) the concordant internal synchronization and (iii), an adapted output in phase. These levels are organized in a hierarchical structure with a master pacemaker, the light sensitive suprachiasmatic nucleus (SCN). The SCN entrains peripheral clocks resulting in a circadian regulation of diverse processes with a stable phase relation to the environmental input (see Figure 1.2)[27,28,19].

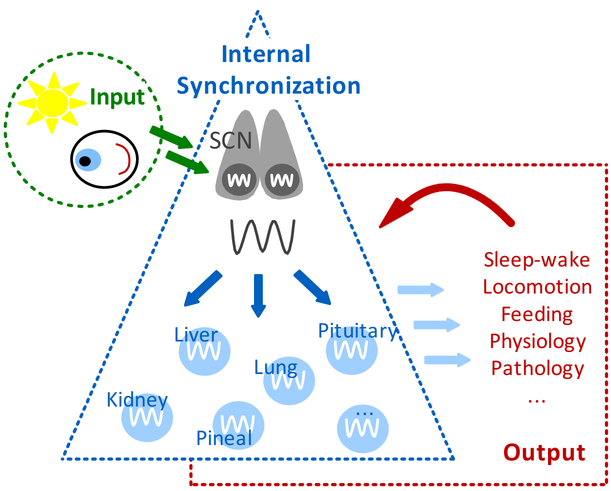


Figure 1.2: Three levels of the mammalian circadian system. Schematic representation of the three major levels of the mammalian circadian timing system. *Input* (green): Daily light information is received by photoreceptors of the eyes' retina and directly forwarded to the suprachiasmatic nucleus (SCN). *Internal synchronization* (blue): In a hierarchically organized manner, the master pacemaker SCN synchronizes peripheral oscillators (e.g. liver, kidney, etc.). *Output* (red): Diverse circadian processes driven by peripheral clocks are in a stable phase relation with the external time. Some of them (feeding, physiology) feed back to the internal synchronization. SCN–suprachiasmatic nucleus. Modified from[19].

The Central Circadian Oscillator

The SCN is a paired structure above the optic chiasm in the anterior ventral hypothalamus. This area was shown to be necessary and sufficient for the generation of circadian activity rhythms in mammals by surgical lesion and transplantation studies[29–31].

The SCN is composed of $\sim$ 20,000 neurons, each containing a cell–autonomous oscillator. However, due to a tight coupling, SCN cells oscillate in a coherent manner and adapt fast to new 'Zeitgeber' signals[27,28].

Input stimuli, the photic information, is received by photosensitive retinal ganglion cells of the eyes' retina that directly project via the retinohypothalamic tract to the SCN. As a master pacemaker, the SCN forwards this time cue to synchronize other brain regions and peripheral oscillators. This is accomplished by autonomic neuronal innervations (sympathetic and parasympathetic) to neighboring brain regions and to peripheral organs, as well as by endocrine and humoral signals. A more indirect synchronization of the periphery is achieved by the modulation of body core temperature and feeding cycles. Both processes are themselves coordinated by the master clock, however can entrain peripheral oscillators as well[27,32].

Interestingly, a light independent, food-sensitive oscillators that can entrain daily rhythms of physiology and behavior in SCN–lesioned rodents has been described, but so far the anatomical location is not defined[33].

Peripheral Clocks

Virtually every cell of the mammalian organism contains a functional molecular clock. Their rhythms persist in isolated primary fibroblast and even in immortalized cultured cells[34,35]. Furthermore, endogenous clocks drive the circadian expression of about 5–10 % of all transcripts. However, the overlap of oscillating clock–controlled genes (CCG) between tissues is surprisingly small[36–40]. In addition, rhythmic abundance of up to 20 % the proteome has been demonstrated[41–45].

Although being self–sustained and cell–autonomous, oscillation of a whole population of single cell clocks dampens very fast *in vitro* due to individual period lengths around 24–hours[46]. *In vivo*, the SCN is required for the synchrony of peripheral clocks in tissues and organs. Moreover, in liver–specific clock 'disrupted' mice, circadian rhythmicity of several transcripts was observed. This indicates that peripheral transcript rhythms are not only synchronized, but in some case as well directly driven by systemic cues orchestrated by the SCN[47]. Nonetheless, peripheral clocks are necessary for tissue physiology, e.g. for glucose homeostasis as shown in an liver–specific clock knockout model[48].

1.2.2 Molecular Architecture of Endogenous Mammalian Clocks

The circadian timing mechanism is generated at the cellular level. The basic principles are conserved across species[49]. Essentially, the core oscillator is based on activating and repressing elements forming transcriptional–translational feedback loops. Positive components activate negative elements that in turn operate as repressors of their own activity (see Fig. 1.3).

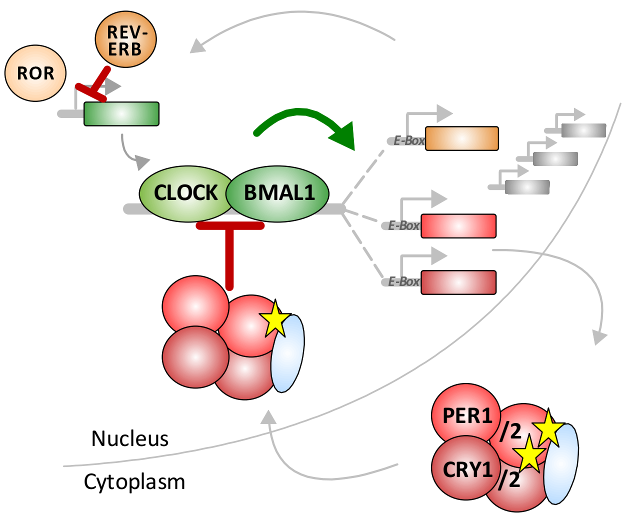


Figure 1.3: Schematic representation of the molecular core clock mechanism. Mammalian circadian rhythms are generated by transcriptional–transcriptional feedback loops. Transcriptional activators CLOCK and BMAL1 (green colors) drive the expression of E–box target genes *Per*, *Cry* (red colors) and *Rev–Erb* (orange), next to further clock controlled genes (gray). PER and CRY proteins form complexes in the cytoplasm and undergo post–translational modifications (yellow stars), primarily mediated by kinases and phosphatases (represented by blue oval). With a certain delay, PER/CRY complexes translocate into the nucleus and inhibit their own transactivation. In contrast to activating ROR proteins, REV–ERB represses the synthesis of its own transactivator *Bmal1* in an additional loop. Abbreviations are explained in subsection 1.2.2. Modified from[50,51].

More specifically, the heterodimeric transcription factor complex of CLOCK and Aryl Hydrocarbon Receptor Nuclear Translocator-like (ARNTL, herein referred as BMAL1) of the positive arm drives the expression of *Period* (*Per1–3*) and *Cryptochrome* (*Cry1–2*) E–box target genes. PER and CRY proteins build up to large complexes that translocate into the nucleus. Here, they directly interfere with CLOCK/BMAL1 mediated transcriptional activity, thereby repressing their own expression[52]. In an auxiliary loop, that is believed to confer robustness to the molecular clock, Nuclear Receptor Subfamily 1, Group D, Member 1 and 2 (NR1D1 and NR1D2, alias REV–ERBα and REV–ERBβ, respectively) directly represses the ROR protein driven gene expression of their own transcriptional activator BMAL1[53,49,51,32].

However, to achieve an approximate 24–hour cycle, certain delay between transcription, translation, complex formation and subcellular localization of positive and negative clock elements is needed. To this end, clock components undergo specific post–translational modifications that time their abundance, formation of complexes, subcellular localization and activity[51,54]. Genome–wide genetic perturbations and pharmacological inhibition screens uncovered the importance of kinases, phosphatases and other enzymes for the generation of about 24–hour periods[55–58]. For example, the F-box protein FBXL3 as part of the Skp1-Cul1 ubiquitin ligase complex mediates the ubiquitination and thus subsequent proteasomal degradation of CRY proteins. Consequently, mutation of loss of FBXL3 leads to a period lengthening at the cellular and behavioral level[59–62].

1.3 The Role of Protein Degradation for Circadian Rhythms

Cellular protein levels are dependent not only on synthesis, but as well on degradation rates. The half–lifes of proteins range from minutes to days[63], whereby proteins that are involved in dynamic regulatory processes are often rather unstable. In this way, they are capable to precisely respond to their environment by quickly changing their cellular abundance levels[64].

Within the molecular clock machinery, targeted degradation of clock proteins is essential to sustain 24–hour rhythms[65]. In addition to the example described above, the *tau* hamster, a circadian mutant with a behavioral activity rhythm of approximately 20–hours, contains a mutation in a kinase (CKIϵ), which is important for the regulation of protein stability of the clock component PER2[66–68].

1.3.1 Protein Degradation Pathways

Protein degradation in eukaryotic cells is mainly conducted by the ubiquitin–proteasome pathway and lysosomal proteolysis (briefly described below).

Lysosomal Proteolysis

One mechanism of cellular protein degradation is lysosomal proteolysis. Membrane-enclosed organelles, the lysosomes, contain diverse digestive enzymes and proteases[64]. The two major pathways for the uptake of proteins into lysosomes are endocytosis and autophagy[69]. Briefly, the endocytic pathway comprises the internalization of plasma membranes. Endocytosed proteins like receptors or channels are either recycled back to the plasma membrane, or in further steps, fused and finally degraded in lysosomes[69]. Within autophagy, membranes derived from the endoplasmic reticulum enclose cytoplasmic particles and organelles in so called autophagosomes, that fuse to lysosomes, resulting in the degradation of their content. This is especially the case for organelles or protein aggregates that cannot be unfolded[70].

In contrast to the specific proteasomal pathway, lysosomal protein degradation appears to be rather non–selective. However, it becomes evident that organelles and structures are selectively removed[70]. Especially in response to cellular starvation, proteins that contain specific consensus sequences are targeted for lysosomal degradation. This possibility involves the binding of chaperones that lead to the unfolding and translocation of proteins across the lysosomal membrane[71,64].

Ubiquitin–Proteasome Pathway

The ubiquitin (Ub)–proteasome pathway is a tightly regulated and highly specific process of targeted protein turnover. In a multistage procedure, proteins are targeted for proteasomal degradation by the attachment of Ub, executed by E1 to E3 enzymes[72].

In an ATP dependent mechanism Ub is loaded on E1 Ub–activating enzymes. In the second step, Ub is transferred to E2 Ub–conjugating enzymes. E3 enzymes are Ub ligases that convey substrate specificity to recognize targeted proteins. Ub can be either directly transferred from E2 enzymes to the substrate while being in complex with an E3 enzyme or are transferred stepwise from E2 via E3 to the substrate protein. Ub is attached to lysine residues and can be further poly-ubiquitinated and thus selectively target proteins for degradation within the multisubunit protease complex, the 26S proteasome.

Whereas most cells contain only one E1 enzyme, they have many E2 and even more E3 enzymes targeting different substrates, thus conveying specificity[64,72]. While some proteins are recognized by primary signals, e.g. degrons within their structure, or via auxiliary proteins, most proteins undergo post–translational modifications that selectively target them for degradation[72].

1.3.2 Specific (and Timed?) Protein Degradation within the Molecular Clock

Protein degradation of most molecular clock components is controlled by the ubiquitin–proteasome pathway. Specific protein turnover is essential to sustain a functional circadian clock work. Not surprisingly, molecular clock components are substrates of diverse enzymes (e.g. kinases or phosphatases) that regulate their for proteasomal degradation[65,54]. Thereby, cellular abundance levels of clock components are precisely controlled over the circadian cycle. This already implies that protein stabilities of core clock components are probably not constant over a 24–hour cycle.

For example, the PER2 protein, a component of the negative feedback loop, contains several phosphorylation sites[73,74]. Dependent on where it is phosphorylated, its stability either increases or decreases, which is essential for maintaining the molecular clock speed[65]. Kinases (e.g. casein kinase 1 epsilon or delta and 2 (CK1ϵ/δ, CK2)) and phosphatases (e.g. protein phosphatase 1) regulate the different phosphorylation patterns[68,73,56,75–77]. These post–translational modifications probably takes place in a time–of–day specific manner, to control the diverse spatio-temporal actions of PER2 and thus might regulate a circadian degradation of PER2 protein. Indeed, although not their major focus, two publications displayed a circadian stability of PER2. Suter *et al.* showed in a supplemental figure that PER2 protein half–life changes over the circadian cycle[78]. Furthermore, Fujimoto *et al.* observed rhythmic PER2 protein levels in synchronized cells constitutively expressing *Per2* mRNA coding sequence driven by an artificial promoter[79]. This indicates a circadian post–translational regulation resulting in time-of day-specific degradation.

Furthermore, CRY proteins of the negative feedback loop are described substrates of kinases (e.g. adenosine monophosphate-activated protein kinase (AMPK), CK1ϵ/δ, DYRK1A or glycogen synthase kinase-3 beta (GSK3β)), that trigger CRY protein degradation[54].

CRY2 protein abundance is a highly rhythmic. Its mRNA levels, however, display only a small amplitude oscillation that, furthermore, peaks already about eight to ten hours before CRY2 protein reaches its maximal abundance within mouse liver tissue[80,53,81]. This discrepancy might be explained by additional modes regulating protein abundance such as circadian post–transcriptional, translational or post–translational events. Indeed, CRY2 is rhythmically phosphorylated by DYRK1A as a priming kinase, followed by GSK3β phosphorylation, thus targeting CRY2 rhythmically for proteasomal degradation[82].

AMPK mediated phosphorylation primes CRY1 for proteasomal degradation. Interestingly, the action of AMPK on CRY1 is limited to a specific spatio-temporal window[83], indicating a circadian degradation and thus time–of–day dependent stability of CRY1 as well.

However, not only circadian post–translational modifications can trigger a time–of–day specific degradation. Furthermore, substrate specific E3 ligases of the ubiquitin–proteasomal pathway (see subsection 1.3.1) have been described to control a time and localization specific degradation of CRY1. Whereas the E3 ligase FBXL3 promotes proteasomal degradation of nuclear CRY1, its paralog FXBL21 regulates CYR1 turnover in the cytoplasm. FBXL21 was described to have a stabilizing effect on CRY1 and even seems to prevent its nuclear degradation[84,59]. Thus, dependent on the subcellular localization of CRY1, which follows a circadian cycles, its protein degradation is altered.

Potential circadian regulation of protein stability is not only found in the negative limb of the molecular clock. BMAL1 and CLOCK of the positive arm are rhythmically phosphorylated[85,86] and are targets of diverse kinases (e.g. CK1ϵ/δ, CK2, GSK3β, mitogen-activated protein kinase (MAPK) and protein kinase C (PKC))[87–92]. For example, GSK3β has a rhythmic activity and its phosphorylation on BMAL1 primes it for proteasomal degradation[90]. Next to phosphorylation, sumoylation as a post–translational modifier has been described. Sumoylation, which occurs rhythmically on BMAL1, triggers its degradation[93]. Recently it was shown that CLOCK and BMAL1 are targets of rhythmic glycosylation that prevent their ubiquitination and thus proteasomal degradation[94,95]. Interestingly, a decreased stability of CLOCK/BMAL1 in the heterodimer coincides with its highest transcriptional activity[96,91]. This is as well proposed in a general model for transcription factors being immediately degraded after activating transcription[97]. Thus, a timed degradation of the circadian transactivators CLOCK and BMAL1 can be assumed.

Altogether, those examples demonstrate that a daytime dependent alteration of stability of components of the core clock is likely. Often it is linked to molecular processes including subcellular localization, protein interaction or activity. Thus, at least within the molecular clock machinery, timed protein degradation seems to be a pervasive attribute. But is this feature as well found at the output of the molecular core clock?

1.3.3 A Circadian 'Stabilome'?

Circadian regulated processes are not only features of the molecular core clock. Large–scale studies revealed that global transcriptional and post–transcriptional mechanisms are circadian[98,99]. In this line, circadian rhythms are found in about 5–10 % of the transcriptome in a given tissue[36–40]. Furthermore, a rhythmic proteome has been shown[41–45]. The existence of a circadian 'stabilome' – that is the time–of–day dependent protein stability and thus protein specific circadian degradation of the proteome – has not be described so far.

Previous results demonstrated that up to 20 % of the hepatic proteome are subjected to circadian control, including pathways of the urea and sugar metabolism[41]. Although, the overall proteome coverage of this and additional studies is low, their results are in discrepancy to only 5–10 % of the transcriptome being rhythmically expressed. Further mass spectrometry analysis of mouse retina and SCN as well as human blood samples identified circadian abundant proteins that surprisingly lack rhythmic transcripts[42–44]. Thus, circadian protein abundance is found also for constitutively transcribed genes. This postulates further rhythmic regulation of post–transcriptional, translational and/or degradation processes.

Next to a circadian abundant proteome that lacks rhythmic expression, hints for a circadian 'stabilome' are found within the degradation pathways themselves. Autophagy by which cytosolic components are conveyed to lysosomal degradation was shown to be circadian[100]. Moreover, genes encoding proteins of the Ub–proteasome pathway are circadian expressed in synchronized rat fibroblasts and mouse liver samples. These include E2 conjugating enzymes, E3 ligases, deubiquitinating enzymes as well as proteasomal subunits[101] (and S. Lück, AG Westermark, ITB Berlin; personal communication about analyzed data from[40]). In addition, Tsuji and colleagues identified a circadian abundant proteasomal subunit with a rhythmic post–translational modification that may regulate its activity[42].

In this context, the first mammalian circadian protein stability of a non–core clock component, the tumor suppressor p53, was demonstrated. Although no observable circadian rhythms in its mRNA, p53 protein levels were rhythmic. Horiguchi and coworkers could show that MDM3 E3 ligase mediated p53 degradation is circadian. This results in a time–of–day dependent stability of p53 and thus circadian abundance levels[102].

1.3.4 Expected Biological and Clinical Relevance of a Circadian 'Stabilome'

Mammalian circadian rhythms are robust yet precisely regulated processes. Orchestrated by the SCN, they are understood to be driven from molecular core clocks that dictate time information to finally result in circadian regulated physiology. To date, this is essentially thought to be a process of circadian synthesis resulting in circadian abundant proteins that drive circadian physiology. Thus, the question arises, what, if at all, might be the biological significance of circadian protein turnover?

Within the molecular core clock, specific circadian regulation at multiple levels including protein turnover is conspicuous. In fact, protein stabilities of core clock proteins have been found to be important for a fine-tuned, non–diseased circadian phenotype. *In vivo* and *in vitro* models showed that altered stabilities of components of the negative loop are linked to changed period lengths of the clock oscillator. Indeed, patients suffering from an inherited form of advanced sleep–phase syndrome (FASPS) have a persistent advanced sleep onset of about 4–hours. This phenotype has been directly linked to an altered phosphorylation of the core clock protein PER2, changing its degradation rate[65]. Features of our modern society including shift-work, air travel and increased nocturnal activity cause circadian dysfunctions. This desynchrony of behavior and physiology from the natural daily cycle is as well characterized by altered periods and/ or phases[65].

Cases are missing to frame general assumptions about the importance of a circadian proteome beyond the core clock. However, circadian stability of p53 makes an example. The tumor suppressor p53 protects cells from uncontrolled proliferation by inducing apoptotic cell death. Indeed, cytotoxicity induced by chemotherapeutic agents was increased at times when p53 protein was most stable.

Altogether, circadian regulated protein stability might display another essential mode of circadian fine tuning. A precise regulation at the level of the proteome might be important for a direct modulation of circadian rhythms, thus adding to the robustness of the circadian system. From a clinical point of view, knowledge about different stabilities of physiological key components might help to understand pathological phenotypes. It will open new targets for pharmacological drugs and approaches for a daytime dependent (chrono-)therapy.

1.4 Aim of this Study

Mammalian circadian clocks are accurate timing devices that control diverse physiological functions. Consequently, their disruption is linked to a variety of diseases including sleep and metabolic disorders, depression and cancer[25]. Molecular clocks persist in almost every cell and drive the circadian expression of up to 10 % of the transcriptome[36–40]. However, in disagreement, 20 % of the proteome is rhythmically abundant[41,44,45,42,43]. This implies further modes of either post–transcriptional, translational and/ or post–translational regulation. Evidence from molecular core clock components (see subsection 1.3.2) and one example of an essential tumor suppressor protein (see subsection 1.3.3) indicate a role for circadian regulation of protein stability by directed and timed protein degradation.

Global circadian stability of the proteome, the level that dictates rhythmic processes of behavior and physiology, has not been studied so far. The existence and extent of a daytime dependent regulation of proteome degradation is unknown.

This PhD project aimed to identify circadian regulated protein stability on a proteome–wide scale. To this end, a fluorescent based reporter method that measures protein abundance as a readout of protein stability should be established and applied to screen about 16.000 proteins of a human open reading frame library. In order to analyze time–of–day dependent degradation the high–throughput approach should be adapted to perform a circadian analysis of protein stability. Screen results should be validated and further analyzed to allow a more comprehensive view on the existence, extent and biological relevance of timed protein degradation.

Investigating global circadian protein stability – the so called circadian 'stabilome' – should reveal a further mode of clock output regulation at an essential level of cellular functionality, the proteome. Time–of–day dependent regulation of protein stabilities probably adds to the precise fine tuning and robustness of circadian rhythms and thus might display essential targets for chronotherapy.

2 Material

2.1 Formulations of Buffer and Media

In the following, buffer and media formulations used for bacterial culture, molecular biological techniques and cell culture are listed.

Buffer and Media	**Formulation**
LB–agar plates	10-15 g agar in 1 LB-medium
LB–medium	10 g NaCl, 10 g Bactotrypton, 5 g Yeastextract, ad 1 l aq. dest., autoclaved
DNA loading buffer, 6x	30 % glycerol, 1 mM EDTA, 0.25 % bromphenol blue
TAE buffer, 50x	2 M Tris-Base, 50 mM EDTA, 1 M 100 % acetic acid, pH 8.5
Lysis–Buffer (for genomic DNA extraction)	10 mM Tris-Cl (pH 8.0), 0.02 M EDTA (pH 8.0), 0.5 % w/v SDS, *add before usage* 0.2 mg/ ml Proteinase K and 25 μg/ ml DNase-free RNase, sterile filtered
PCR Lysis Buffer	DirectPCR-Cell Lysis Reagent, 10 % Proteinase K
RIPA–Buffer	1x PBS, 1 % Igepal CA-630, 0.5 % sodium deoxicholat, 0.1 % SDS, *add before usage* 1 % Protease Inhibitor Cocktail
PBS, 10x	1.37 M NaCl, 27 mM KCl, 100 mM Na_2HPO_4, 20 mM NaH_2PO_4 in aq. dest., pH 7.2; for 1x PBS 100 ml 10x PBS ad 1 l aq. dest., autoclaved
BCA–Solution	50:1 solution A:B; solution A (2 % w/v Na_2CO_3 x H_2O, 1 % w/v BCA-Na_2, 0.95 % w/v $NaHCO_3$, 0.4 % w/v NaOH, 0.16 % w/v Na_2-Tartrat, ad 1 l aq. dest., pH 11.25), solution B (4 % w/v $CuSO_4$ x $5H_2O$)
TBS, 10x	1.37 M NaCl, 100 mM Tris-Base, pH 7.3
TBST, 1x	100 ml 10x TBS, 0.05 % Tween 20, ad 1 l aq. dest.
HEPES–Buffer	50 mM HEPES, 140 mM NaCl, 1.5 mM Na_2HPO_4, aq. dest., pH 7, sterile filtered
Freezing Medium	90 % FBS, 10 % DMSO
Protamine–sulfate	8 mg/ ml in aq. dest., sterile filtered
PEG–Solution, 5x	34 % w/v PEG600, 1.6 M NaCl ad aq. dest, sterile filtered
FACS–Buffer, 1x	1x PBS, 0.1 % NaN_3, 0.5 % FBS

Buffer and Media	Formulation (*continued*)
PFA–Solution, 4 %	10 g PFA in 225 ml H_2O heated to 65 °C, for dissolving 2M NaOH added drop–wise, ad 250 ml 10x PBS, pH 7.4, sterile filtered, aliquots frozen to −20 °C
Measurement–Medium	phenol red-free DMEM, 10 % FBS, 1 % Penicillin/Streptomycin, depending on application 27.7–250 μM D-luciferin
Dexamethasone	*stock* 1 mM in EtOH, *usage* 1 μM
D–Luciferin	*stock* 25 mM in H_2O
Passive Lysis Buffer, 5x	diluted to 1x buffer in aq. dest.

2.2 Antibiotics

Antibiotic		Source	Order No.
Ampicillin	*stock* 100 mg/ml in EtOH, *usage* 1:1000	Carl Roth	HP62.1
Blasticidin	*stock* 10 mg/ ml in aq. dest., *usage* 1:1,000	Life Technologies	R21001
Kanamycin	*stock* 50 mg/ ml in aq. dest., *usage* 1:500	Carl Roth	T832.1
Penicillin/ Streptomycin	10,000 U/ ml, *usage* 1:100	Life Technologies	15140122
Puromycin	*stock* 10 mg/ml in EtOH, *usage* 1:1,000	Sigma-Aldrich	P9620

2.3 Antibodies

Antibody	Dilution	Source
Primary Antibody		
mouse anti-βactin	1:200,000	Sigma-Aldrich, A-5441
rabbit anti-mCRY1 3818, serum, 3^{rd} bleeding	1:100 of pre–cleared serum*	self-generated within laboratory, done at Eurogentec
Secondary Antibody		
goat anti-mouse IgG-HRP	1:1,000	SantaCruz Biotechnology, sc-2005
donkey anti-rabbit IgG-HRP	1:1,000	SantaCruz Biotechnology, sc-2305

*Rabbit serum was pre–cleared applying the Melon Gel IgG Spin Purification Kit according to the manual.

2.4 Enzymes

2.4.1 Restriction Endonucleases

Enzyme		Source	Order No.
*BamH*I-HF	20,000 U/ ml	NEB	R3136
*Bsr*GI	10,000 U/ ml	NEB	R0575
*Bsr*I	10,000 U/ ml	NEB	R0527
*Dpn*I	20,000 U/ ml	NEB	R0176
*Eco*RV-HF	20,000 U/ ml	NEB	R3195
*EcoR*I	10,000 U/ ml	NEB	R0101
*Nco*I	10,000 U/ ml	NEB	R0193
*Sac*I	20,000 U/ ml	NEB	R0156
*Spe*I-HF	20,000 U/ ml	NEB	R3133
*Xho*I	20,000 U/ ml	NEB	R0146

2.4.2 Further Enzymes

Enzyme		Source	Order No.
Calf Intestinal Alkaline Phosphatase	10,000 U/ ml	NEB	M0290
M-MLV Reverse Transcriptase	200 U/ μl	Life Technologies	28025021
Phusion High-Fidelity DNA-Polymerase	2,000 U/ ml	Biozym Scientific	F-530L
RevertAid H Minus Reverse Transcriptase	200 U/ μl	Thermo Scientific	EP0451
RiboLock RNase Inhibitor	40 U/ μl	Thermo Scientific	EO0384

2.5 Oligonucleotides

For primer design Primer3 Input online software was used. Primer pairs were controlled for hairpin formation or self-annealing with the online tool OligoCalc. Melting temperatures (T_M) were calculated using the online tool OligoCalc (T_M of TOPO cloning primers was calculated without the non–coding CACC recognition sequence). Primers were synthesized at Eurofins MWG Operon. FW–forward primer, RV–reverse primer, ampli.–amplification.

2.5.1 Amplification Primers

Primer name	**Sequence (5′→3′)**	**T$_M$**
outer FW *nested PCR*	AACCACTACCTGAGCACCCAGT	64 °C
outer RV *nested PCR*	GCCAGAGGCCACTTGTGTAG	63 °C
nested FW *nested PCR*	GGGATCACTCTCGGCATGGACGA	68 °C
nested RV *nested PCR*	TAATACGACTCACTATAGGGAGAGG	64 °C
attB FW 5' *spike–in ampli.*	GGGGACAAGTTTGTACAAAAAAGCAGGC	69 °C
attB RV 3' *spike–in ampli.*	GGGGACCACTTTGTACAAGAAAGCTGG	69 °C
FW-ORF-GPS5.1 *Touch–Down PCR*	AACCACTACCTGAGCACCCAGT	64 °C
RV-ORF-GPS5.1 *Touch–Down PCR*	GCCAGAGGCCACTTGTGTAG	63 °C
mCry1 mutagenesis FW	GGCCAAATGGGCAGAAGGCCGGACA*G**A**C*TT CCCGTGGATTGACGCC	87 °C
mCry1 mutagenesis RV	GGCGTCAATCCACGGGAA*G**T**C*TGTCCGGCC TTCTGCCCATTTGGCC	87 °C
hAP4M1-TOPO-FW	**CACC**ATGATTTCCCAATTCTTCATTCTG	58 °C
hAP4M1-TOPO-RV	GATCCGAATGACATAGGCGTC	61 °C
hARFGAP3-TOPO-FW	**CACC**ATGGGGGACCCCAGCAAG	60 °C
hARFGAP3-TOPO-RV	AGAACCGTAGCGATCCTGAATT	60 °C
hARG1-TOPO-FW	**CACC**ATGAGCGCCAAGTCCAGAAC	60 °C
hARG1-TOPO-RV	CTTAGGTGGGTTAAGGTAGTCAA	60 °C
hFADS1-TOPO-FW	**CACC**ATGGGAACGCGCGCTGC	59 °C
hFADS1-TOPO-RV	TTGGTGAAGATAGGCATCTAGC	60 °C
hMR1-TOPO-FW	**CACC**ATGGGGGAACTGATGGCGTT	60 °C
hMR1-TOPO-RV	TCGATCTGGTGTTGGAAGGTAG	62 °C
hNDE1-TOPO-FW	**CACC**ATGGAGGACTCCGGAAAGAC	60 °C
hNDE1-TOPO-RV	GCAGGAGCTGGACGACCT	60 °C
hNPM1-TOPO-FW	**CACC**ATGGAAGATTCGATGGACATGG	61 °C
hNPM1-TOPO-RV	AAGAGACTTCCTCCACTGCC	60 °C
hNUTF2-TOPO-FW	**CACC**ATGGGAGACAAGCCAATTTGG	60 °C
hNUTF2-TOPO-RV	GCCAAAGTTGTGCAGGGCG	61 °C
hPIH1D1-TOPO-FW	**CACC**ATGGCGAACCCGAAGCTG	58 °C
hPIH1D1-TOPO-RV	AGAAGGCACCGGCAGAAGC	61 °C
hPLK1-TOPO-FW	**CACC**ATGAGTGCTGCAGTGACTGC	60 °C
hPLK1-TOPO-RV	GGAGGCCTTGAGACGGTTG	61 °C
hRANBP3-TOPO-FW	**CACC**ATGGCGGACCTGGCGAAC	60 °C
hRANBP3-TOPO-RV	TGTGCTCCCGGTCGTCTG	60 °C
hSEC61A-TOPO-FW	**CACC**ATGGGCATCAAATTTTTAGAAGTTA	57 °C
hSEC61A-TOPO-RV	GTAGAAGAGTATCCTTTTTGTAAAATG	60 °C
hTHOC6-TOPO-FW	**CACC**ATGGAGCGAGCTGTGCCG	60 °C
hTHOC6-TOPO-RV	GAAGGACAGGGAGAAGGCTC	62 °C

Primer name	**Sequence (5′→3′)** *(continued)*	**T$_M$**
Oligo d(T)	TTTTTTTTTTTTTTTTTT	–
Random Pentadecamers	NNNNNNNNNNNNNNN	–

2.5.2 Sequencing Primers

Primer name	**Description**	**Sequence (5′→3′)**
mCry1-450-FW	within mCry1 coding sequence	TCAGACTCTCGTCAGCAAG
M13-FW	in pENTR/D backbone, sequencing of 5' end of TOPO insert	GTAAAACGACGGCCAGTG
M13-RV	in pENTR/D backbone, sequencing of 3' end of TOPO insert	CAGGAAACAGCTATGAC
FW-MSCV-GPS	in pLenti6-GPS-*Dest* backbone, sequencing of 5' end of insert	GGGATCACTCTCGGCATGGACGA
RV-ORF-GPS5.1	in pLenti6-GPS-*Dest* backbone, sequencing of 3' end of insert	GATCAGTTATCTAGATCCGGTGGA

2.5.3 Quantitative Real–Time Primers

qRT–PCR primers were designed to amplify products between 80–120 bp at a T$_M$ of ≈ 60 °C. Efficiency of self-designed primer pairs was tested prior to analysis. Commercial qRT–PCR primers were purchased from Qiagen.

Primer name	**Sequence (5′→3′)/ Order No.**
self-designed primers	
AP4M1-qPCR–FW	GGGTCGATGAAGTCTCGTTT
AP4M1-qPCR–RV	TACCGCATCACAGTCAGCTC
ARFGAP3-qPCR–FW	TTGCCTCTCACGTTTCTCCT
ARFGAP3-qPCR–RV	GTTTCCACAGGCCTTGATGT
ARG1-qPCR–FW	TCCAAGGTCTGTGGGAAAAG
ARG1-qPCR–RV	ATTGCCAAACTGTGGTCTCC
FADS1-qPCR–FW	CTGCTGTACCTGCTGCACAT
FADS1-qPCR–RV	AGAGGAGGAAGGGCAAAAAG
MR1-qPCR–FW	CTGGCTGGAGTTGGTGTTCT
MR1-qPCR–RV	ATCGATCTGGTGTTGGAAGG
NDE1-qPCR–FW	GCAGCACTCTGAAGGCTACC
NDE1-qPCR–RV	GCTTGCTCCAGCTCTCTGAT
NPM1-qPCR–FW	TGGAGGAAGATGCAGAGTCA
NPM1-qPCR–RV	TGGAACCTTGCTACCACCTC
NUTF2-qPCR–FW	ACCATCAGCCCACTCCAGAT
NUTF2-qPCR–RV	ATGATGGGGTCTTCATCCGC
PIH1D1-qPCR–FW	CCAGACCAGAATCGACACAA

Primer name	Sequence (5′→3′)/ Order No. *(continued)*
PIH1D1-qPCR–RV	ATAGAGGGGGAGTGGCAGAT
pLenti6-backbone-FW	GGTGATAATTCTGCAGTCGAC
pLenti6-backbone-RV	CGACTCACTATAGGGAGAGG
PLK1-qPCR–FW	GGGCAACCTTTTCCTGAATGA
PLK1-qPCR–RV	TCCCACACAGGGTCTTCTTC
RANBP3-qPCR–FW	GGGCAGAACTTGAGGGACAGA
RANBP3-qPCR–RV	AAATAGTTCGTTGCGGTTGGCG
SEC61A-qPCR–FW	GAGGATCCTGTCCATGTCGT
SEC61A-qPCR–RV	CTTTGGCTGAGGAACCAGAC
THOC6-qPCR–FW	CTTGCAGCGGCTCCATATGAC
THOC6-qPCR–RV	AGAGGACAAGCTGAAGATGGCAA
TTC9C-qPCR–FW	CTACCGGGAAGGGAAGTACC
TTC9C-qPCR–FW	CTACCGGGAAGGGAAGTACC
TTC9C-qPCR–RV	TAGGTAACGGAGAGGGCAGA
TTC9C-qPCR–RV	TAGGTAACGGAGAGGGCAGA
hGapdh-qPCR-FW	TGCACCACCAACTGCTTAGC
hGapdh-qPCR-RV	ACAGTCTTCTGGGTGGCAGTG
commercial primers	
hBmal1	QT00011844
hClock	QT00054481
hCry1	QT00025067
hCry2	QT00094920
hCsnk2β	QT00012446
hDbp	QT00055755
hDec1	QT00026642
hDec2	QT00032697
hNfil3	QT00013944
hNono	QT01677963
hNpas2	QT00032480
hNr1d1	QT00000413
hPer1	QT00069265
hPer2	QT00011207
hRorα	QT00072380
hRorβ	QT00082026
hRorγ	QT00097888
hWdr5	QT00055524
mBmal1	QT00101647
mClock	QT00197547
mCry1	QT00117012
mPer2	QT00198366

2.6 Coding Sequences

Coding sequences are available with and without stop codon. All coding sequences were verified by analytical restriction digest and DNA sequencing. Coding sequences marked with * were cloned within this study and marked with ‡ were previously cloned within the laboratory.

Gene	Entrez GeneID	Source
h*Ap4m1*	9179	*
h*Arfgap3*	26286	*
h*Arg1*	383	*
h*Bhlhe40*	8553	kind gift from [103]
h*Bhlhe41*	79365	kind gift from [103]
h*Fads1*	3992	*
h*Mr1*	3140	*
h*Nde1*	54820	*
h*Npas2*	4862	kind gift from [103]
h*Npm1*	4869	*
h*Nr1d1*	9572	kind gift from [103]
h*Nutf2*	10204	*
h*Pih1d1*	55011	*
h*Plk1*	5347	*
h*Ranbp3*	8498	*
h*Ror*α	6095	kind gift from [103]
h*Sec61A2*	55176	*
h*Thoc6*	79228	*
m*Bmal1*	11865	‡
m*Clock*	12753	‡
m*Cry1*	12952	‡
m*Cry1mut*	modified from m*Cry1*, single point mutation at bp 1006 resulting in amino acid change G336D	* according to [104]
m*Per2*β*TrCP*	modified from m*Per2*, amino acid changes S477A and G479A	‡ [105]
m*Per2*	18627	‡
m*Per2mut7*	modified from m*Per2*, amino acid changes in S659G, S662A, S665A, S668A, S670A, S671A and T672A	‡ [73]
m*PPP1CA*	19045	kind gift from [103]
d1/ d4Egfp	destabilized EGFP variants	kind gift from [106]
DsRed	red fluorescent protein (available only with stop codon)	kind gift from [106]
Egfp	green red fluorescent protein	kind gift from [106]
lacZ	gene of β-galactosidase	Life Technologies

2.7 Vector Backbones

The following vector sequences were verified by analytical restriction digest and partial DNA sequencing. A vector map of the fluorescent reporter pLenti6-GPS-*Dest* is attached in the appendix A.14. Vector maps of basic backbones are described in detail at the suppliers websites. Modified plasmid backbones have been generated within the laboratory if not otherwise stated. euk.–eukaryotic; prok.–prokaryotic; resis.–resistance.

Vector backbone	**Description**	**Prok./ Euk. Resis.**	**Reference/ Source**
pABpuro–BluF	luciferase reporter of *Bmal*1 promoter, based on lentiviral pWPI backbone, 2^{nd} generation	Amp/ Puro	addgene
pDest26	Gateway® destination vector, used within co–transactivation assays	Amp, Chl/ Neo	Life Technologies
pENTR/D–TOPO	Gateway® entry vector	Kana/ –	Life Technologies
pGL3–E6S	firefly luciferase reporter of 6x cis-regulatory E–box motifs, used within co–transactivation assays	Amp/ –	modified from pGL3, Promega by [73]
phRL–SV40	encodes renilla luciferase, used within co–transactivation assays	Amp/ –	Promega
pLenti6–*Dest*-Luc	Gateway® destination vector, based on lentiviral backbone of 2^{nd} generation, fuses C–terminal luciferase to the insert	Amp/ Bla	modified from pLenti6-V5
pLenti6–GPS–*Dest*	bicistronic Gateway® destination vector, based on lentiviral backbone of 3^{rd} generation, fuses N–terminal EGFP, encodes DsRED, used for GPS analysis	Amp/ Bla	modified from pLenti6-V5 by [106]
pLenti6–V5	Gateway® destination vector, based on lentiviral backbone of 2^{nd} generation, used for overexpression	Amp/ Bla	Life Technologies
PM2	lentiviral envelope plasmid 3^{rd} generation	Amp/ –	kind gift from [107]
pMD2G	lentiviral envelope plasmid 2^{nd} generation	Amp/ –	addgene
psPAX	lentiviral packaging plasmid 2^{nd} generation	Amp/ –	addgene
REV	lentiviral packaging plasmid 3^{rd} generation	Amp/ –	kind gift from [107]
TAT	lentiviral packaging plasmid 3^{rd} generation	Amp/ –	kind gift from [107]
VSVG	lentiviral envelope plasmid 3^{rd} generation	Amp/ –	kind gift from [107]

2.8 Cell Lines

Cell line		**Reference/ Source**
HEK293	human embryonic kidney cell line	ATTC No. CRL-11268
HEK293T	human embryonic kidney cell line expressing the SV40 large T-antigen	ATTC No. CRL-11268
U–2 OS	human osteosarcoma cell line	kind gift of AG Hagemeier, Charité Berlin
U–2 OS-GPS-hORFeomeV5.1	U–2 OS cells stably expressing the GPS-hORFeome library V5.1 [107]	provided by the Elledge group, Harvard Medical School, Boston, MA; USA

2.9 Competent Bacterial Strains

Bacterial Strain	**Source**
E.coli DB3.1	Life Technologies
E.coli DH5α	Life Technologies
E.coli DH10β	Life Technologies
E.coli TOP10	Life Technologies
E.coli XL1 blue	Agilent Technologies

2.10 Material and Consumables

Consumable	**Source**	**Order No.**
2-mercaptoethanol	Carl Roth	4227
30 μm cell strainer	Partec	QT00055524
96–well plate, white	Thermo Fisher Scientific	136101
Agarose	Serva	11404
Exon Array 4x 180 k	Agilent	
Blotting Paper	VWR	732-4216
Cell Counting Chamber	Carl Roth	T729.1
Cycloheximide, ready made solution	Sigma-Aldrich	C4859
Dexamethasone	Sigma-Aldrich	D4902
Diamont Seal^{TM} for 96–well plates	Thermo Fisher Scientific	AB-0812
DirectPCR-Cell Lysis Reagent	PeqLab	31-301-C
D-Luciferin	P.J.K	102112

Consumable (*continued*)	**Source**	**Order No.**
DMEM, High Glucose	PAA	E15-810
DMEM, Phenol red-free, High Glucose, HEPES-buffered	Life Technologies	21063029
DMSO	AppliChem	A3672
DNA Ladder, 1 kb	NEB	N3232
DNA Ladder, 100 bp	NEB	N3231
Dual Luciferase Reporter Assay	Promega	E1960
Ethidium bromide 1 % solution	Roth	2218.1
FBS	Life Technologies	10270-106
Grease	Dow Corning	
HEPES	Life Technologies	15630058
Lipofectamine 2000 reagent	Life Technologies	11668019
Luciferase Assay reagent	Promega	E1483
MagicMark^{TM} XP Western Protein Standard	Life Technologies	LC5603
Melon Gel IgG Spin Purification Kit	Thermo Fisher Scientific	45206
Nitrocellulose Membrane	Thermo Fisher Scientific	77010
Nunc Cryo Mr. Frosty Freezing Container	Thermo Fisher Scientific	5100-0001
NuPAGE® 4–12 % Bis–Tris Protein Gels	Life Technologies	NP0335BOX
NuPAGE® LDS Sample Buffer, 4x	Life Technologies	NP0007
NuPAGE® MES SDS Running Buffer, 20x	Life Technologies	NP0002
NuPAGE® Transfer Buffer, 20x	Life Technologies	NP00061
Opti–MEM	Life Technologies	31985047
Orange Loading Dye, 6x	Thermo Scientific	R0631
Paraformaldehyde	Sigma-Aldrich	P6148
Passive Lysis Buffer, 5x	Promega	E1941
PEG6000	Sigma-Aldrich	81253
Protamine-sulfate	Sigma-Aldrich	P4020
Protease Inhibitor Cocktail	Sigma-Aldrich	P8340
Proteinase K	Roth	7528.1
PureLink® DNase Set	Life Technologies	12185010
Slim Milk	Becton Dickinson	232100
SuperSignal^{TM} West Pico Chemiluminescent Substrate ECL solution	Thermo Fisher Scientific	34080
Trypan blue	Sigma-Aldrich	T8154
Trypsin/EDTA	PAA Laboratories GmbH	L11-004

2.10.1 Consumable Kits

Kits were used according to the manufacturer instructions if not otherwise.

Consumable kit	**Source**	**order No.**
CalPhos^{TM} Mammalian Transfection Kit	Clontech Laboratories	631312
Fast-Link DNA Ligation Kit	Biozym	133625
Gateway® LR-Clonase® II Enzyme mix	Life Technologies	11791020
GeneJET Gel Extraction Kit	Thermo Fisher Scientific	K0691
HotStarTaq *Plus* Master Mix Kit	Qiagen	203645
Maxima^{TM} SYBR Green qPCR Master Mix	Thermo Fisher Scientific	K0223
MSB® Spin PCRapace	Invitek	1020220300
PureLink® HiPure Plasmid Filter Midiprep Kit	Life Technologies	K210015
PureLink® Quick Plasmid Minikit	Life Technologies	K210011
PureLink® RNA Mini Kit	Life Technologies	12183018A
QuickChange™ Site–directed Mutagenesis Kit	Agilent Technologies	200518

2.11 Technical Equipment

Device		**Company**
ALPS 50^{TM}	heater for sealing of 96–well plates	Thermo Fisher Scientific
Boxes, light-tight with single photomultiplier tubes	35–mm dish luminometer, temperature entrainable	Hamamatsu Photonics
ChemoCam Imager 3.2	chemiluminescence and UV-light detection, including ChemoStar software	INTAS
FACS Canto II	flow cytometry device	Becton Dickinson
JumoImago 500	control unit for temperature entrainment (setup was built by the Technische Werkstätten Charité, Berlin)	Jumo
LumiCycle	35–mm dish luminometer 37 °C	Actimetrics
NanoDrop 2000c	spectrophotometers	Thermo Fisher Scientific
Nunc Galaxy CO_2 incubator	cell culture incubator, used for cyclic temperature settings	Eppendorf
Orion II	plate luminometer, including Simplicity software	Berthold Detection Systems
qRT–PCR cycler CFX96	real–time detection system	Bio-Rad
Tecan Reader Infinite® 200 PRO	filter-based microplate reader, including software iControl	Tecan
Thermo cycler	uno cycler used for PCR	VWR
TopCount	plate luminometer at 37 °C	PerkinElmer

2.12 Software

Software		Company/ Source
CFX Manager	software for qRT–PCR cycler	Bio-Rad Laboratories
ChronoStar 2.0	in–house developed software for data evaluation of circadian bioluminescent online monitoring	Stephan Lorenzen, Bert Maier
FACS Diva	software to FACS Canto II	Becton Dickinson
FCS Express 4 Plus Research Edition	flow cytometry data analysis	DeNovo Software
Galaxy Commander	program controlling temperature cycles of cell incubator	RS Biotech, Eppendorf
GENtle	viewing and editing of DNA and amino acid sequences	open source
GO-Elite software	Gene Ontology enrichment analysis tool	open source, from [108]
ImageJ	image editing and analysis	open source
KyPlot 5.0	plot program	KyensLab Inc.
LaTex	text processing	open source
LumiCycle	LabView vi, control of LumiCycle luminometer	Actimetrics
Office 2010		Microsoft
Origin 7	graphing and data analysis software	OriginLab
photoNgraph	LabView vi, control of luminometer boxes	AutoMessTec
R	software for data analysis and statistical computing	open source

2.13 Databases, Websites and Online Tools

Tool	URL	last use
GO browser QuickGO	http://www.ebi.ac.uk/QuickGO/	05.11.2013
hORFeome V5.1	http://horfdb.dfci.harvard.edu/hv5/	14.11.2013
OligoCalc	http://www.basic.northwestern.edu/biotools/OligoCalc.html	14.11.2013
Primer3 V.0.4.0	http://bioinfo.ut.ee/primer3-0.4.0/	14.11.2013
UniHI	http://193.136.227.168/UniHI/pages/unihiSearch.jsf	01.11.2013

2.14 Company Register

Company	
Actimetrics	Wilmette, IL; USA
addgene	Cambridge, MA; USA

Company (*continued*)	
Agilent Technologies	Santa Clara; CA; USA
AppliChem GmbH	Darmstadt; Germany
Becton Dickinson	Franklin Lakes, NJ; USA
Berthold Detection Systems	Pforzheim; Germany
Bio-Rad Laboratories	Hercules, California; USA
Biozym Scientific GmbH	Hessisch Oldendorf; Germany
Clontech Laboratories	Mountain View, CA; USA
Dow Corning	Wiesbaden; Germany
Eppendorf	Enfield, CT; USA
Eurofins MWG Operon	Ebersberg, Germany
Eurogentec	Seraing; Belgium
Hamamatsu Photonics	Herrsching am Ammersee; Germany
INTAS	Göttingen; Germany
Jumo	Fulda; Germany
Life Technologies	Carlsbad, CA; USA
New England BioLabs (NEB)	Ipswich, MA; USA
P.J.K	Kleinblittersdorf; Germany
PAA Laboratories GmbH	Cölbe; Germany
Partec	Münster; Germany
Peqlab Biotechnologie GmbH	Erlangen; Germany
PerkinElmer	Rodgau; Germany
Promega	Madison, WI; USA
Qiagen	Venlo; Netherlands
Roche	Mannheim; Germany
SantaCruz Biotechnology	Dallas, TX; USA
Serva	Heidelberg; Germany
Source BioScience	Berlin; Germany
Tecan	Männedorf; Switzerland
Thermo Fisher Scientific	Waltham, MA; USA
VWR	Darmstadt; Germany

2.14.1 Suppliers of Chemicals

General laboratory chemicals were obtained from the following companies if not otherwise stated.

Company	
Carl Roth	Karlsruhe; Germany
Merck KGaA	Darmstadt; Germany
Sigma-Aldrich	St. Louis, MO; USA

3 Methods

Commercially available kits were used according to the manufactures instructions if not otherwise stated or described in designated sections.

3.1 Molecular Biology Methods

3.1.1 Polymerase Chain Reaction

Polymerase chain reactions (PCRs) according to Mullis et al.[109] were performed to amplify desired DNA fragments specifically. As template either 100–200 ng genomic DNA, 30–50 ng plasmid DNA or cDNA corresponding to 2 μg reverse transcribed RNA were used per reaction. Annealing temperature was chosen according to the lowest melting temperature (T_M) of used primers if not indicated otherwise (see oligonucleotide list in 2.5.1).

Table 3.1: PCR setup

(a) General PCR composition

component	volume (μl)
template	*variable*
High Fidelity Buffer [5x]	10
dNTP's [10 mM]	1
forward primer [10 μM]	2
reverse primer [10 μM]	2
H_2O	ad 50
Phusion Polymerase [2 U/μl]	0.3

(b) General PCR conditions

step	temperature	time
Initial denaturation	98 °C	1 min
30–35 cycles	98 °C	10 sec
	T_M	30 sec
	72 °C	30 sec per kb
Final extension	72 °C	10 min

Nested PCR

In order to specifically amplify single ORFs that are inserted in the cellular genome within the fluorescent reporter construct (see subsection 4.1.3) the following nested PCR protocol was applied. For a pre–amplification of each ORF per cellular genome, a first primer pair was designed specific to ORF adjacent regions in the reporter backbone. To guarantee a pre–amplification of all 16,000 library ORFs 2 μg genomic DNA were used in ten single 1^{st} PCR reactions (corresponds to $\sim$ 300,000 cells that represent a 19–fold overrepresentation of each ORF). PCR amplificates were pooled and purified for the 2^{nd} nested PCR. Here, a second primer pair with sequence homology to ORF flanking sides within the 1^{st} PCR product was designed. PCR products were purified with the MSB® Spin PCRapace kit.

1^{st} PCR

200 ng genomic DNA used as template per reaction,
10 single 1^{st} PCR reactions,
outer primer pair
⇓

Table 3.2: 1^{st} PCR conditions

step	temperature	time
Initial denaturation	98 °C	1 min
5 cycles	98 °C	10 sec
	61 °C	30 sec
	72 °C	5 min
Final extension	72 °C	10 min

mix all 10 PCR reactions
⇓
PCR purification of 50 μl PCR reaction mix,
elution in 115 μl H_2O
⇓

2^{nd} PCR

34.7 μl of purified 1^{st} PCR used as template,
nested primer pair
⇓

Table 3.3: 2^{nd} (nested) PCR conditions

step	temperature	time
Initial denaturation	98 °C	1 min
20 cycles	98 °C	10 sec
	59 °C	30 sec
	72 °C	4.5 min
Final extension	72 °C	10 min

⇓
PCR purification of nested PCR reaction,
elution in 100 μl H_2O

Touch–Down PCR for Sequencing of hORFeome V5.1 Single Clones

For sequencing of hORFeome library V5.1 constructs of single cell clones, open reading frames (ORFs) needed to be pre–amplified from genomic DNA of lysed cells (see subsections 3.1.10 and 3.1.12). This amplification was done in a Touch–Down PCR. PCR reaction composition and conditions are listed in Tab. 3.4. Prior to DNA sequencing, PCR products were purified with MSB® Spin PCRapace kit and eluted in 20 μl H_2O.

Table 3.4: Touch–Down PCR setup

(a) Touch–Down PCR composition

component	volume (μl)
lysed cell mix	2
forward primer [10 μM]	1.5
reverse primer [10 μM]	1.5
H_2O	5
HotStar Taq Master Mix	10

(b) Touch–Down PCR conditions

step	temperature	time
Initial denaturation	95 °C	5 min
	97 °C	30 sec
3 cycles	64 °C	30 sec
	72 °C	5.5 min
	94 °C	30 sec
3 cycles	61 °C	30 sec
	72 °C	5.5 min
	94 °C	30 sec
3 cycles	58 °C	30 sec
	72 °C	5.5 min
	94 °C	30 sec
35 cycles	57 °C	30 sec
	72 °C	5.5 min

3.1.2 Site–Directed Mutagenesis

Site–directed mutagenesis was done to introduce a substitution of amino acid 336 from Glycine (G) to Aspartic Acid (D) within the m*Cry1* coding sequence at bp 1006 (according to[104]). To this end, the QuickChange™ Site–directed Mutagenesis Kit was applied. As template pENTR/D_m*Cry1* plasmid DNA was used for the mutagenesis PCR (see composition and condition in Tab. 3.5). Primers were designed with sequence homology to the base surrounding the site to be mutated (G to A, see bold marked nucleotides in the primer sequences in the oligonucleotide list in subsection 2.5.1). The point mutation was verified by DNA sequencing. Plasmids whit the correct mutation were preparatively digested with *Sac*I and *Eco*RV to cut out a DNA fragment surrounding the mutated site. This fragment was cloned into a sequence verified pENTR/D_m*Cry1* plasmid, where the corresponding wild type fragment was removed by the same restriction enzymes.

Table 3.5: Mutagenesis PCR

(a) Mutagenesis PCR composition

component	volume (μl)
pENTR/D_m*Cry1* [50 ng/μl]	1
Buffer [10x]	5
dNTS's [10 mM]	1
forward primer [2.5 μM]	5
reverse primer [2.5 μM]	5
H_2O	ad 50
PfuTurbo Polymerase [2.5 U/μl]	1

(b) Mutagenesis PCR conditions

step	temperature	time
Initial denaturation	95 °C	2 min
	95 °C	30 sec
16 cycles	55 °C	1 min
	68 °C	4 min 20 sec

3.1.3 TOPO® Cloning

Human coding sequences were shuttled into pENTR/D plasmids via TOPO® cloning. Desired coding sequences were amplified from Oligo(dT) reverse transcribed human RNA from U–2 OS cells (see subsections 3.1.1 and 3.1.13). For TOPO® cloning 1 μl of purified PCR product (preparative gel electrophoresis, see subsection 3.1.7) was used. Reaction mix was setup according to the manufacturer's protocol, however, TOPO® cloning reaction was incubated for 30 min at room temperature (RT). Transformation into competent *E.coli* TOP10 strains was done as described in section 3.1.8.

3.1.4 Gateway® Cloning

The Gateway® reaction is an *in vitro* recombination that allows to shuttle coding sequences from entry vectors (pENTR/D_TOPO) into destination plasmids (see vector list in section 2.7). After Gateway® reaction (see Tab. 3.6 for conditions) the mixture was transformed into competent *E.coli* DH10β bacterial strains as described in section 3.1.8.

Table 3.6: Gateway® reaction and composition

component	volume (μl)
pENTR/D_*ORF* [30 ng/ μl]	1
destination vector [30 ng/ μl]	1
LR Clonase® II	0.5
incubate 60 min at 25 °C	
Proteinase K [2 μg/ μl]	0.25
incubate 10 min at 37 °C	

3.1.5 DNA Restriction Digestion

DNA restriction digestion was done analytically or preparatively. In general, all reactions were carried out with 1 mg/ ml bovine serum albumin (BSA), 1x restriction enzyme buffer and filled up with HPLC–grade water to the indicated volume. Restriction digestion fragments were size-separated using gel electrophoresis, visualized using ethidium bromide and if required extracted from the agarose gel (see subsection 3.1.7).

For an analytical digest, approximately 500 ng plasmid DNA and 1–3 U restriction enzyme were incubated in a final volume of 10 μl for at least one hour at the indicated restriction enzyme activity temperature (mostly 37 °C).

For a preparative restriction digest of plasmid DNA whose DNA fragments were used in further steps, about 5–10 μg plasmid DNA were digested with 5–10 U restriction enzyme in a final volume of 50–100 μl for at least three hours at the appropriate restriction enzyme activity temperature.

3.1.6 DNA Ligation

Ligations of DNA inserts into linearized plasmid backbones were done with the Fast-Link DNA Ligation Kit. Beforehand, linearized plasmids were dephosphorylated to prevent recircularization. To this end, 1 μl Calf Intestinal Alkaline Phosphatase (CIP) was directly added to the preparative restriction digestion mixture and incubated for 30 min at 37 °C. This step was repeated once.

DNA inserts and plasmid backbones were purified via preparative agarose gel electrophoresis (see subsection 3.1.7). After ligation reaction (see composition and condition in Tab. 3.7) the reaction mixture was transformed into competent bacterial cells (see subsection 3.1.8).

Table 3.7: Composition and condition of DNA ligation

component	volume (μl)
Fast-Link Buffer [10x]	1
ATP [10 mM]	1
plasmid backbone [5–20 ng/ μl]	1
DNA insert [*final* molar ratio 2:1 of insert:vector]	*variable*
H_2O	ad 4.5
Fast-Link ligase [2 U/ μl]	0.5
incubate 10 min at RT	
incubate 15 min at 70 °C for heat-inactivation	

3.1.7 Agarose Gel Electrophoresis

DNA gel electrophoresis was done to size-separate linearized DNA fragments. Depending on the expected size of DNA fragments, agarose gels with 0.8–2 % w/v agarose in 1x TAE and 0.05 μl/ ml 1 % ethidium bromide (EtBr) solution were used. DNA samples were mixed with 6x DNA loading buffer or 6x Orange Loading Dye (used if DNA fragments around 300–400 bp were expected), loaded on the agarose gel and run at 80–120 V for 20–40 min. Simultaneously, DNA ladders (100 bp or 1 kb DNA Ladder) were run. Intercalated EtBr into DNA fragments allowed a visualization by UV-light.

For preparative gel electrophoresis, DNA fragments of expected sizes were cut out and extracted from agarose by the use of a DNA gel extraction kit (GeneJET). DNA fragments were eluted in 10–30 μl HPLC–grade water.

3.1.8 Transformation of Plasmid DNA into Competent Bacterial Strains

For transformation, either 200–500 ng of plasmid DNA or whole reaction mixtures of DNA ligation, Gateway® or TOPO® cloning were gently mixed with 100 μl thawed competent bacterial cells and incubated for 30 min on ice. After heat-shock at 42 °C for 90 sec, 500 μl LB-medium without antibiotics was added. Transformation mixture was incubated for 30–40 min at 37 °C on a horizontal shaker (750 rpm).

100–500 μl of transformed bacterial cells were plated on LB–agar plates containing the appropriate antibiotic according to the transformed plasmid DNA. Plates were incubated over night at 37 °C.

3.1.9 DNA Preparation from *E.coli*

For plasmid DNA isolation a single bacterial colony was inoculated in LB–medium containing the appropriate antibiotic and incubated at 37 °C over night in a horizontal shaker (250 rpm). Plasmid DNA was prepared applying Mini- or Maxiprep kits (PureLink®) with 5 or 50 ml over night culture, respectively. DNA was eluted in 50 or 100 μl HPLC–grade water and spectrophotometrically analyzed and quantified. DNA sequence was verified via analytic restriction digests and/ or DNA sequencing.

3.1.10 DNA Sequencing

DNA sequencing was done by the service provider Source BioScience. Used sequencing primers are listed in Tab. 2.5.2.

For sequencing of plasmid DNA, 5 μl of 50–150 ng/ μl concentrated DNA and 5 μl of 3.2 pmol/ μl sequencing primer were used.

Sequencing of Single Cell Clones of hORFeome V5.1 Library

For sequencing of single V5.1 library clones, ORF information was pre–amplified from genomic DNA of lysed cells (see details within subsection 3.1.12 and 3.1.1). 6 μl of column purified PCR products were used per sequencing reaction.

3.1.11 DNA Cryoconservation

For longtime storage of plasmid DNA, bacterial overnight culture carrying the desired plasmid were mixed 1:1 with DMSO, vortexed and chilled for at least one hour on ice. DNA cryos were stored at −80 °C.

3.1.12 Isolation of Genomic DNA

High pure genomic DNA from fresh or frozen pellets of cell culture cells was isolated via phenol–chloroform extraction as depicted in Tab. 3.8. The amount and purity of genomic DNA was spectrophotometrically determined.

Table 3.8: Isolation of genomic DNA

Cell lysis	
add 600 μl Lysis–Buffer to up to 4 Mio. pelletized cells	
incubate at 55 °C, horizontal agitation (750 rpm), 1 h	
Extraction of genomic DNA	
add sodium chloride to final 0.3 M	
add 600 μl phenol/chloroform/isoamyl alcohol (1:1:1)	repeat 3x
invert 5x, *centrifugate* 10 min at 10,000 rcf and RT	
transfer aqueous upper phase to new tube	
Removal of phenol impurities	
add 600 μl chloroform/isoamyl alcohol (24:1)	repeat 2x
mix gently, *centrifugate* 5 min at 10,000 rcf and RT	
transfer aqueous upper phase to new tube	
Precipitation of genomic DNA	
add 1.5 ml ice-cold EtOH	
mix gently, *precipitate* for 1 h at −80 °C	
centrifugate 30 min at 13,000 rcf and 4 °C	
wash with ice-cold 70 % EtOH	
centrifugate 30 min at 13,000 rcf and 4 °C	
discard supernatant	
incubate 20 min at 37 °C to dry DNA pellet	
add 50 μl H_2O and resolve DNA pellet over night at RT	

Isolation of Genomic DNA form hORFeome V5.1 Library Single Cell Clones

For sequencing of single library clones genomic DNA was isolated from confluent single cell clones grown in 96–well plates. Cells were treated with 25 μl Trypsin/ EDTA and incubated at 37 °C until cells were detached. For cell lysis, 5 μl of detached cells were mixed with 20 μl of PCR Lysis Buffer and incubated for 15 min at 55 °C, followed by 85 °C for 45 min. Lysed samples were stored at −20 °C.

3.1.13 Isolation of RNA and Reverse Transcription

For total RNA isolation PureLink® RNA Mini Kit was used. RNA of approximately 2 Mio. cells (living or frozen either as cell pellet or in cell culture dish) was isolated. Optional DNase treatment was performed during RNA purification as recommended in the RNA Mini Kit protocol. RNA amount and purity were spectrophotometrically determined.

cDNA Synthesis

RNA was either reversely transcribed with Random Pentadecamer or Oligo(dT) primers. For gene expression analysis, cDNA was synthesized with Random Pentadecamers applying the M-MLV reverse transcriptase (see Tab. 3.9 a). Reverse transcription from total RNA with Oligo(dT) primer applying RevertAid H Minus reverse transcriptase was done to obtain cDNA used as template for amplification of coding sequences (TOPO® cloning, see 3.1.3). This reaction was not heat-inactivated to prevent the cleavage of long cDNA strands (see Tab. 3.9 b).

Table 3.9: cDNA synthesis protocol and composition

(a) ... with Random Pentadecamers

component	volume (μl)
total RNA [2 μg]	*variable*
Reaction Buffer [5x]	10
dNTS's [10 mM]	1
RNase Inhibitor [40 U/ μl]	0.5
Pentadecamers [500 pmol]	5
H_2O	ad 50
M-MLV RT [200 U/μl]	1
incubate 10 min 25 °C 50 min 37 °C 15 min 70 °C	

(b) ... with Oligo(dT) primers

component	volume (μl)
total RNA [2 μg]	*variable*
Oligo(dT) [100 pmol]	1
H_2O	ad 12.5
incubate at 65 °C for 5 min, chill on ice *add*	
Reaction Buffer [5x]	4
RNase Inhibitor [40 U/ μl]	0.5
dNTS's [10 mM]	2
RevertAid H Minus RT [200 U/μl]	1
incubate at 42 °C for 60 min, chill on ice	

3.1.14 Quantitative Real–Time PCR

Quantitative real–time PCR (qRT–PCR) was applied for endogenous gene expression analysis or quantification of PCR products. cDNA samples for gene expression analysis and PCR products of amplified ORFs from the hORFeome library (see sub–subsection in 3.1.1) were diluted 1:10 in HPLC–grade water. qRT–PCR reactions were setup and performed as depicted in Tab. 3.10. For the detection of primer dimers, a melting step was included after the qRT–PCR reaction. qRT–PCR primers are listed in Tab. 2.5.3.

Table 3.10: Compositions and conditions of qRT–PCRs

(a) Composition with self-designed qPCR primers

component	volume (μl)
SYBR Green Master Mix	10
self-designed qPCR primer FW [10 μM]	0.3
self-signed qPCR primer RV [10 μM]	0.3
H_2O	1.4
diluted sample	8

(b) Composition with commercial qPCR primers

component	volume (μl)
SYBR Green Master Mix	10
commercial primers	2
diluted sample	8

(c) qRT–PCR conditions

step	temperature	time
	50 °C	2 min
Initial denaturation	95 °C	10 min
40 cycles	95 °C	15 sec
	60 °C	60 sec
Melting curve	95 °C	10 sec
	65 °C to 95 °C	increment of 1 °C/ sec
	95 °C	0.5 sec

3.2 Protein Biochemistry Methods

3.2.1 Whole Cell Protein Extraction

Proteins were extracted from adherent cell culture, grown to confluence in 6–well plates. Cells were washed with ice-cold 1x PBS and lysed in 600 μl RIPA–Buffer for 30 min at 4 °C with slight horizontal agitation. Cells were scraped from the culture dish surface and the genomic DNA was sheared through a 20 gauge cannula. Lysed cells mix was transferred to a reaction tube and cell debris were pelleted for 30 min at 13,000 rcf and 4 °C. Supernatant, containing the protein lysate, was transferred to a new reaction tube.

3.2.2 Determination of Protein Concentration

Total protein concentration was determined by the colorimetric bicinchoninic acid (BCA) assay. To this end, 5 μl protein lysate and 200 μl BCA–Solution were incubated for 30 min at 37 °C. After incubation, colorimetric change was measured at

560 nm. Protein concentrations were calculated based on a calibration curve of serial BSA dilutions in RIPA–Buffer (0–10 μg/ μl) that were measured simultaneously.

3.2.3 SDS Polyacrylamid Gel Electrophoresis

Proteins were size-separated by SDS–Polyacrylamid Gel Electrophoresis (SDS–PAGE). To this end, 30–50 μg protein lysate were heat-denatured for 10 min at 95 °C in 4x NuPAGE® Sample Buffer. Denatured protein samples and 5 μl protein ladder were loaded on 4–12 % Bis–Tris polyacrylamid gels and run for 45 min at 200 V.

3.2.4 Western Blotting and Immunodetection

For the immunodetection, size-separated proteins were transferred from the polyacrylamid gel to a nitrocellulose membrane in 90 min at 90 V. After the wet–transfer, nitrocellulose membranes were blocked for 1 h in 5 % w/v skim milk in 1x TBST at RT and constant slight horizontal agitation. Membranes were washed 3x for 10 min with 1x TBST and incubated in the primary antibody dilution at 4 °C over night with slight horizontal agitation (see subsection 2.3 for antibodies and dilutions). After 3x 10 min washing with 1x TBST, diluted HPRT–conjugated secondary antibody was incubated for 1 h at RT and slight horizontal agitation. Again, membranes were washed 3x for 10 min with 1x TBST. In a chemiluminescent reaction HPRT–conjugates were detected by addition of their substrate and monitored with the ChemoCam Imager. For quantification of detected protein bands, ImageJ software was applied.

3.3 Cell Biology Methods

3.3.1 Cell Cultivation and Cryoconservation

Cell lines were kept in Dulbecco´s modified Eagle Medium (DMEM) supplemented with 10 % fetal bovine serum (FBS), 1 % Penicillin/Streptomycin and 25 mM HEPES–Buffer at 37 °C and 5 % CO_2 in a humidified cell culture incubator. Cells were grown to 80–90 % confluence and splitted 1:6 through 1:8. To that, cells were washed in 1x PBS and detached by the incubation with Trypsin/EDTA solution at

37 °C for 5 min. After resuspension in pre–warmed supplemented DMEM, cells were diluted to new culture flasks.

For cell cryoconservation, 1–5 million cells were pelleted at 300 rcf for 7 min at 4 °C. Cell were resuspended in 1 ml Freezing Medium and incubated for one day in a freezing container at −80 °C that cools the cells approximately 1 °C/ per h.

For long-term storage, cells were transferred to liquid nitrogen tanks. For the re-cultivation, cryoconserved cells were thawed for 2 min in a 37 °C water bath. Pre–warmed supplemented DMEM was added drop–wise until cells were completely thawed. Cells were carefully platted in appropriate cell culture dishes. One day after thawing, medium was changed.

Reporter Library Cell Cultivation

To maintain the complexity of 16,000 ORFs, with one ORF per cellular genome, at least 40 million hORFeome V5.1 library cells were cultured in parallel and combined and mixed with each splitting step.

3.3.2 Lentiviral Packaging and Transduction (stable overexpression)

For a stable overexpression of desired ORFs or reporter plasmids into immortalized cells lentiviral transductions were performed.

Lentiviral Packaging

The production of lentiviral particles was done in HEK293T cells in 75 square centimeter cell culture flasks to harvest about 20 ml lentiviral supernatant (the protocol was also adapted to larger or smaller cell culture flasks). To this end, lentiviral packaging and expression plasmids were transiently transfected (CalPhos Transfection Kit) into HEK293T cells of 70–80 % confluency (seeded one day before). Transfection compositions for the production of 2^{nd} and 3^{rd} lentiviral particles are depicted in Tab. 3.11 (see section 2.7 for plasmids of the 2^{nd} and 3^{rd} generation of lentiviral particles). While vortexing, DNA/ Ca^{2+} transfection mixtures were added drop–wise to 600 μl 2x HEPES–buffered Saline (HBS). After 20 min incubation at RT, the whole transfection solution was added to HEK293T cells (supplemented with fresh medium prior to this step). After over night incubation at cell culture conditions medium was replaced. Cell supernatants containing the viruses were collected 48 and 72 hrs post–transfection. To remove cell debris, supernatant was spun at 4,100 rcf for 15 min and filtered through a 45 μm filter. Viral supernatants were stored in aliquots at $-80\,^{\circ}$C.

Table 3.11: Lentiviral packaging

(a) ... of 2^{nd} Generation lentiviral particles

plasmid	volume (μl)
lentiviral expression plasmid [*final* 8.4 μg]	*variable*
psPAX [1 μg/ μl]	6
pMD2G [1 μg/ μl]	3.6
H_2O	ad 600
Ca^{2+}	74.4

(b) ... of 3^{rd} Generation lentiviral particles

plasmid	volume (μl)
lentiviral expression plasmid [*final* 9.6 μg]	*variable*
TAT [1 μg/ μl]	2.4
VSVG [1 μg/ μl]	2.4
REV [1 μg/ μl]	2.4
PM2 [1 μg/ μl]	2.4
H_2O	ad 600
Ca^{2+}	74.4

Precipitation of Lentiviral Supernatant

In order to concentrate lentiviral containing supernatant, one part 5x PEG–Solution was mixed with 4 parts lentiviral supernatant and incubated at $4\,^{\circ}$C for one day. The mixture was inverted from time to time. Lentiviral/ PEG particles were precipitated at 4,100 rcf for 30 min and $4\,^{\circ}$C. Supernatant was aspirated and residual supernatant was carefully removed after an additional centrifugation step at 4,100 rcf for 5 min, at $4\,^{\circ}$C. Lentiviral/ PEG600 pellets were resuspended with 1x PBS in $\frac{1}{100}$ of the starting volume.

Lentiviral Transduction

Lentiviral transduction was performed in 6–well plates transducing 2 x 10^5 cells per well. For infection, lentiviral supernatant (depending on desired MOI and application) were added in a final volume of 2 ml culture medium including final 8 μg/ ml protamine–sulfate. Lentiviral supernatant was incubated on cells for at least 24 hrs at cell culture conditions. Depending on the transduced expression plasmid (see section 2.7) cells were selected with either 10 μg/ml puromycin for at least 24 hrs or 10 μg/ml blasticidin for at least 72 hrs. Prior to blasticidin selection, cells were detached (as described in subsection 3.3.1) and re-seeded in supplemented DMEM including blasticidin. A selection control (cells without overexpression of the selection marker) was always cultured in parallel.

Titration of Lentiviral Particles

To determine the transduction unit per ml (TU/ ml) of lentiviral supernatant, serial two–fold dilutions (depending on the viral supernatant ranging from 0–1:512) were transduced as described above. After selection, surviving cells were counted. TU/ ml was calculated according to equation 3.1 and a mean was build with all dilutions where a linear relationship between amount of transduced viral supernatant and percent of surviving cells was found (e.g. with increasing volume of viral supernatant used for transduction, more cells survived). Volume of lentiviral supernatant needed for a certain Multiplicity of infection (MOI) was calculated according to equation 3.2.

$$\text{TU/ ml} = \frac{\text{No. of transduced cells} \times \text{\% surviving cells}}{\text{used volume of lentiviral supernatant [ml]}} \tag{3.1}$$

$$\text{volume}_{\,lentiviral\ supernatant} = \frac{\text{MOI} \times \text{no. of cells to be transduced}}{\text{mean TU/ ml}} \tag{3.2}$$

Transduction of hORFeome Reporter Library Cells

To the 'clamp' the molecular circadian clock, CRY1 and and as a control CRY1mut were stably overexpressed in hORFeome reporter library cell. To this end, lentiviral particles with expression plasmids of CRY1 and CRY1mut were concentrated and TU/ ml were determined. To maintain the complexity of the library, 60 x 10^6 library cells were transduced with an MOI of one in 30 single 175 square centimeter cell culture flasks each with a final volume of 20 ml.

3.3.3 Bioluminescent Live–Cell Monitoring

Bioluminescent online measurements were done with synchronized or asynchronous cell populations carrying a luciferase reporter. Measurements were performed either with sealed white 96–well plates (Nunc) containing 2 x 10^4 cells per well in 150 μl Measurement–Medium or with 35–mm dishes (Nunc) containing 2 x 10^5 cells in 2 ml Measurement–Medium, sealed with grease. 96–well plates were measured in the TopCount plate luminometer, 35–mm dishes were monitored either in the LumiCycle luminometer or in light-tight boxes with single photomultiplier tubes. For online measurements of cells luminometers were used at 37 °C.

Cell Synchronization with Dexamethasone

To characterize circadian clock dynamics after perturbation (see subsection 4.2.1), synchronized U–2 OS cells carrying a stably integrated clock gene promoter fused to luciferase (*Bmal1*–promoter luc) were monitored over 4–8 days. To this end, cells were synchronized with 10 μM dexamethasone for 30 min, washed with 1x PBS and cultured in Measurement–Medium containing 250 μM luciferin.

Cell Synchronization by Temperature Entrainment

Temperature entrainment was applied to cells carrying the ORF::luc reporter constructs (see subsection 4.3.4). Cells were synchronized with 10 μM dexamethasone for 30 min, washed with 1x PBS and cultured in Measurement–Medium containing 27.7 μM luciferin. After dexamethasone synchronization, cells were kept for 18 hrs at 37 °C before 4x temperature cycles of 12 hrs 33 °C/ 12 hrs 37 °C were applied. Temperature cycles were generated using a heating plate at the bottom of individual boxes controlled by a JumoImago 500 control unit. After release into constant 37 °C, bioluminescent measurements were continued for at least 3 days.

In addition, for experiments without bioluminescent online monitoring, temperature cycles were applied to a cell culture incubator without humidification, controlled by Galaxy Commander software.

Data Evaluation of Bioluminescent Online Monitoring

For data evaluation of bioluminescent live–cell monitoring of synchronized cells, ChronoStar 2.0 software was used. Raw data were trend–eliminated by dividing them with their 24 hr running average. Based on a fitted sine curve circadian parameters of period, phase, amplitude and damping were calculated.

ChronoStar 2.0 software was not feasible for data evaluation of ORF::luc online monitoring after temperature entrainment due to low amplitude rhythms in highly trend afflicted raw data (ChronoStar software was programmed and trained for high amplitude rhythms). Thus, trend–elimination as described above was calculated in Microsoft Excel 2000 for the raw data after release into constant 37 °C.

Data evaluation of not synchronized cells (see experiment in Fig. 4.12 B, ORF::luc cells after overexpression) was done with Microsoft Excel 2000. To this end, raw data of the second day of live–cell monitoring were averaged.

3.3.4 Bioluminescence Measurements in Harvested Cells

In addition to online bioluminescent recordings of living cells, luciferase activity was determined in cell lysates (e.g. for protein decay experiments, see subsection 3.3.5). Cells were lysed with 1x Passive Lysis Buffer. To this end, per well of cells grown to confluency in a 12–well plate, 100 μl 1x Passive Lysis Buffer was added and incubated for 5 min at RT with slight horizontal agitation. For a complete lysis, 12–well plates with cells in lysis buffer were incubated for at least one hour at $-80\,°$C. Luciferase measurements were done in a plate luminometer (Orion II). Of thawed and vortexed lysed cells, 5 μl were pre–layed per well of a white 96–well plate (Nunc). Within the luminometer, 25 μl Luciferase Assay Reagent I were added per well, incubated for 2 sec and luciferase counts were measured over 10 sec.

3.3.5 Protein Decay Measurements after Cycloheximide Administration

In order to determine the stability of proteins integrated into cells as reporter fusions, protein translation was blocked and protein decay measured via EGFP fluorescence (see experiments in Fig. 4.2 B) or luciferase activity (see experiments in Fig. 4.12 D and appendix fig:proof) in harvested cell samples. To this end, protein translation was blocked by the direct administration of 0.2 mg/ ml cycloheximide (CHX) to the growth media of confluent cells. As a control, solvent (DMSO) treated cells were handled in parallel. Depending on the experiment cells were harvested in regular intervals over 6–8 hours. For a fluorescence based readout cell were fixed with paraformaldehyde (see subsection 3.3.8) and EGFP decay was monitored by flow cytometry analysis (see subsection 3.3.7). To readout protein decay via luciferase activity, cells were lysed in 1x Passive Lysis Buffer and luciferase counts were measured in a luminometer (described in subsection 3.3.4).

Fluorescence intensities or luciferase counts are given relative to intensities/ counts of solvent treated cells over the time course. Exponential decay functions were fitted to calculate protein half–lifes according to the formulas 3.3 and 3.4.

$$N(t) = N_0 e^{-\lambda t} \tag{3.3}$$

$$t_{1/2} = \frac{ln(2)}{\lambda} \tag{3.4}$$

N(t) = quantity at time t
N_0 = initial quantity
λ = decay constant

3.3.6 Co-Transactivation Assay

To test the inhibitory action of CRY1 and of a CRY1 mutant on the transactivation activity of CLOCK/BMAL1 on E–boxes, co–transactivation assays were performed (see experiments in Fig. 4.6 A). To this end, 1 x 10^5 HEK293 cells were seeded into 24–well plates (final volume 500 μl). Twenty-four hours later, growth medium was changed and cells were transfected using Lipofectamine 2000 reagent with the following constructs: 50 ng of pGL3–E6S construct (contains a 6x cis–regulatory E–box motif followed by firefly luciferase) were co–transfected with 300 ng of pDest26–Clock and of pDest26–Bmal1, as well as with 80 or 250 ng of pDest26–Cry1 or Cry1mut as indicated. In addition, as a transfection control, 50 ng of phRL–SV40 renilla luciferase plasmid were co–transfected. Each transfection reaction was supplemented with pDest26–lacZ to reach 1,200 ng DNA in total. DNA was diluted in final 150 μl Opti–MEM per reaction. In addition, per transfection 3 μl Lipofectamine transfection reagent was incubated for 5 min at RT in 147 μl Opti–MEM. Lipofectamine– and DNA– OptiMEM mixtures were mixed, incubated for 20 min at RT and added drop–wise to the cells. Forty–eight hours after transfection, cells were harvested in 100 μl 1x Passive Lysis Buffer as described in subsection 3.3.4. Luciferase activities were measured using the Dual Luciferase Reporter Assay System and the Orion II luminometer. Within the Orion II luminometer, 25 μl Luciferase assay reagent II was added to 5 μl cell lysate, incubated for 2 sec and measured over 10 sec to determine firefly activity. Then 25 μl Stop & Glow reagent was added, incubated for 2 sec and measured over 10 sec to determine renilla activity. For data evaluation firefly luciferase counts were normalized to renilla luciferase counts.

3.3.7 Flow Cytometry

Analytical flow cytometry was done to quantify single cell fluorescence of U–2 OS cells carrying the GPS reporter (for details see section 4.1). To this end, cells were grown to confluence. Four days after medium change, cells were trypsinized and pelleted at 300 rcf for 7 min and 4 °C. Cell pellets were resuspended in 50–200 μl 1x FACS–Buffer and chilled on ice. Alternatively, paraformaldehyde-fixed cells were analyzed (see subsection 3.3.8).

Flow cytometry analysis was done using a FACS Canto II with the software application FACSDiva. Settings for the analysis of U–2 OS cells carrying the fluorescent reporter are depicted in Tab. 3.12. Samples were excitated with an argon laser at 488 nm. DsRED fluorescence was readout via the PE filter (585 nm/ width 42 nm) and EGFP fluorescence with the FITC filter (530nm/ width 30 nm). Per sample at least 3,000 DsRED positive cells were analyzed.

Table 3.12: Flow cytometry settings

Filter	Voltage	Compensation
FCS	120	
SSC	450	
FITC	300	FITC – % PE 2.34
PE	270	PE – % FITC 14.76

Flow Cytometry Data Analysis

FCS files were exported from the FACSDiva software for data analysis within FCS Express 4 Plus software. Using different gates mean fluorescence intensities (MFIs) were extracted and exported to text files for ratio calculations within Microsoft Excel 2010. Furthermore, raw data of flow cytometry analysis were exported as text files for analysis and visualization with Microsoft Excel 2010 or KyPlot 5.0 (e.g. EGFP/DsRED ratio histograms). For the determination of the EGFP/DsRED ratio where the distribution peaks (as done in appendix Fig. A.3), a Lorentzian function was fitted and the maximum calculated (done within Orion 7).

Preparative Flow Cytometry

Preparative flow cytometry was done at the DRFZ Flow Cytometry Core Facility using the sorter FACSDiva. To prevent cell clumping, cells prepared as described for analytical flow cytometry, were strained through a 30 μm filter directly before sorting. In addition, a 100 μm nozzle was used for the preparative flow cytometry sort. Cells were preparatively sorted according to their EGFP/ DsRED ratio (as depicted in Fig. 4.4 A) into eight tubes where supplemented DMEM was pre–layed. *Note:* Due to technical reasons compensation as depicted in Tab. 3.12 was not possible for the EGFP/ DsRED ratio sort.

3.3.8 Cell Fixation

Cells were fixed with paraformaldehyde (PFA) for subsequent analytical flow cytometry analysis (e.g. time series or decay experiments). To this end, cells were detached as described in subsection 3.3.1 and resuspended in a defined volume of pre–warmed 1x PBS. For fixation the same volume of 4 % PFA–Solution was added. The cell–PFA–mixture was transferred to reaction tubes and incubated for 30 min at RT with constant vertical rotation. Fixed cells were pelleted for 3 min at 2,000 rcf. The supernatant was discarded and the fixed cells were washed with 1x FACS–Buffer. After centrifugation for 3 min at 2,000 rcf the supernatant was discarded and fixed cells were resuspended in 200–400 μl 1x FACS buffer and store dark at 4 °C for up to 4 weeks.

3.4 Bioinformatic Analysis

3.4.1 Microarray Data Analysis

Microarray Data Pre–Processing

Microarray data pre–processing was done in cooperation with Karsten Jürchott (Berlin-Brandenburg Center for Regenerative Therapies, Charité Berlin). Raw data were background subtracted, normalized within each, and over all four arrays applying the limma package[110] in the R software. Pre–processed microarray data of each color were graphically depicted in dot plots for visual inspection. Each dot represents the signal intensity measured for a probe set (ps). All four arrays show an overall correlation of Cy3 versus Cy5 intensities. Clouds on the upper left and lower right corner represent probe sets with an altered intensity of one color, indicating differential regulation of their corresponding ORF (appendix Fig. A.6 A).

Threshold Determination

A threshold determining detected ORFs was based on probe set intensities of genes that are not denoted within the hORFeome library V5.1. To this end, for each labeling color (Cy3 and Cy5) and fraction (pooled fractions 1–4) the intensity value was determined for which 90 % of all ORFs that are not in the hORFeome library are below. The mean of this cut-off value was 10.9 ± SD 0.26 for Cy3 labeled samples and 10.9 ± SD 0.232 for Cy5 labeled samples.

Microarray Data Processing

In addition to the microarray data processing described in the results subsection 4.3.2, the following considerations and calculations were done.

Probe set intensities per fraction were corrected according to the amount of cells per fraction for each overexpression condition (thus for each labeling color). This was done because library cells display a bell-shaped distribution over EGFP/DsRED ratios (see Fig. 4.4 A). Thus, the number of cells categorized in fractions one to four varies. As similar numbers of cells were sorted per fraction, the probability of single cells to be preparatively sorted is not similar between fractions. This was normalized by the multiplication of probe set intensities per fraction times the relative fraction content. The relative fraction content is the percent of cells categorized per fraction relative to the fraction with the least percent of cells.

The relative protein stability index (r.PSI), Δr.PSI, z–score of Δr.PSI as well as the euclidean distance (ED) and centered intensity were calculated for each probe set with log2 scaled pre–processed and fraction content corrected intensities as described in subsection 4.3.2 and according to the formulas in Fig. 4.10 A. Instead of absolute values, relative values were used for the calculation of PSI and Euclidean Distance to exclude high Δr.PSI and ED values that are only based on different overall intensities between both sub–libraries.

To exclude bimodal intensity distribution pattern, probe sets, for which the difference between the r.PSI and the fraction number with the highest intensity is greater than one, were removed before probe set condensation.

Calculation of Δr.PSIs for qRT–PCR derived Values

Intensity profiles of the microarray analysis were validated by qRT–PCR in samples with more fractions (not pooled samples of fractions 1–8, same samples as analyzed on microarray) and different spacing of fractions (independently sorted library samples; see experiments in Fig. 4.11). Cq–values were normalized to the amount of reporter backbone (as determined by a qRT–PCR primer pair detecting a sequence region of the GPS backbone) and corrected for fraction content (see above). For the calculation of r.PSIs according to the formula 4.10 $A_{(1)}$ 2^{dCq} values were used and following modifications were applied. First of all, 2^{dCq} were not additionally delogarithmized for calculation of r.PSI values. Secondly, to calculate r.PSI of qRT–PCR data from the not pooled samples, fraction numbers 0.5, 1, 1.5, 2, 2.5, 3, 3.5 and 4 were used fractions 1–8, respectively. For the independently sorted samples with different spacing of the fractions, parameters for r.PSI calculations of qRT–PCR derived values were adapted based on the size of the fraction. Thus, for fraction 1 to 5 following parameters were used 0.5, 1, 1.5, 2.75 and 3.25, respectively.

3.4.2 Gene Ontology Enrichment Analysis

Gene ontology (GO) enrichment analysis was done using the GO-Elite software[108]. Settings depicted in Tab. 3.13 were applied. Enriched GO terms were manually combined to major terms under consideration of their ancestor charts and child terms. To this end, the GO browser 'QuickGO' was used.

Table 3.13: GO-Elite Analysis Settings

Criteria	Setting
Input ID List	540 CAAPs
Denominator ID	6,059 detected gs
Algorithm	Fisher Exact Test
Primary ID system	EntrezGene
Ontology terms ranked by	z–score
Analyzed resources	all
z–score cutoff for initial filtering	1.96
Permuted p–value cutoff	0.05
Minimum no. of changed genes	3
Exclude terms with >	10000

3.4.3 Statistics

For statistical testing of two unpaired groups the nonparametric Mann–Whitney U test was performed. The Kolmogorov–Smirnov was performed to test if two sets of data were sampled from populations of different distribution. For a statistical comparison of two unpaired groups, the nonparametric Spearman Rank Correlation test was performed. All test were done with a significance level of 0.05. Calculations were done in Excel 2010, KyPlot 5.0 or R.

4 Results

The comparison of circadian transcriptome and proteome studies revealed a discrepancy between the amount of circadian transcripts and circadian abundant proteins[36–41,44,45,42,43]. This indicates further rhythmic control at post–transcriptional, translational and/ or post–translational levels. To elucidate the extent of post–translational mechanisms controlling rhythmic protein abundances, circadian protein degradation of the proteome, the so called circadian 'stabilome' needs to be investigated.

This study aimed to establish and perform a proteome–wide screen to analyze circadian protein stabilities. To this end, firstly a method to analyze protein stabilities, feasible for proteome–wide high–throughput approaches, is established (section 4.1). The second part describes how the method was implemented to identify proteins with potential circadian stabilities (section 4.2). In the last section, evaluation of the high–throughput screen as well as validation experiments of selected candidates are presented (section 4.3).

4.1 A Method to Screen Protein Stabilities

In order to screen for global circadian protein stabilities, a method introduced by the Elledge group[106], allowing to analyze protein stabilities in high–throughput was applied. The 'Global Protein Stability' (GPS) method is based on a bicistronic reporter plasmid. The plasmid encodes for a DsRED fluorescent protein and an EGFP fluorescent protein that can N–terminally fuse to any open reading frame (ORF). An internal ribosome entry site (IRES) allows for the simultaneous translation of both proteins from one mRNA transcript (see Fig. 4.1).

DsRED and EGFP are stable proteins with half–lifes of about eight and four days, respectively[111]. The red fluorescent protein serves as control for transcription rate and amount of reporter construct integrated in cells. The stability of the EGFP–fusion, however, is determined by the more unstable protein of the fusion. Thus, ratio

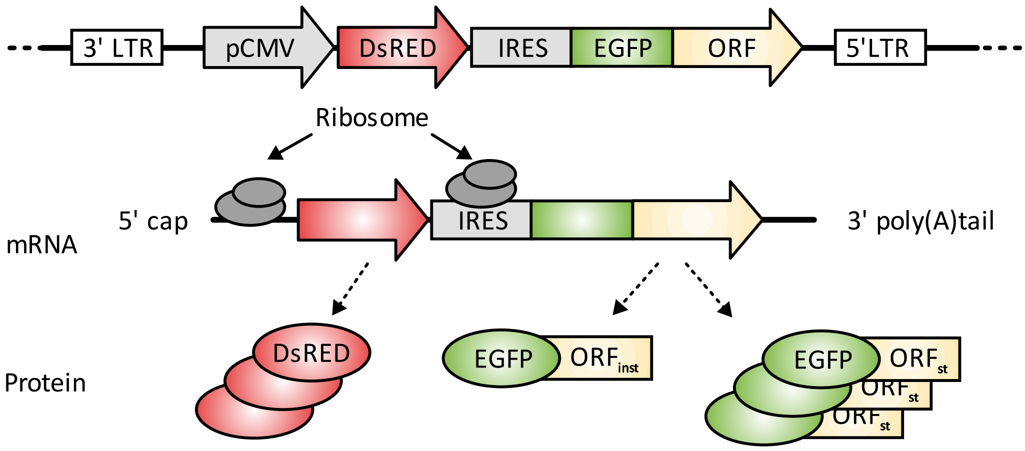


Figure 4.1: Bicistronic reporter to monitor protein stability. Schematic representation. The GPS reporter construct is transcribed under the control of a constitutive CMV promoter. Ribosomes binding simultaneously at the 5′ cap and IRES translate DsRED and EGFP–fusion proteins from one transcript. Stable EGFP–fusion proteins are more abundant than unstable fusion products. Thus, fluorescent intensities reflect the stability of the EGFP–fusion relative to the long-lived DsRED protein. Modified from [106]. DsRED–*Discosoma species* red fluorescent protein, EGFP–enhanced green fluorescent protein, GPS–Global Protein Stability, IRES–internal ribosome entry site, LTR–long terminal repeats, ORF–open reading frame, pCMV–cytomegalovirus promoter, st–stable, unst–unstable.

changes of green relative to red fluorescence reflect different abundances of EGFP–fusion proteins. Fused ORFs lack motifs of 5′ and 3′ untranslated regions (UTRs) that might regulate protein level post–transcriptionally. Hence, translation rate of different EGFP–fusions is alike, independent of the fused ORFs. Thus, different protein abundances most likely reflect diverse stabilities of fused proteins.

4.1.1 Characterization of the GPS Method

The GPS method to analyze protein stability is based on fluorescent reporter proteins that are encoded from exogenous DNA stably integrated in cells. The method was thoroughly characterized to investigate whether ratio changes of EGFP/DsRED are solely based on the abundance (stability) of fused ORFs.

As described above, both fluorescent proteins are translated from one mRNA transcript. Translation efficiencies starting from the 5′ cap and IRES are reported to result in a molar ratio of roughly 1:1 for both cistron proteins [112]. The maturation time of both fluorescent proteins is described to be similar [113,114]. Thus, median fluorescence intensities (MFIs) and their resulting ratio of EGFP/DsRED are supposed to be constant over time. To test this, fluorescent intensities were analyzed by flow cytometry over several days in human osteosarcoma cells (U–2 OS) carrying the GPS reporter. DsRED MFI slightly increased over seven days after medium change. In contrast, EGFP MFI displayed a strong decline reaching a plateau at the fourth day after medium change (see Fig. 4.2 A). The same trend was observed for different ORFs analyzed within the reporter (data not shown). Hence, a relatively stable ratio of EGFP/DsRED MFIs adjusts four days after medium change and was defined as the earliest time of measurement for following analysis applying this method.

Different EGFP and DsRED intensities are observed between single cells of a population expressing the same ORF within the GPS reporter. This might be due to the amount of reporter integrated per cellular genome and/ or the transcription rate depending on the genomic integration site. As expected, an increase of relative DsRED fluorescence is in accordance with increasing relative EGFP fluorescence as observed for different fused ORFs (see Fig. 4.2 B). Thus, stable ratios of EGFP/DsRED MFIs can be assumed over all DsRED intensities. To test this, EGFP–DsRED positive cells were clustered into five classes of increasing DsRED intensity, each containing 20 % of cells. Surprisingly, EGFP/DsRED ratios decreased with increasing DsRED intensity (see Fig. 4.2 C). This indicates an absence of a linear correlation between EGFP and DsRED fluorescence within a cell population. As this trend was observed for all analyzed ORFs within the GPS reporter, a general reporter rather than an ORF specific effect is assumed. EGFP/DsRED ratios of cell populations, however, cluster around different ratios, specific for each ORF (see Fig. 4.2 D). Thus, either the peak of the distribution or the quotient of EGFP/DsRED MFIs can be used as an approximation for the ORF specific ratio where most of the analyzed cells accumulate.

4.1.2 Application of the GPS Method to Analyze Protein Stability

In order to show that the bicistronic fluorescent reporter can be used to measure stabilities, proteins with different half–lifes were analyzed. To this end, the core clock component mPERIOD2 (PER2) and two PER2 mutants with described stabilities were cloned into the GPS reporter and stably introduced into U–2 OS cells. The PER2mut7 variant is mutated in stabilizing phosphorylation sites and thus less stable than wild type PER2 protein[73]. In contrast, mutation of the F–box protein binding region (PER2βTrCP) prevents its proteasomal degradation, thus stabilizes PER2[105].

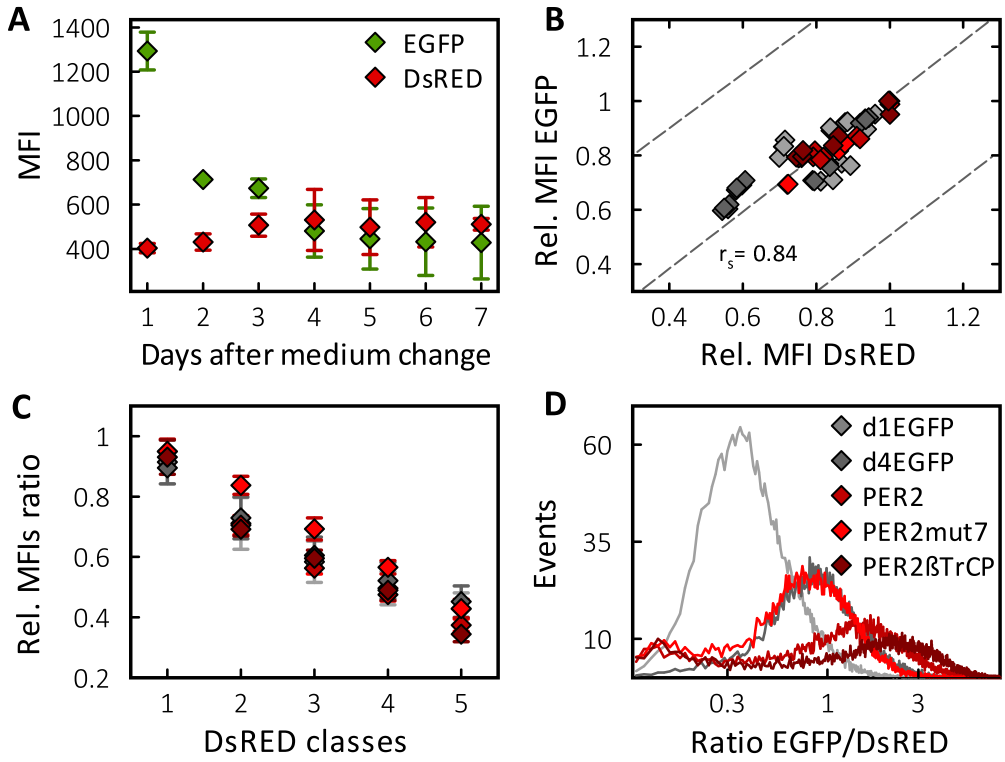


Figure 4.2: Characterization of the GPS method to screen protein stability. Flow cytometry analysis were preformed with U–2 OS cells stably expressing the GPS reporter. (A) EGFP/DsRED ratios change over time. Green and red MFIs of an EGFP variant (d4EGFP) within the reporter were measured over seven days after medium change. Mean values ± SD of three to four replicates of at least two independent measurements are given. *(B–D)* EGFP/DsRED ratios were analyzed in cells harboring either a destabilized EGFP variant (d1/d4EGFP) or the core clock protein PERIOD2 (PER2) or one of its destabilized (PER2mut7) or stabilized (PER2βTrCP) mutants. *Color code in (D).* (B) Relative EGFP versus relative DsRED MFIs correlate significantly ($P < 0.001$, Spearman Rank Correlation test, $n = 56$ per group). MFIs are given relative to the maximum intensity of each, EGFP and DsRED per analyzed ORF. (C) Cells were binned into five classes, each containing 20 % EGFP-DsRED positive cells. Increasing class numbers represent increasing DsRED fluorescence. MFI ratios are set relative to the maximum ratio per analyzed ORF. Plotted are mean ratios ± SD of at least four independent experiments. (D) Distribution plots of cells according to their EGFP/DsRED ratio. r_s–Spearman's Rank Correlation coefficient, rel.–relative.

As expected, MFI ratios of EGFP/DsRED were decreased for PER2mut7 and increased for PER2βTrCP compared to PER2 fusion proteins. Furthermore, two EGFP variants with known protein stabilities of 1– and 4–hrs (d1– and d4EGFP, respectively[115,116]) were analyzed. MFI ratios for d1EGFP were decreased compared to d4EGFP MFI ratios, as expected (see Fig. 4.3 A and 4.2 D; see appendix Fig. A.1 for representative flow cytometry raw data). Thus, stabilities of different proteins can be approximated with the GPS method.

To test if changing stabilities of a protein can be detected (as it would be the case for circadian protein stabilities), genetic perturbation experiments were performed. Overexpression (OX) of protein phosphatase 1 alpha (PPP1CA) was shown to interact and destabilize BMAL1 protein. Indeed, reduced MFI ratios of BMAL1 but not of its binding partner CLOCK were measured after OX of PPP1CA compared to an overexpression control (see Fig. 4.3 B, results were published in[117]).

In order to test whether the results of the GPS method correlate with values derived from 'classical' protein stability measurements, half–lifes of different proteins were determined and compared to their EGFP/DsRED ratios. To this end, protein *de novo* synthesis in cells was blocked by the administration of cycloheximide (CHX). Decay of MFI ratios of EGFP/DsRED were monitored over seven hours. Exponential decay functions were fitted to calculate protein half–lifes. A significant correlation of protein half–lifes to MFI ratios of EGFP/DsRED was observed (see Fig. 4.3 C; see appendix Fig. A.2 for an additional correlation plot of EGFP/DsRED ratios versus protein half–lifes obtained from C–terminally tagged luciferase fusion proteins). Together, these findings demonstrate that the GPS method can be applied to distinguish between different stabilities of proteins encoded by ORFs.

4.1.3 A hORF Library within the GPS reporter for a Proteome–Wide Analysis

To perform a proteome–wide analysis of protein stabilities, a hORF library was used. The hORFeome V5.1 consists of 15,483 ORFs covering about 12,794 non–redundant human genes. These include 1,502 genes with multiple splice variants and 814 polymorphic genes[118,119]. The library is based on PCR amplified cDNA clones from the Mammalian Gene Collection[120,121]. The hORFeome V5.1 library was Gateway® cloned into the GPS reporter construct and stably introduced into U–2 OS cells. Thereby, viral transductions were done with a multiplicity of infection (MOI) of approximately 0.05 to guarantee the insertion of only one ORF per cell*.

Flow cytometry analysis of library cells revealed a wide distribution of EGFP/ DsRED ratios (see Fig. 4.4 A). This represents various EGFP–ORF abundances indicating different protein stabilities.

Analysis of EGFP/DsRED ratios in cell populations expressing single ORFs within the reporter showed a good correlation to their protein half–lifes (see Fig. 4.3 C

*The reporter hORFeome cell library in U–2 OS cells was kindly provided by the Elledge group, Harvard Medical School, Boston, MA; USA.

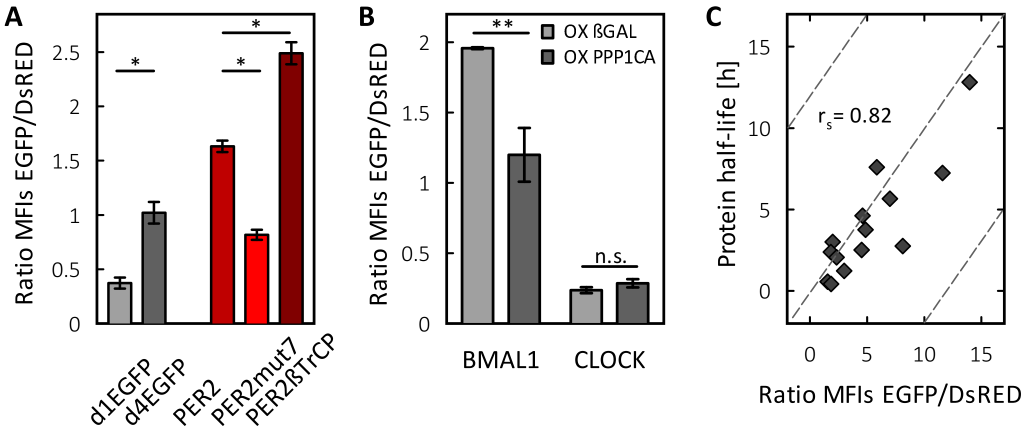


Figure 4.3: The GPS method analyzes protein stabilities. Depicted are MFI ratios of EGFP/DsRED analyzed in U–2 OS cells carrying the stably integrated fluorescent reporter. (A) Two destabilized EGFP variants with known protein half–lifes of 1– and 4–hrs as well as PER2 together with a destabilized (PER2mut7) and a stabilized (PER2βTrCP) mutant were analyzed. Mean values $\pm$ SD of at least four independent experiments are given. *P $<$ 0.05 (Mann–Whitney U test). (B) PPP1CA destabilizes BMAL1 protein. PPP1CA or a control (βGAL) were overexpressed in cells carrying either BMAL1 or CLOCK within the GPS reporter. Mean values $\pm$ SD of at least two independent experiments with each four to six replicates given. **P $<$ 0.01 (Mann–Whitney U test, data are published in [117]). (C) Correlation plot of EGFP/DsRED ratios (x–axis) versus protein half–lifes (y–axis). Decay of the MFI ratios after pharmacological block of protein synthesis relative to a solvent control was determined over seven hours. Protein half–lifes were calculated from fitted exponential decay functions (P $<$ 0.001, Spearman Rank Correlation test, n = 14 per group). βGAL–β-galactosidase, n.s.–not significant.

and appendix Fig. A.2). In order to investigate whether fluorescent intensities within the library largely reflect proteome half–lifes, a comparison to proteome stability data recently published by Schwanhäusser *et al.*[63] was done. To this end, protein half–lifes estimated in a SILAC (stable isotope labeling by amino acids in cell culture) mass spectrometry approach in NIH3T3 mouse fibroblasts were compared to protein stability indices (PSIs) of corresponding ORFs of the hORFeome library. Essentially, increasing PSI values reflect increased protein abundances and thus stabilities of the EGFP–ORF fusion (detailed explanation of the PSI is described in subsection 4.3.2). Among 1,873 proteins that were analyzed in both studies, no overall correlation of protein half–lifes to PSIs was found. Nonetheless, overall half–lifes obtained in the mass spectrometry approach are significantly different for proteins categorized according to their PSI values into a stable or an unstable group (two groups of proteins according to the 10 % smallest or largest PSIs; see Fig. 4.4 B). Thus, taking into consideration that both screens differ in their origin of proteins

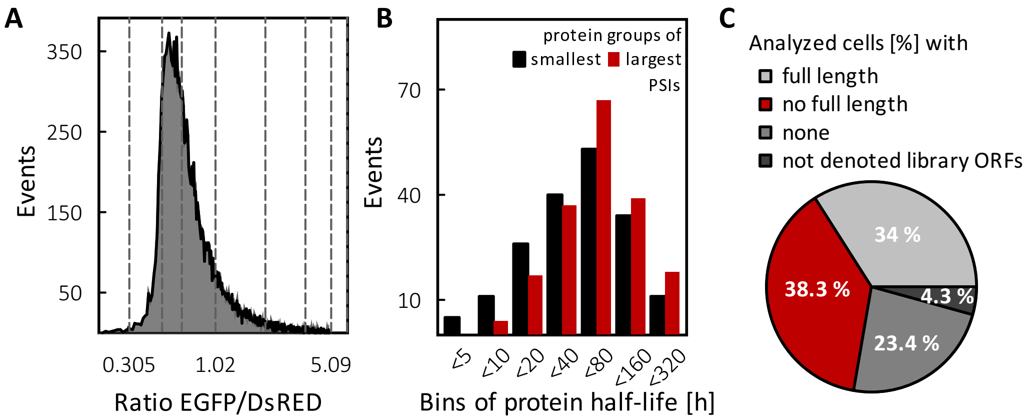


Figure 4.4: Characterization of the hORF library within the GPS reporter. (A) Distribution plot of EGFP/DsRED ratios of the hORFeome V5.1 library within the GPS reporter in U–2 OS cells. Dotted lines border eight fractions. (B) Half–lifes analyzed in [63] of proteins categorized into groups of the 10 % smallest (black) and largest (red) PSIs were binned (see x–axis). A significant difference was detected between both groups (*$P < 0.05$ Mann–Whitney U test and Kolmogorov–Smirnov test, $n = 187$ per group; PSI data are from the control library, see subsection 4.3.2). (C) Individual cell clones of the reporter library were sequenced. The pie chart shows percents of cells that were found as correct full length (light gray), fragments (red), none (medium gray) or not denoted (dark gray) ORFs within the hORFeome V5.1 library ($n_{total} = 47$). PSI–protein stability index.

(mouse versus human), study setup (endogenous proteins versus overexpressed ORFs) and measurement technology (mass spectrometry versus GPS method), hORFeome EGFP–ORF abundances can be assumed as a good estimation for protein stability.

Quality Assessment of the hORFeome Library

Quality assessments of earlier hORFeome library releases revealed a sequence mutation rate of 8–17 %[118,119]. Additionally, not denoted or empty clones were found as a result of unspecific PCR byproducts and failures of the Gateway® cloning, respectively. In order to study the quality and consistency of the reporter library used here, ORFs of isolated cell clones were sequenced. To this end, primers specific for the reporter backbone at either the 5′ or 3′ adjacent end of the inserted ORFs were designed. Of 47 sequenced individual clones, only one third were validated as full length ORFs according to the hORFeome V5.1 database†. More than one third of sequenced single clones were found to be a fragment, i.e. either without the correct 5′ end or an incomplete 3′ end. Nearly one fifth of analyzed single clones contained no ORF at all. Furthermore, about 4 % were not denoted to be present

†http://horfdb.dfci.harvard.edu/hv5/

within the hORFeome V5.1 (see Fig. 4.4 C). Thus, probably less than 50 % of the library represent full length ORFs – a result that needs to be taken into consideration when interpreting results derived from this library.

In summary, the GPS reporter method is feasible to detect different protein stabilities and can be applied for a proteome–wide analysis. Particularly with regard to screen for circadian changes in protein stability, the method seems to be appropriate. It uncouples protein abundance from protein specific synthesis allowing to detect changes of protein abundance that are solely based on post–translational events. Thus, protein specific transcriptional, post–transcriptional and translational processes that might be under circadian control resulting in circadian abundant proteins can be excluded.

4.2 High–Throughput Screen for Circadian Protein Stabilities

The following section describes the implementation of the GPS reporter method to screen for proteins with circadian stability in a proteome–wide high–throughput. To this end, a 'comparative one–time point approach' where the molecular circadian clock of library cells is 'arrested' was developed. To identify and quantify changes in EGFP–ORF abundances, 'clock arrested' and control cells were preparatively sorted into fractions according to their EGFP/DsRED ratios (see indicated fractions in Fig. 4.4 A). ORF coding sequences were specifically amplified from genomic DNA of fractionated cells for the subsequent quantitative analysis on microarrays (see flow chart in Fig. 4.5).

4.2.1 Comparative One–Time Point Approach

A protein sample taken at one circadian time point contains protein pools of different ages and circadian times. The steady state between protein synthesis and degradation is probably not reached. Consequently, the expected amplitude of putative circadian abundance rhythms is likely low. In a first attempt to screen for time–of–day dependent protein stabilities, synchronized cells were sampled around the clock. EGFP/DsRED ratios of control ORFs (d1– and d4EGFP) showed high non–circadian variability over the circadian cycle (see appendix A.3). Thus, slight changes of circadian protein abundances might not be detectable with such an approach.

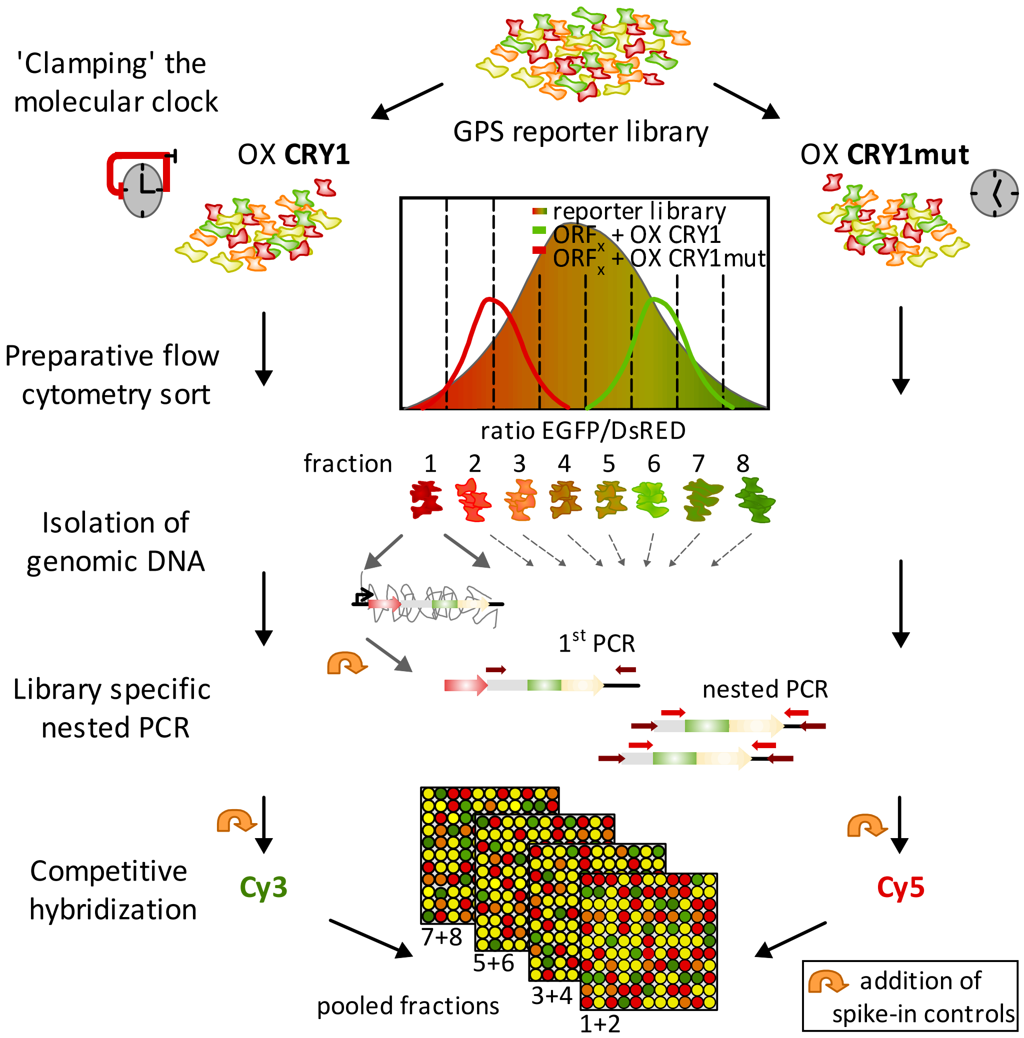


Figure 4.5: Proteome–wide high–throughput approach to analyze circadian protein stabilities. The flow chart depicts essential steps of the high–throughput approach to screen for proteome–wide circadian protein stabilities. The GPS reporter library covering about 16,000 hORFs consists of a stably integrated bicistronic reporter construct encoding DsRED and N–terminally fused EGFP–ORF proteins (see Fig. 4.1). The ratio of EGFP/DsRED describes the abundance of the EGFP fusion protein. Due to the design of the reporter, different protein stability is indicated by altered protein abundance that results in a changed EGFP/DsRED ratio (see example of ORF_x). To screen for time–of–day dependent changes in protein stability, a comparative one–time point approach was performed. Stable overexpression of CRY1 'clamps' the molecular clock at a specific phase compared to a the overexpression of a control (CRY1mut). To analyze altered protein stabilities, cells of both sub–libraries were preparatively sorted into eight fractions according to their EGFP/DsRED ratios. Genomic DNA of sorted cells was isolated and a library specific nested PCR amplifying ORF sequences was applied. To identify and quantify amplified ORFs, PCR products of two consecutive fractions were pooled, labeled with fluorescent dyes (Cy3 and Cy5) and competitively hybridized on exon microarrays. Spike–in controls were added before (as genomic DNA) and after nested PCR reaction (as PCR products) as indicated by the orange arrows. Cy3/5–cyanine dyes 3/5, GPS–Global Protein Stability, OX–overexpression, PCR–polymerase chain reaction.

Overexpression of CRY1 'Clamps' the Circadian Clock

In order to increase the difference of protein abundances, a comparative one–time point approach was developed. Library cells were 'clamped' at a specific phase of the circadian cycle. Thereby, the steady state between production and degradation can be reached and potential differences in protein abundances between two conditions ('clamped' versus unsynchronized clocks) are easier to be detected.

How can the molecular clock be 'clamped' at a specific circadian phase? Rhythmic abundances of most core clock proteins are crucial for maintaining cellular clock dynamics. Their protein peak represents a specific circadian time[80,122]. Thus, to 'arrest' the clock at a defined circadian phase, clock proteins of the negative and positive arm were dose–dependently overexpressed in U–2 OS cells. Circadian dynamics and expression of core clock and clock output genes were analyzed to identify 'clamped' clocks and the phase they were 'arrested' (see appendix A.4).

To this end, stable exogenous expression of CRY1 protein, a component of the negative arm, 'arrested' the molecular clock in a phase of low E–box driven gene expression. Indeed, CRY1 is described to interact with and reduce CLOCK/BMAL1 mediated transcriptional activity of E–box driven genes[123]. To control for clock unspecific effects of overexpression, a CRY1 mutant was cloned. CRY1mut, with the amino acid change G336D, was shown to lack the repressive potential towards CLOCK/BMAL1 activity[104,124]. In this line, in co–transactivation assays, dose–dependent repression of CLOCK/BMAL1 mediated activation of luciferase fused to E–box target sequences could be shown for CRY1, but not for its mutant (see Fig. 4.6 A).

OX of CRY1 'Clamps' the Circadian Clock and Stabilizes PER2 Protein

In order to show that a 'clamped' clock model can be used to detect altered protein abundances of proteins with circadian stability, PER2 protein, known to be circadian degraded[78,79], was analyzed in cells with an 'arrested' clock. Overexpression of CRY1, but not of its mutant, increased EGFP–PER2 and PER2–V5 tagged protein abundances (see Fig. 4.6 B, upper panel and C), as expected[125,73]. Thus, a 'clamped' clock in a comparative one–time point approach seems to be a suitable model to detect altered protein abundances of circadian stable proteins.

Thus, to 'clamp' the molecular circadian clock, CRY1 and – as a control – the Cry1 mutant were stably overexpressed in GPS reporter library cells. Viral infections were performed with a MOI of one to transduce approximately each cell of the library (each expressing a single ORF) but avoiding a too strong overexpression of exogenous proteins. As expected, stable expression of CRY1 repressed CLOCK/BMAL1 driven E–box target genes (h*Per1*, h*Cry1*, h*Nr1d1* and h*Dbp*) and altered circadian dynamics compared to library cells stably expressing the CRY1mut (see Fig. 4.7 A and B). *Note:* To exclude effects of residual rhythmicity, both library cell lines were kept in constant conditions for four days before harvesting.

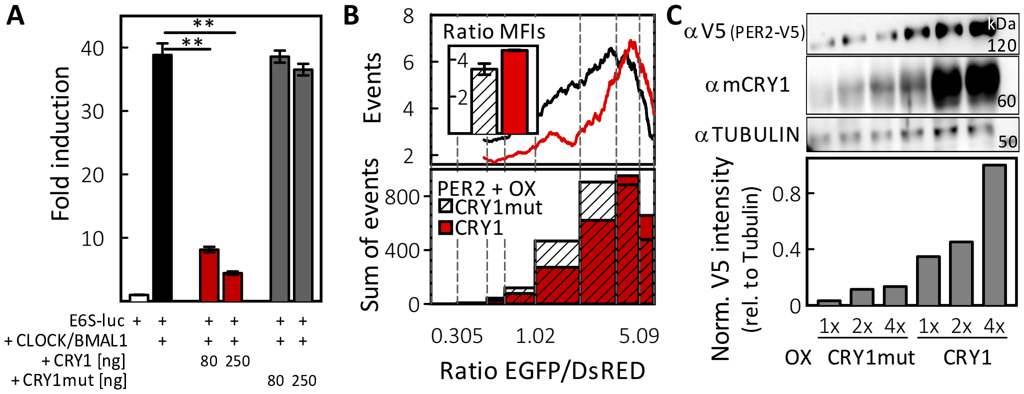


Figure 4.6: CRY1 overexpression 'clamps' the clock and stabilizes PER2. (A) Co–transactivation assay in HEK293 cells. Co–transfection of *Clock* and *Bmal1* together with a six-times repeat of an E–box motif fused to luciferase (E6S–luc) results in an increase of luciferase activity, as measured by the emission of photons. The CLOCK/BMAL1 mediated induction was dose–dependently repressed by the co–transfection of Cry1 but not by Cry1mut. Depicted is the fold induction of emitted photons relative to E6S–luc alone. Mean values $\pm$ SD ($n = 3$) of representative results of two independently performed experiments are given. **P $<$ 0.01 (Mann–Whitney U test). *(B, C)* CRY1 stabilizes PER2, resulting in an increased PER2 protein abundance[125,73]. (B) Increased protein abundance of EGFP–PER2 in the GPS reporter upon overexpression (OX) of CRY1 but not CRY1mut was observed. The upper panel depicts the distributions of EGFP–PER2/DsRED ratios upon OX of CRY1 (red color) or its mutant (black color). The inset shows the quantification of MFI ratios ($n = 3$, mean values $\pm$ SD). Dotted lines indicate eight fractions of the preparative cell sort (see subsection 4.2.2). Expected number of cells in each fraction after preparative cell sort is depicted in the lower panel. For visualization, box width is sized according to the fraction size (log2 scale). (C) Western blot of protein lysates from U–2 OS cells expressing PER2–V5 and different doses of CRY1 or CRY1mut. Blots were immunostained with antibodies against V5 to detect PER2–V5 (upper blot, expected size 137 kDa), against mCRY1 to control for CRY1 and CRY1mut OX (middle blot, expected size 66 kDa), as well as against TUBULIN, serving as loading control (lower blot, expected size 55 kDa). Quantification below is depicted as V5 detected intensities relative to the loading control, normalized to the maximal V5 intensity. Representative blot of two independent experiments. *Note:* PER2 protein was described to stabilize CRY1 protein[125]. Thus, detected CRY1 intensities are elevated compared to CRY1mut. Amount of OX of CRY1 and of CRY1mut proteins are comparable for each dosage as seen in equally treated control cells without additional OX of PER2–V5 (data not show). α–antibody against, HEK293–human embryonic kidney cells 293, kDa–kilo Dalton, Norm.–normalized, OX–overexpression.

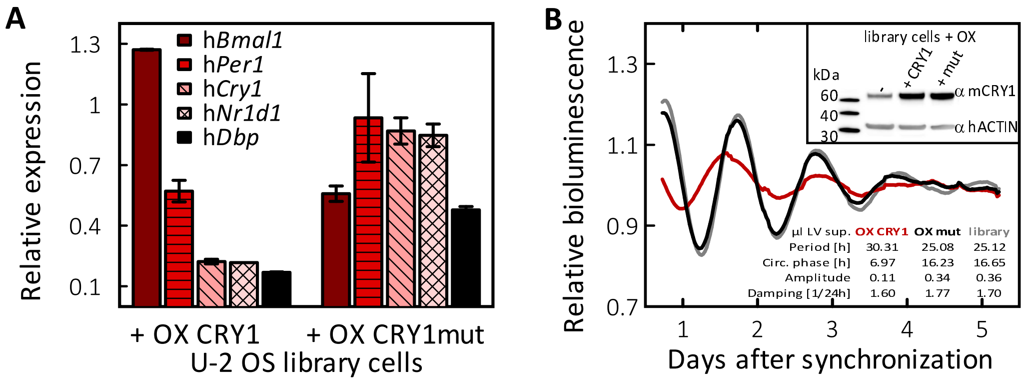


Figure 4.7: Effect of CRY1 overexpression on circadian dynamics in library cells. Stable overexpression of CRY1 or its mutant in U–2 OS hORFeome V5.1 GPS reporter library cells. (A) Gene expression analysis in library cells with OX of CRY1 or CRY1mut was done four days after medium change. CRY1 OX reduces gene expression of CLOCK/BMAL1 driven E–box target genes (h*Per1*, h*Cry1*, h*Nr1d1* and h*Dbp*) and increases h*Bmal1* expression compared to OX of CRY1mut. Gene expression is normalized to h*Gapdh* and depicted relative to no OX. Mean values ± SD of six replicates of two independent experiments are given. (B) Bioluminescence recordings of a *Bmal1*–promoter luciferase reporter over five days after synchronization. Circadian dynamics are altered upon CRY1 OX compared to library cells alone or OX of CRY1mut. Circadian parameters are plotted in the bottom right corner. Depicted are smoothed 24–hrs running average trend–eliminated data (mean of n = 8 replicates). Black curve – library cells, red curve – with OX of CRY1, gray curve – with OX of CRY1mut. The inset shows protein levels of overexpressed CRY1 and its mutant in library cells. Given is one representative blot of two.

4.2.2 Preparative Sort of Library Cells

The hORFeome library is embedded in a fluorescent reporter, where the EGFP/DsRED ratio is used as a proxy for abundance and thus stability of the fused ORF (see subsections 4.1.2 and 4.1.3). To extract information of EGFP/DsRED profiles for individual cells within the library (representing individual ORFs), cells were preparatively sorted into fractions of increasing EGFP/DsRED ratios. Fractions were spaced into similar sizes according to log2 scaled ratios (see dotted lines in Fig. 4.4 A and 4.6 B; *Note:* equal sizing was not possible due to technical limitations). To guarantee a coverage of potential 16,000 ORFs at least 300,000 cells per fraction were collected, corresponding to a potentially 19–fold overrepresentation of the whole library per fraction.

4.2.3 Specific Amplification of Library ORFs

The reporter library is stably integrated into U–2 OS cells, whereby each cell carries one ORF (see subsection 4.1.3). Thus, sequence information of expressed EGFP–

fusion proteins can be retrieved from genomic DNA. To identify and quantify ORFs of cells sorted into fractions, a library specific nested PCR protocol was established. In order to guarantee a quantitative amplification the amount of template, number of cycles and reproducibility of both single PCR reactions was rigorously determined in a test setup. To this end, five cell populations, each carrying a different ORF within the GPS reporter (m*Bmal1*, m*Cry1*, m*Per2* and h*Ttc9c*, a single clone derived from the library) were used. Cells were counted, serially diluted and added to 16,000 library cells. In the end, five cell mixtures of library cells containing different dilutions of different spike–ins were obtained (see Fig. 4.8 A).

A nested PCR protocol with five cycles of pre–amplification followed by 20 cycles of the second PCR showed to be efficient for a quantitative amplification, as revealed by real–time PCR analysis of spiked–in ORFs (see Fig. 4.8 B; for details on the nested PCR optimization process see appendix Fig. A.5 and methods subsection 3.1.1). According to the three–fold dilutions, spiked–in ORFs were detected with an expected Cq value differences of 1.5. Thus, the established nested PCR protocol is sensitive enough to quantitatively amplify one to 81 copies within 16,000 cells and can be applied for a quantitative amplification of ORFs from library cells.

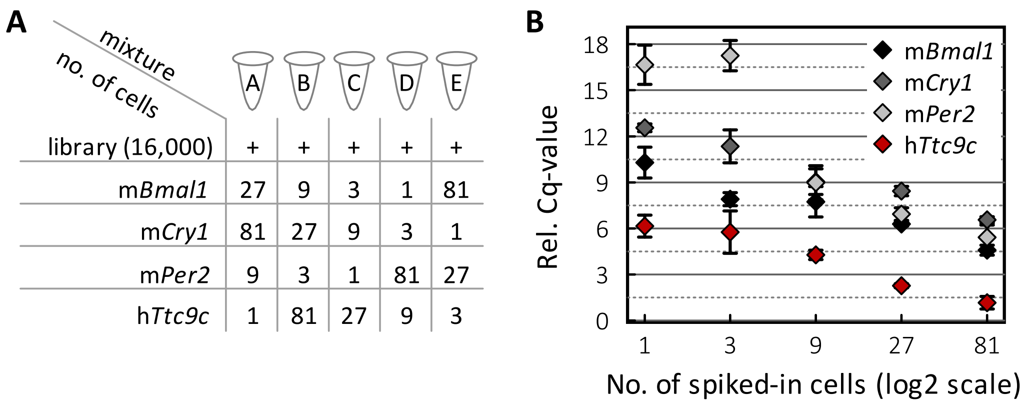


Figure 4.8: Establishment of a quantitative nested PCR. (A) Design of the test setup. Cells containing single ORFs within the GPS reporter (m*Bmal1*, m*Cry1*, m*Per2* and h*Tcc9c*) were serially diluted in five three–fold steps. Each dilution was added to a cell mixtures of 16,000 library cells. In the end, each mixture contains different dilutions of added spike–ins. (B) The established nested PCR protocol can be used to quantitatively amplify spiked–in ORFs of the test setup. The graph shows relative Cq–values for the indicated ORFs plotted according to their number of spiked–in cells. Cq–values are relative to Cq–values obtained by a qRT–PCR primer pair detecting a part of the reporter backbone (loading control). Given are mean values ± SD of three independently performed test setups. Cq–cycle in quantitative real–time PCR, based on regression analysis.

4.3 Microarray Data Processing, Analysis and Validation

This section describes the analysis and validation of microarray data from the comparative one–time point approach to screen for proteins with circadian protein stability (see section 4.2). To this end, mathematically calculated parameters were introduced to characterize processed microarray data. Microarray results were validated and potential circadian protein stability was confirmed for selected candidates. Furthermore, for a deeper characterization of the impact and reliability, screen results were compared to other large–scale data.

4.3.1 Pre–Processing of Microarray Data

Specifically amplified ORFs of cells sorted into eight fractions were identified and quantified on exon microarrays (Agilent 4x180k). For both sub–libraries (overexpressing either CRY1 or CRY1mut), purified PCR products of two consecutive fractions were pooled, labeled and competitively hybridized with corresponding pooled fractions[‡].

For microarray data pre–processing, raw data were background subtracted and normalized within each and over all four arrays (done in cooperation with K.Jürchott, for details see methods sub–subsection in 3.4.1 and appendix Fig. A.6 A).

Quality control of microarray sample (pre–) processing

In order to control for sample processing (nested PCR, labeling and hybridization), ORF spike–in controls that are not present in the library but detectable on the microarray were applied (see Fig. 4.5 orange arrows).

To control for the variability of the nested PCR between fractions and libraries, spike–in controls were added to genomic templates of each fraction before the 1^{st} PCR amplification. To this end, genomic DNA of two U–2 OS cells lines stably expressing human ORFs within the GPS reporter were used. DNA amounts representing one and ten copies of h*Rorß* and h*Csnk2ß* were applied, respectively.

Additional spike–in controls were added to the purified PCR products to control for variability after nested PCR amplifications (e.g. labeling of probes or microarray hybridization). Three human ORFs were selected (h*Bhlhe40*, h*Bhlhe41* and h*Npas2*). PCR products of their coding sequences were serial diluted and quantified to use

[‡]Sample labeling and competitive hybridization was done by the service provider Source Bioscience, Berlin.

different concentrations of their PCR products as spike–in controls (see appendix Fig. A.7). In summary, detected intensities on the microarray of the spiked–in ORFs were in accordance with the added amounts. Intensities between fractions and between sub–libraries were in a similar range (deviation below 5 % of their mean values), indicating an equal processing of all samples before and within the microarray analysis (see Fig. 4.9 A). However, for one coding sequence (h*Bhlhe40*) large deviations of microarray signals were detected. This might be due to inappropriate probe sets detecting this gene.

Coverage of library ORFs

To extract information on how many library ORFs were be detected, a threshold for minimal fluorescent intensity was set to 10.9 for both labeling colors (log2 scale, for details see methods sub–subsection in 3.4.1 and appendix Fig. A.6 B). Probe sets (ps) whose intensities were below the threshold in all four fractions of one sub–library were removed. The remaining probe sets were condensed to gene symbols (gs) if they fulfilled the following 'probe set' criteria: (i) more than one probe sets is present for this gene symbol (ps $>$ 1) and (ii) more then 50 % of probe sets per gene symbol present on the array were detected above the threshold (ps $>$ 50 %). The latter filter was applied to reduce the number of non–full length ORFs form further analysis (see subsection 4.1.3). Together, 54,808 probe sets were condensed to 6,059 detected gene symbols of 16,521 present on the array. Of the detected, 5,655 gene symbols are denoted within the library. 77 % of library ORFs are present on the array. Of those, 47 % were detected (see Fig. 4.9 B).

4.3.2 Microarray Data Processing

Probe sets with background–adjusted and normalized intensities above the threshold were further processed in order to identify ORFs with altered abundance upon CRY1 OX. Intensities of a single probe set over four consecutive pooled fractions describe the EGFP/DsRED profile of the corresponding ORF (see exemplary distribution of PER2 in 4.6 B, lower panel; or appendix Fig. A.8). In order to select gene symbols with profile differences between both tested conditions (representing altered protein abundance), two mathematically calculated indices were defined, that (i) quantify the difference of two intensity distributions and, (ii) determine whether OX of CRY1 stabilizes or destabilizes the corresponding protein (right or left shift on EGFP/DsRED axis, respectively).

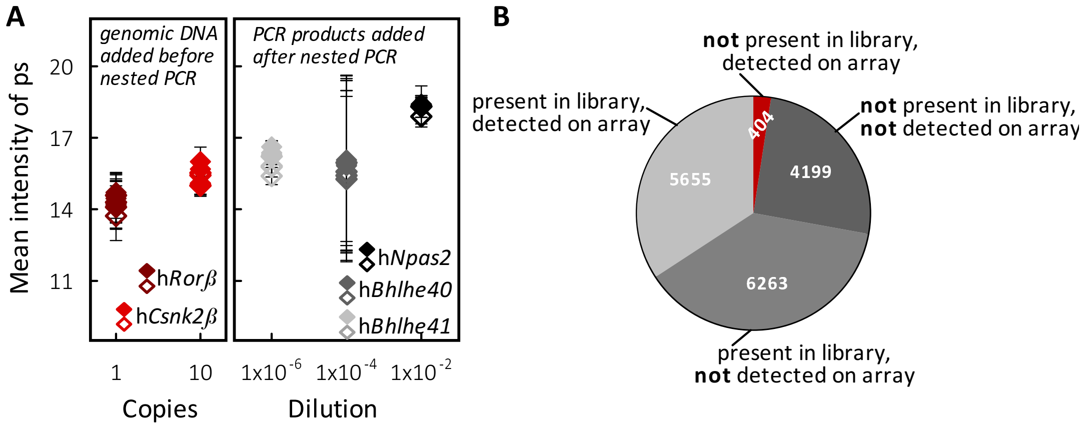


Figure 4.9: Microarray data pre–processing. (A) Quality control of microarray samples. Spike–in controls were added to the fraction samples of either library before and after the nested PCR (see Figure 4.5). Depicted are mean intensities of probe sets from the microarray analysis (y–axis) versus the amount of added spiked–ins (x–axis, either as the amount of cell copies or the dilution of the PCR product). Red and brown colors represent spike–in controls added before the nested PCR (left plot), gray to black colors depict spike–in controls added after the nested PCR (right plot). Filled symbols represent data obtained from the four pooled fractions of CRY1mut overexpressing library, open symbols of CRY1 overexpressing library cells. Given are mean values ± SD. (B) Given is the distribution of 16,521 gene symbols (gs) with more than on probe set (ps) present on the microarray. All ps with intensities above the threshold were condensed to gs according to the 'probe set' criteria*. The number of gs per group are given in the pie chart. *gs that fulfill 'probe set' criteria: more than one ps is present on the array and at least 50 % of ps per gs are detected.

The Protein Stability Index

The Protein Stability Index (PSI, introduced by[106]) describes where most intensities of the EGFP/DsRED profile accumulate. Relative intensities of each fraction are multiplied with the fraction number and summed up (see Fig. 4.10 $A_{(1)}$). Thus, low relative PSI (r.PSI) values describe maximum intensities observed in earlier fractions and vice versa. The difference between r.PSI values of CRY1 minus CRY1mut overexpression condition describes the direction of change (e.g. negative Δr.PSIs values for CRY1 destabilized proteins; see Fig. 4.10 $A_{(2)}$). The resulting Δr.PSIs indicate how far intensity profiles are apart, but cannot be used for a quantitative ranking because changes of equal intensities from fraction one to two are less scored than from fraction three to four. For an example calculation see appendix Fig. A.8.

The Euclidean Distance

In addition, to quantitatively rate how far distribution profiles are apart, the Euclidean Distance (ED) was introduced. The ED calculates the distance between two points in a four-dimensional space. Hereby, measured intensity of each fraction per probe set describes its location within four dimensions. The ED between both overexpression conditions per probe set was calculated according to the formulas in Fig. 4.10 $A_{(4,5)}$[§]. For an example calculation see appendix Fig. A.8.

Thresholds Defining CAAPs

In order to create a list of '**C**RY1 mediated **a**ltered **a**bundant **p**roteins' (CAAPs), thresholds for the ED and Δr.PSI were defined. To this end, z–scores were calculated for Δr.PSI per probe set (see formula in Fig. 4.10 $A_{(3)}$). All probe sets with Δr.PSI values ± 1.5 standard deviations apart from the mean of the population and above an ED of 3.5 were selected (see Fig. 4.10 B, dashed area). Probe sets were condensed to gene symbols if they fulfilled the 'probe set' criteria (see subsection 4.3.1). This resulted in a list of 540 CAAPs (9 % of detected gs) with 281 being stabilized and 259 being destabilized upon CRY1 overexpression (see Fig. 4.10 C). Data were ranked according to the ED which was directed based on the corresponding Δr.PSI value (e.g. negative ED value if Δr.PSI is negative; see whole list of CAAPs in appendix Tab. A.1; for more details on data processing see methods sub–subsection in 3.4.1).

For further validation experiments, 13 candidates with various degrees of protein stability alterations upon OX of CRY1 were randomly selected if (i) their genes were not circadian transcribed (according to Hughes *et al.*[40]) and, (ii) an antibody, tested for endogenous protein was available (see highlighted dots in Fig. 4.10 B and appendix Tab. A.2). *Note:* Candidates without circadian transcripts were chosen to demonstrate in subsequent experiments that circadian protein abundance can be achieved via circadian degradation.

4.3.3 Validation of Microarray Data

The condensation of probe sets that passed the thresholds of detection, 'probe set' criteria, z–scores$_{\Delta r.PSI}$ and ED resulted in list of gene symbols with specific EGFP/DsRED distribution profiles over four fractions. To validate the observed intensity profiles by a microarray independent method, qRT–PCR analysis was performed for several randomly selected gene symbols. Here, all eight fractions (before pooling, see Fig. 4.5) were analyzed. Quantitatively measured amounts of coding sequences resembled the distribution profiles of the microarray data for most

[§]The use of the ED was initiated in cooperation with K. Jürchott.

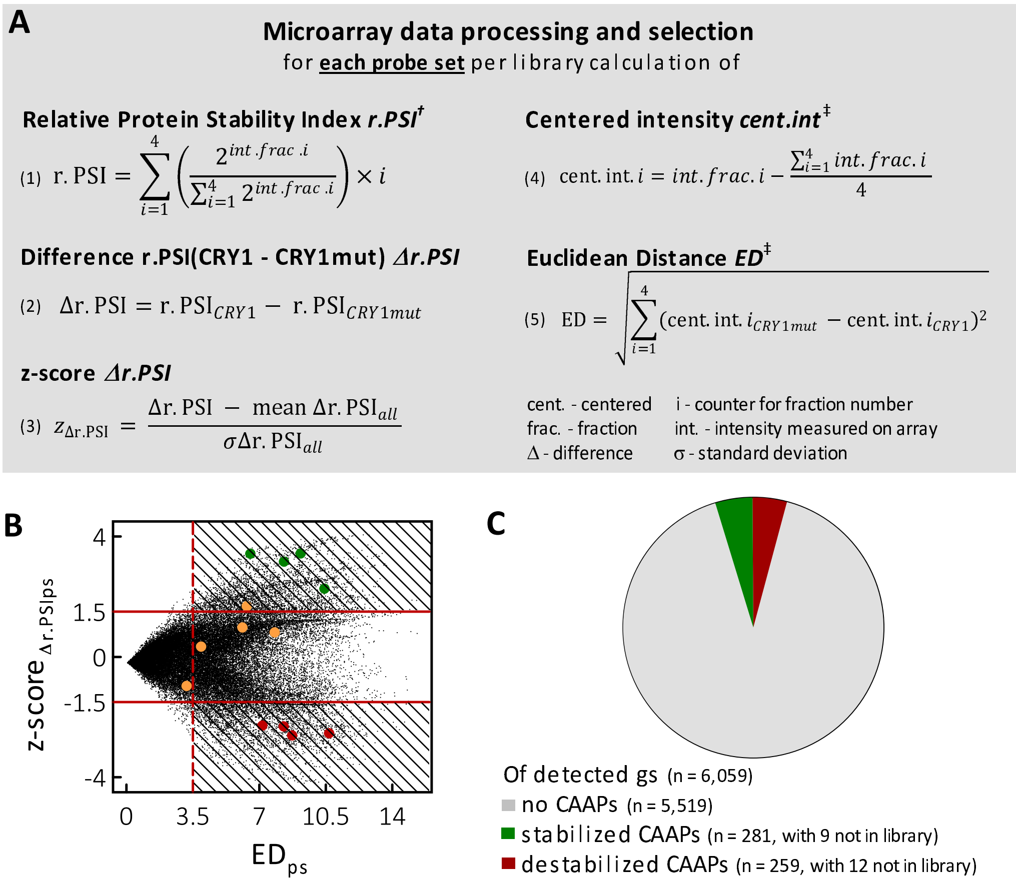


Figure 4.10: Microarray data processing. *(A, B)* Depicted are calculations and thresholds of parameters to define CAAPs. (A) Relative Protein Stability Indices (r.PSIs), their difference between ps of CRY1 and CRY1mut overexpressing libraries (Δr.PSI) and z–scores$_{\Delta \mathrm{r.PSI}}$ were calculated for each detected ps according to the depicted formulas A$_{(1-3)}$. † PSI method was adapted from[106]. The Euclidean Distance (ED) for each ps was calculated according to the formulas A$_{(4-5)}$. ‡ The usage of the ED was initiated in cooperation with K. Jürchott. For an example calculation see appendix Fig. A.8. (B) The dot plot visualizes EDs (x–axis) and z–scores$_{\Delta \mathrm{r.PSI}}$ (y–axis) of all detected ps. Red solid lines and red dashed line represent the defined thresholds to select CAAPs (hatched areas). Colored dots represent condensed gs of candidates selected for further validation experiments. Green and red dots are CAAPs that are stabilized or destabilized, respectively. Orange colored dots represent ORFs that were not selected as CAAPs. *Note:* Highlighted dots are enlarged compared to ps. One orange gs appears within the area of CAAPs, however was dismissed as not 50 % of its ps fulfilled the selection criteria. (C) The pie chart depicts CAAPs that are found to be stabilized (green) or destabilized (red) of all detected gene symbols.

samples (see Fig. 4.11 A–C, upper and middle panels). Δr.PSI values calculated for both, microarray and qRT–PCR data, correlated highly significantly ($P < 0.001$, Spearmann Rank Correlation, Fig. 4.11 D, left plot).

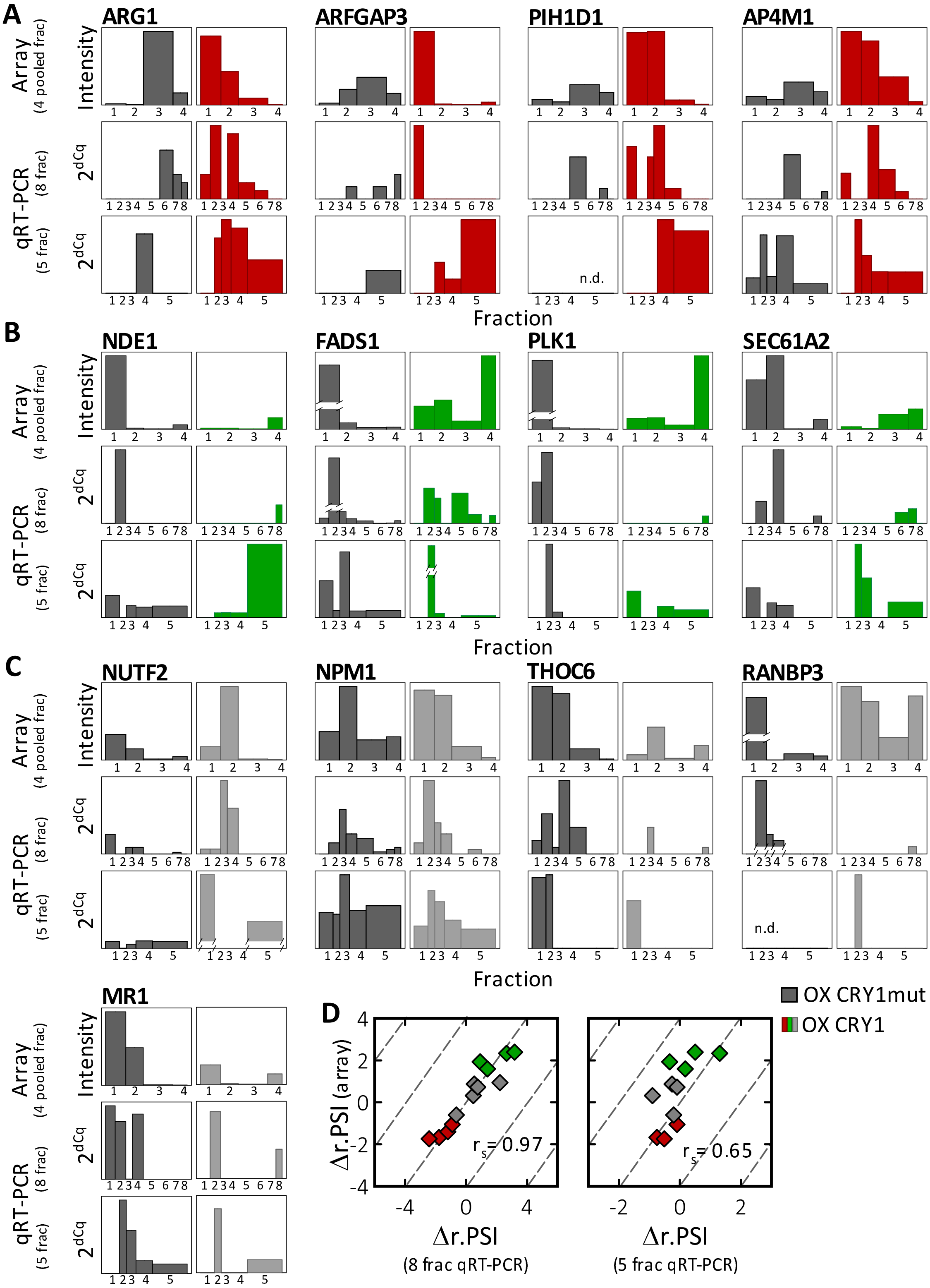
For an additional validation of observed distribution profiles, qRT–PCR analysis for selected gene symbols was performed with independently sorted library cells overexpressing CRY1 or its mutant. Although the size and number of fractions differ substantially from the samples analyzed on the microarray, a rather similar distribution profile and correlation of Δr.PSIs was observed for the tested candidates ($P < 0.05$, Spearmann Rank Correlation, Fig. 4.11 A–C, upper and lower panels and D, right plot; for details on Δr.PSI calculation see methods sub–subsection in 3.4.1).

4.3.4 Validation of Circadian Protein Stability

Cellular protein abundance is controlled at various levels from transcription to degradation[64,63,99]. In order to test whether selected candidates from the microarray analysis are targets of circadian degradation, a method monitoring specifically protein degradation uncoupled from ORF specific synthesis was applied. To this end, a reporter construct driven by a constitutive promoter where each ORF is C–terminally fused to luciferase (LUC) was introduced (see Fig. 4.12 A, upper scheme). The reporter was stably integrated into U–2 OS cells and luciferase signals were monitored

---

Figure 4.11 *(following page)*: Validation of microarray intensity profiles. *(A–C)* Intensity profiles of candidates selected for validation. *Upper panel* show the delogarithmized intensities detected on the microarrays for each pooled fraction (one to four) of the indicated ORFs. *Middle panel* depict 2^{dCq} values of qRT–PCR analysis in eight fractions corresponding to the pooled samples in the panels above. *Lower panel* show results of qRT–PCR analysis done with five independently sorted fractions. 2^{dCq} values are determined in triplicates. All values are normalized to the reporter backbone and adjusted to the corresponding amount of cells per fraction at the preparative sort. Y-axes of corresponding graphs per ORF and sample pair are equally scaled, except for those with marked axis–breaks. Here, the lower axis break corresponds to the maximum scale of the corresponding plot. Box widths (x–axis) are scaled according to the fraction sizes at the preparative sort (log2 scale). Dark gray colors represent samples with overexpression of CRY1mut. Red (A) and green colors (B) represent CAAPs that are destabilized or stabilized (left or right shift of intensity profile after OX of CRY1, respectively). Light gray (C) samples refer to candidates that are not defined as CAAPs. (D) Correlation plots of Δr.PSI values calculated from microarray intensities (y–axis) or 2^{dCq} values of qRT–PCR (x–axis) of eight (left, $P < 0.001$, Spearman Rank Correlation test, $n = 13$) or five (right, $P < 0.05$, Spearman Rank Correlation test, $n = 11$) fractions. For details on calculation see methods sub–subsection in 3.4.1.



A
B
C
D
ARG1
ARFGAP3
PIH1D1
AP4M1
NDE1
FADS1
PLK1
SEC61A2
NUTF2
NPM1
THOC6
RANBP3
MR1
Array (4 pooled frac)
Intensity
qRT-PCR (8 frac)
(5 frac)
2^{dCq}
Fraction
n.d.
OX CRY1mut
OX CRY1
Δr.PSI (array)
Δr.PSI
(8 frac qRT-PCR)
(5 frac qRT-PCR)
r_s = 0.97
r_s = 0.65

serving as a readout of protein abundance of the fused ORF. Changes in luciferase counts are presumably mediated by post–translational events altering the degradation rate of the fused protein, as ORF specific transcriptional, post–transcriptional or translational regulation can be excluded. Indeed, the ratio of luciferase counts after overexpression of CRY1 relative to CRY1mut were increased for PER2–LUC but not for a control protein fused to luciferase[125,73] (Fig. 4.12 B; CRY1 mediated stabilization of PER2 was shown as well in Fig. 4.6 B and C).

The ORF–Luciferase Reporter Monitors Circadian Protein Stability

To detect circadian degradation, cells carrying the luciferase reporter were synchronized via temperature entrainment over four days (see Fig. 4.12 A, lower scheme). The entrainment of cells to cyclic 12–hrs temperature changes has been shown to increase the ensemble amplitude of circadian transcript rhythms[126] and is assumed to be also conveyed to the proteome level. Monitoring of CLOCK– and CRY1mut–LUC fusion proteins after temperature entrainment revealed antiphasic circadian profiles of luciferase signals but not for a control protein fused to luciferase (see Fig. 4.12 C). To test whether the observed circadian luciferase signals are result of phase-specific degradation rates, protein half–lifes of the fusion proteins were analyzed at two time points after release into constant conditions (see Fig. 4.12 C, t1 and t2). To this end, protein *de novo* synthesis was blocked at the indicated times by the administration of cycloheximide (or solvent as a control) and cells were harvested in two hour intervals over six hours. Measured decays of luciferase signals in cell lysates and the resulting calculated half–lifes showed a clear difference between the two time points for the fusion of CRY1mut but not for the control (see Fig. 4.12 D, left and right panels and insets). This might indicate circadian stability of the CRY1mut fusion protein. Although showing a circadian profile after temperature entrainment, no clear difference of protein stability was observed for the CLOCK fusion. This might be due to its lower amplitude rhythm at the second day after release from temperature entrainment (see Fig. 4.12 D, middle panel). Nonetheless, the luciferase fusion reporter represents a valuable tool to analyze protein abundances uncoupled from protein specific synthesis and allows for the detection of circadian abundance changes indicating rhythmic stability.

Circadian Analysis of Selected Candidates within the Luciferase Reporter

Next, the selected candidates of the microarray analysis (including CAAPS and proteins that were not defined as CAAPs, see Fig. 4.10 B) were monitored as luci-

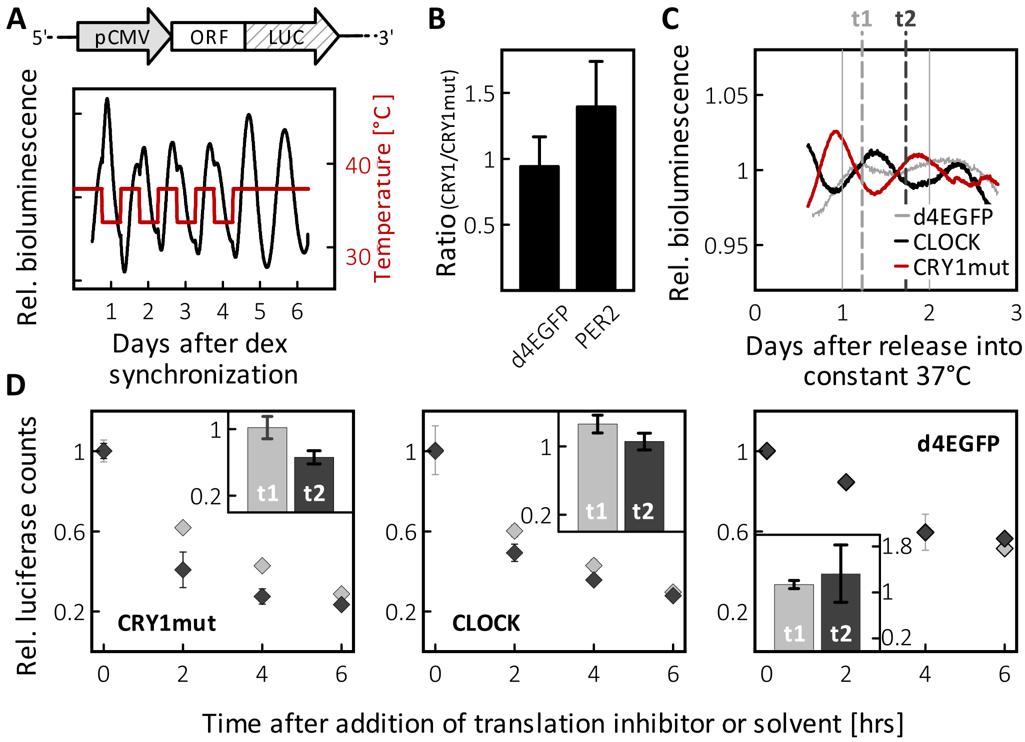


Figure 4.12: Luciferase reporter to analyze circadian protein stability. (A) Schematic representations. *Luciferase reporter* (upper panel): any ORF is C–terminal fused to luciferase and transcribed by a constitutive CMV promoter. *Temperature entrainment setup* (lower panel): U–2 OS cells stably expressing ORF::luc reporter constructs were temperature entrained. After dexamethasone synchronization, cells were kept at 37°C for 18–hrs, followed by 12–hrs 33°C—12-hrs 37°C cycles over four days before being released into constant 37°C conditions. Meanwhile, luciferase counts were monitored online. (B) U–2 OS cells stably expressing PER2 and a control protein as luciferase fusions. Depicted are ratios of luciferase signals after overexpression of CRY1 relative to CRY1mut (n = 5, mean ± SD). *(C, D)* CRY1mut, CLOCK and d4EGFP luciferase fusion proteins were analyzed as described in *(A)*. (C) Depicted are 24–hrs running average trend–eliminated bioluminescent data after release into constant 37°C. (D) Analysis of protein half–life at two times (t1 = 29–hrs and t2 = 41–hrs after release into constant conditions, see light and dark gray vertical dashed lines in *(C)*). After addition of translation inhibitor CHX or its solvent (DMSO) at the indicated time points, cells were harvested in 2–hrs intervals over 6–hrs. Luciferase counts were measured in the supernatant of lysed cells. Depicted are normalized luciferase counts (CHX/DMSO) relative to their maximum value. The insets shows calculated protein half–lifes, relative to half–lifes obtained from unsynchronized cells. Given are mean values of n = 2 biological samples with each 2 technical replicates; error bars represent minimum/maximum value of two biological samples. *Note:* Due to the low number of biological replicates, statistical testing was not done.

ferase fusions after temperature entrainment as described above. Indeed, circadian oscillations with higher amplitudes rhythms compared to the control were detected. Slightly different phases were observed, however, do not distribute over the whole circadian cycle. Interestingly, no clear visual differences of amplitude or damping between CAAPs and candidates that are not defined as CAAPs were observed (see Fig. 4.13, CAAPs are green and red colored). Only NPM1, the ORF of the selected validation candidates whose abundance was changed the least after OX of CRY1 compared to CRY1mut (as indicated by the euclidean distance) did not show a circadian luciferase profile after temperature entrainment (see Fig.4.13).

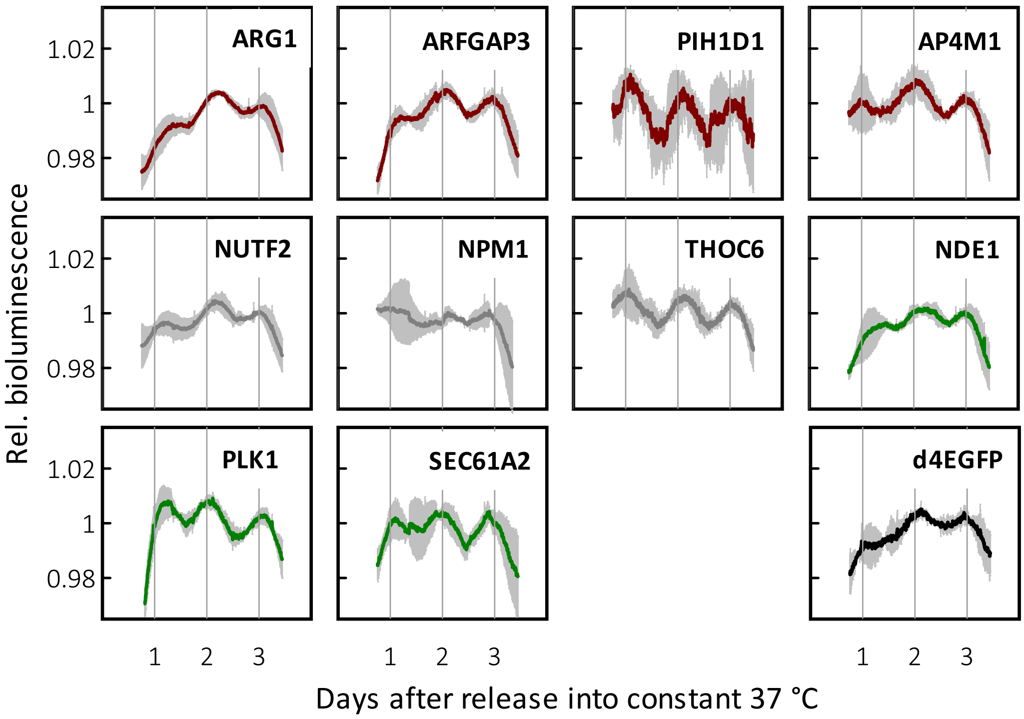


Figure 4.13: Validation of selected candidates within the luciferase reporter. ORF luciferase fusion proteins in temperature-entrained U–2 OS cells. Depicted are 24–hrs running average trend–eliminated mean bioluminescent data after release into constant 37°C. Red and green data represent destabilized or stabilized CAAPs, respectively. Dark gray data are candidates that did not fulfill the selection criteria of CAAPs. The black data represent an artificial control protein (d4EGFP). Light gray areas indicate the standard deviation of three independent measurements.

4.3.5 Bioinformatic Analysis of CAAPs

The high–throughput screen revealed various extents of altered protein abundance indicating different protein stabilities upon overexpression of CRY1. In further investigations, to test whether screen results represent proteins with circadian stability changes that most likely affect their endogenous protein abundance, bioinformatic analysis applying other circadian large–scale studies were performed. Furthermore, potential biological relevance of CAAPs within the molecular core oscillator or enrichment in biological functions, processes or compartments was investigated.

CRY1 Overexpression Stabilizes the Proteome

Overexpression of CRY1 'clamped' the molecular clock of library cells at a specific circadian phase. In liver, CRY1 protein level rise at the end of the dark phase (circadian times (CTs) 18–24; CT 0 indicates the beginning of a subjective day, CT 12 is the beginning of a subjective night)[84,80]. Bioinformatic analysis of circadian proteome and transcriptome data predicted an enrichment of rhythmically abundant proteins at CT 18 (S. Lück, AG Westermark, ITB Berlin; personal communication). This is in accordance with overall peak phases of the circadian liver proteome[41] (and Robles *et al.*, *Plos Genetics*, in press). Thus, overexpression of CRY1 should represent a circadian phase of overall increased protein abundance potentially mediated by increased protein stabilities. Indeed, r.PSI values were significantly larger in cells overexpressing CRY1 compared to CRY1mut indicating a stabilization of many proteins (see Fig. 4.14).

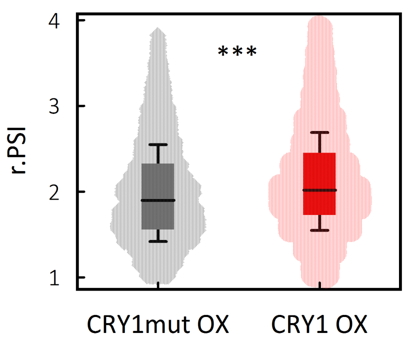


Figure 4.14: CRY1 overexpression stabilizes the proteome. r.PSI values after OX of CRY1 (red data) are significantly different and larger compared to r.PSI values after OX of CRY1mut (gray data). Depicted are scatter plots of r.PSI values of all detected gs with box plots centered at the median, bordering the interquartile range. Error bars represent the standard deviation of the mean. (***P < 0.001 Kolmogorov–Smirnov and Mann–Whitney U test, $n_{CRY1mut,CRY1} = 6059$).

Rhythmically Abundant Proteins Are Enriched Among CAAPs

This study aimed to identify proteins with circadian stability. Such proteins should exhibit circadian abundance profiles if their production is not antiphasic to a circadian degradation, and if the overall protein stability does not masks circadian degradation. Consequently, proteins with rhythmic abundance profiles might be enriched among CAAPs. To test this, unique proteins that were found to be rhythmically abundant in either human blood[43], rat pineal gland[45], mouse retina, SCN or liver[42,43,41](and Robles *et al.*, *Plos Genetics*, in press) were analyzed (see appendix Tab. A.3). Indeed, rhythmically abundant proteins were enriched among CAAPs above the expectation value[¶] (see Tab. 4.1, a).

Circadian abundant proteins identified in other studies, however, do not exclude rhythmically synthesized proteins. To test if rhythmic proteins that are constitutively transcribed are as well enriched among CAAPs, a circadian liver transcriptome analysis was consulted[40]. Although being limited towards a tissue–specific analysis, this study provides high resolution data of circadian transcripts over two circadian cycles. A slight enrichment above the expectation value for rhythmically abundant proteins without rhythmic transcripts was observed within the list of CAAPs. Interestingly, a slight enrichment above the expectation value was also observed for rhythmic proteins that are circadian transcribed (see Tab. 4.1, b and c).

Table 4.1: Circadian abundant proteins

(a) ...with rhythmic or non–rhythmic transcripts

	number
no CAAPs	5519
of those rhythmic proteins	90
CAAPs	540
expectation value	8.8[¶]
found within CAAPs	16

(b) ...with non–rhythmic transcripts

number
2418
57
271
6.39[¶]
10

(c) ...with rhythmic transcripts

number
635
26
90
3.69[¶]
6

CAAPs do not Regulate Overall Circadian Dynamics

This screen aimed to identify proteins whose stability is modulated by the circadian clock. To test whether CAAPs are themselves modifiers feeding back to the molecular clock machinery, results were compared to data of a short hairpin RNA mediated knockdown screen that analyzed circadian dynamics. Within overlapping gs of both studies, no significant enrichment of altered circadian parameters was found between CAAPs and proteins not defined as CAAPs (Fisher's Exact test with gs that are $\pm$ 2.5 SD apart from the knockdown–screen mean values of period, amplitude, damping, phase, magnitude or arhythmicity; data not shown; knockdown data were generated in our laboratory and provided by B. Maier). Thus, CAAPs presumably do not regulated overall circadian core clock dynamics.

[¶]Please note that all expectation values are based on gene symbols detected within this screen (same parent population). They cannot be used for statistical evaluation since the number and identity of analyzed proteins of the published studies are unknown.

CRY1 Interactors – Promising Targets of Circadian Protein Stability

The proteome–wide screen was done in cells where the molecular clock was 'clamped' by the OX of CRY1. Alterations of protein abundance upon OX of CRY1 might indicate circadian protein stability (see subsection 4.2.1). *In vivo*, CRY1 protein levels are circadian[123,80]. Thus, CRY1 protein interactions are limited to a specific time–of–day[117]. Consequently, CRY1 interactors among CAAPs might represent promising candidates exhibiting circadian stability, as seen for PER2 which is known to be in complex with CRY1[125] (see Fig. 4.6 B and C, and Fig. 4.12 B).

Of 17 CRY1 interactors (from Unified Human Interactome (UniHI) database, see appendix Tab. A.4) detected in the microarray analysis, two were found within the list of CAAPs – PPP2R1B and HSPA8. They represent interesting candidates for further validation experiments revealing circadian protein stability.

Functional Classification of CAAPs

CAAPs might represent proteins that exhibit circadian stability. A proteome–wide occurrence and involvement of circadian degradation in specific pathways or molecular processes is unknown. Thus, a functional analysis was done to characterize if CAAPs are enriched in specific biological processes, molecular functions or among cellular components. To this end, Gene Ontology enrichment analysis and subsequent pruning of related terms was performed using the GO-Elite software[108] (for details on software settings see methods subsection 3.4.2). Resultant GO terms with p–values < 0.05 (Fisher's Exact; see whole list in appendix Tab. A.5) were manually combined to major terms under consideration of their GO ancestor charts and child terms. These major terms are 'Mitosis', 'Vesicles and secretion (including related subterms 'Transport', 'Proteolyse' and 'Cell-cell contact'), 'Protein binding', 'Immune response', 'Development and growth', 'RNA processing' and 'Lipid biosynthesis' (see Fig. 4.15). The major terms describe a common theme of the belonging GO–IDs, however, can be as well related to GO–IDs classified into other major terms.

Interestingly, GO terms associated with 'Vesicles, secretion and transport' were as well found in previous studies investigating the circadian proteome of mouse retina, SCN and liver tissue[42,43] (and Robles *et al.*, *Plos Genetics*, in press).

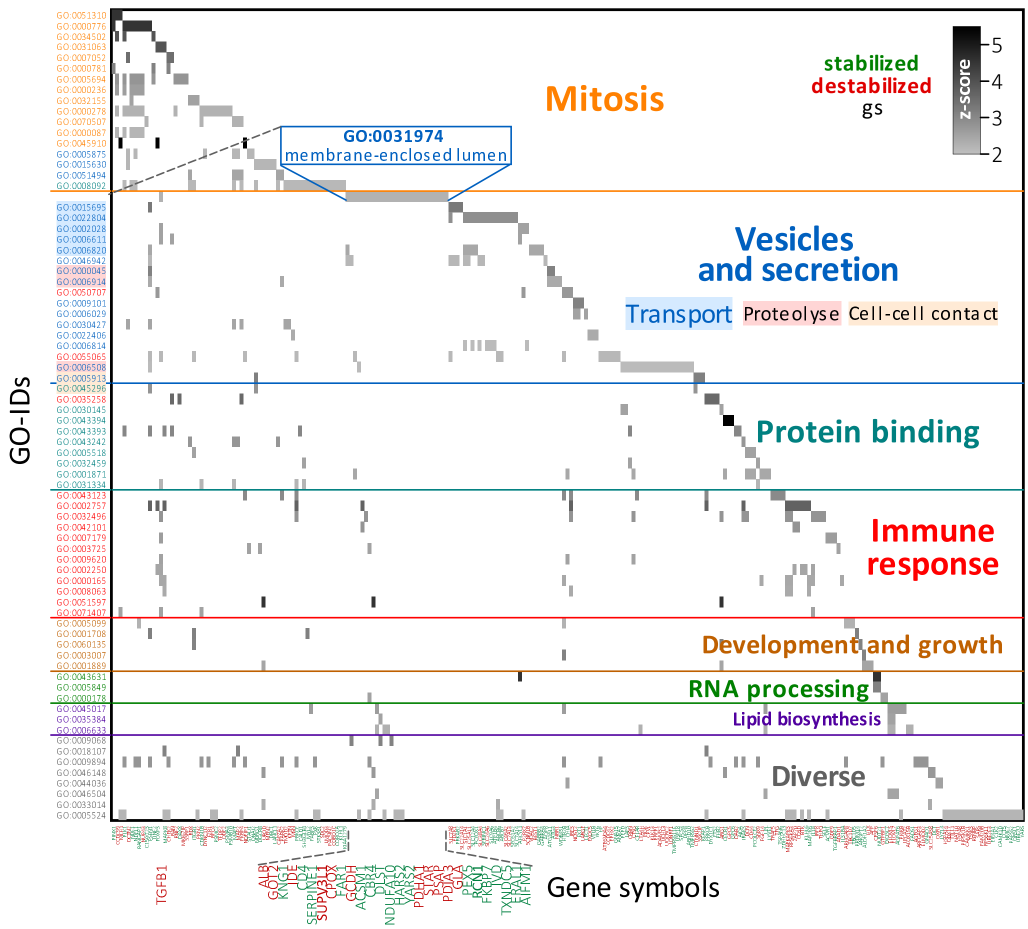


Figure 4.15: Classification of CAAPs in GO terms. Gene Ontology analysis of CAAPs and reduction to non–redundant terms was done with GO-Elite software[108]. Associated GO terms with $P < 0.05$ (Fisher's Exact test) were manually categorized in seven major terms (bold, GO–IDs are color coded accordingly). Font size of term name is adjusted according to the percent of CAAPs of all detected gene symbols corresponding to the major term (e.g. 10 detected gs of the microarray analysis correspond to term X and of those 7 are CAAPs). GO terms with affiliation to more than one major term are categorized in one and highlighted with the color of the other term. Within the term 'Vesicles and secretion' subterms exist and are marked with background colors. In the matrix, gs corresponding to a GO–ID are highlighted from light to dark gray according to their overrepresentation z–score. Gs names are colored green or red if they were stabilized or destabilized by CRY1, respectively. GO–ID:0031974 and corresponding gs are emphasized exemplarily. *Note:* Six GO terms that were high in the GO hierarchy and contained > 200 gs were excluded form visualization.

5 Discussion

Endogenous, self–sustained 24–hrs cycles are omnipresent in organisms living on earth. In mammals, a variety of physiological processes are circadian and a loss or disruption of their circadian control is associated with adverse health effects and diseases[24,26]. Not surprisingly, endogenous molecular rhythms driving diverse physiological functions are precisely controlled.

Circadian control is already found at the transcriptome level. High–throughput screens revealed rhythmic control of 10 % of the transcriptome[36–40]. In discrepancy to the amount of rhythmic transcripts, analysis of the circadian proteome, however, showed that up to 20 % are rhythmically abundant[41–45] (and Robles *et al.*, *Plos Genetics*, in press). In addition, circadian proteins with rhythmic transcripts that display a 12–hrs delay between mRNA and protein abundance peak have been detected and cannot be explained by consecutive transcription–translation processes[41] (and Robles *et al.*, *Plos Genetics*, in press; S. Lück, AG Westermark, ITB Berlin; personal communication). Together, this indicates the involvement of additional mechanisms regulating circadian proteins that miss, or have antiphasic transcript rhythms.

What are additional modes controlling the rhythmicity of a protein? Different regulations are conceivable: (i) rhythmic post–transcriptional mechanisms – indeed, circadian control of global post–transcriptional events was demonstrated in mouse liver. Poly(A) tail lengths allow cells to alter protein level rapidly without *de novo* transcription. Kojima and colleagues found transcripts with circadian poly(A) tail lengths, potentially resulting in rhythmic protein abundances. However, only below 1 % of expressed transcripts with circadian poly(A) tail lengths did not show a detectable circadian rhythm in their pre– and /or mRNA level.

Furthermore, (ii) circadian translation is a potential mode regulating rhythmic protein abundances – Jouffe *et al.* described the rhythmic control of ribosome biogenesis and translation initiation. Of expressed genes without circadian transcript, about 2 % were rhythmically translated in mouse liver[127].

Together, the described extent of rhythmic post–transcriptional and translational regulations do not completely cover the gap of about 10 % of circadian proteins that are not rhythmically transcribed. An other possibility, (iii) the involvement of post–translational events regulating circadian protein abundances, however, has not been studied globally so far. Up to now, circadian regulation of protein degradation has been demonstrated for only two examples (PER2 and p53)[78,79,102].

This study aimed to identify the existence and extent of global circadian protein stability – the circadian 'stabilome'.

5.1 A Method to Screen Global Circadian Protein Stability

To answer the question on the existence of a circadian 'stabilome', a method feasible to analyze (i) on a proteome–wide scale, (ii) protein stability and, (iii) under circadian conditions, is needed. Is the used Global Protein Stability (GPS) method suitable for a circadian 'stabilome' analysis?

5.1.1 A Proteome–wide High–Throughput Analysis

For a global analysis of protein stability, a method allowing to investigate the proteome in a reasonable frame of time and costs is needed. In this study, a library consisting of about 16,000 human open reading frames (ORFs) was used. ORFs are embedded in a fluorescent reporter construct and are stably integrated into cells, with one ORF per cell. Encoded proteins are identified by their ORF sequences within the the cellular genomic DNA. Thus, in contrast to proteome arrays[128], protein specific analysis is possible, independently of the availability of antibodies. Furthermore, library ORFs are uncoupled from endogenous protein specific synthesis and are constitutively transcribed by an artificial promoter. Thus, proteins with low endogenous levels can be analyzed with the GPS method in contrast to mass spectrometry based approaches where low abundant proteins are hardly detectable[129,130]. Consequently, depending on the ORF library, the fluorescent reporter based method allows for a larger coverage of the proteome.

Qualitative analysis of the hORFeome V5.1 library used in this study, however, revealed less than 50 % of library ORFs being full length clones. Most of the sequenced library ORFs are truncated clones or empty reporters (see Fig. 4.4 C). This high failure rate is common within ORF cDNA collections. They are mainly based on errors of the PCR amplification caused by polymerase mutation rates or sequence mutations within PCR primers[118,119]. In addition, the propagation of ORF library clones into reporter plasmids via the Gateway® technology is error prone and may account for empty reporters.

Consequently, large–scale high–throughput approaches hold method specific limitations that impede with a proteome–covering analysis. Probably, the combination of various approaches would be best for a complete proteome–wide screen.

5.1.2 Analysis of Global Protein Stability

Next to a proteome–wide coverage, the analysis of a 'stabilome' requires a method allowing to analyze protein stabilities in high–throughput. Methods that analyze cellular protein stability share a required feature. They stall or uncouple endogenous protein synthesis to consider solely the decay, thus stability of the protein[63,106,131]. Protein *de novo* synthesis can be stalled by the inhibition of translation. The administration of pharmacological inhibitors like cycloheximide or lactimidomycin however have adverse side effects. They are cytotoxic and the overall block of translation similarly affects regulatory proteins that eventually control the stability of proteins of interest[132].

On the other hand, 'uncoupling' freshly synthesized proteins from the existing pool of proteins is achieved by pulse–chase labeling. For a limited period of time, markers are present and incorporated into proteins within *de novo* synthesis. In mass spectrometry based approaches, especially the labeling with heavy amino acids has been successfully applied to monitor global protein decay[133,63]. This procedure, however, has so far not been done for a circadian analysis of protein stability, probably because of the high expenditure of time and costs[128].

Another possibility to 'uncouple' proteins from endogenous synthesis is the use of cellular reporter plasmids. Thereby, only protein coding sequences are expressed under the control of an artificial promoter. Synthesis rates of proteins within the reporter are similar due to the lack of both 5' and 3' regulatory regions. Consequently, differences in protein abundances are exclusively based on protein specific degradation[131,106].

The method applied in this study is based on a fluorescent reporter, stably integrated into cells (see Fig. 4.1). As described above, this reporter allows to investigate protein stability 'uncoupled' from protein specific synthesis. The fluorescent proteins are used for the readout of protein abundance, thus stability. Their usage, however, holds as well limitations.

Specific Characteristics of the Applied Fluorescent Reporter

Within the GPS method, in a flow cytometry based approach, proteins are analyzed as EGFP fusions relative to DsRED. EGFP fusion and DsRED are cistron proteins being independently translated. Although translation efficiency of the IRES (for EGFP–ORF) was shown to result in a 1:1 molar ratio to the 5' cap initiated translation (for DsRED)[112], a cell line specific usage of translation start sites cannot be excluded. Moreover, maturation of DsRED is slightly faster than of EGFP (about 18 min)[113,114]. In addition, DsRED protein half–life (about eight days) is twice as high as EGFP[111] half–life. As a consequence, DsRED protein accumulates over time within a cell (as seen over 21 days after medium change, data not shown), while EGFP fusions will reach the steady state abundance, when the rate of production equals that of degradation, earlier.

Altogether, the point mentioned above certainly contribute to the fact that EGFP/DsRED ratios vary over time and in single cells of a population (see Fig. 4.2, A–D). The standardized handling, however, minimizes such variability effects allowing to analyze and compare EGFP/DsRED ratios.

The N–Terminal EGFP Tag

Although EGFP fusion proteins allow to determine fused proteins' stability based on the fluorescence ratio EGFP/DsRED, some limitation need to be taken into consideration. In this study, EGFP is N–terminally fused to ORFs. As mentioned above, approximately 50 % of library ORFs are truncated or empty reporter clones. Their detection could have been avoided by the use of a C–terminal tag. Furthermore, large tags such as fluorescent proteins hold the risk to interfere with protein specific function and/ or subcellular localization. This could as well influence protein specific stability. In addition, an N–terminal tag might be disadvantageous for protein stability analysis. N–termini of proteins display important degrons that regulate their stability (N–end rule). This involves the N–terminal acetylation and eventual conjugation of amino acids by aminoacyl-tRNA transferases[134], which is prevented by a N–terminal tag. Although up to 80 % of mammalian proteins are N–acetylated and harbor thus potential degradation signals according to the N–end rule, the importance of N–end mediated stability is not known on a proteome–wide scale[135,134]. Furthermore, a circadian protein stability as analyzed in here, is assumed to be mediated by protein specific timed post–translational modifications rather than global regulations. Nonetheless, mechanisms controlling circadian protein stability on a proteome–wide scale are also possible and cannot be excluded so far.

Despite these potential limitations, proteins with described different stabilities were successfully classified with the applied fluorescent reporter method (see Fig. 4.3 A and B; and 4.6, B). Protein half–lifes correlated significantly with observed ratios of EGFP/DsRED, indicating that the abundance of a EGFP fusion reflects its stability (see Fig. 4.3 C and appendix Fig. A.2). In this line, protein groups of very stable and unstable proteins investigated in an independent mass spectrometry based approach, corresponded reasonably well to stable and unstable proteins groups analyzed in this study (see Fig. 4.4 B). Together, this indicates that the GPS method is suitable to analyze proteome–wide stabilities.

5.1.3 Analysis of Global Circadian Stability

To investigate circadian protein stability, the artificial nature of the applied method is beneficial. First of all, circadian rhythms control as well transcriptional, post–transcriptional and translational processes making it difficult to analyze rhythmic protein stability endogenously. Secondly, from a physiological perspective, circadian differences of stability probably not affect highly abundant and stable proteins. Moreover, circadian degradation is assumed to be found for rather low abundant and unstable proteins. As mentioned, ORFs within the GPS method are constitutively transcribed allowing to analyze proteins whose endogenous protein level are presumably below the detection limit of other approaches.

Circadian analysis of EGFP/DsRED ratios of control ORFs (d1– and d4EGFP) in synchronized cells, however, revealed non–circadian fluctuations (see appendix Fig. A.3). This might represent a consequence of the above described variability of the fluorescent reporter and prevents the robust detection of abundance changes of time–of–day dependent degradations. To circumvent these limitations, a 'clamped' circadian clock model was introduced.

'Clamping' the Clock – A Cell Model to retrieve Circadian Information

A 'clamped' clock represents a snapshot of a specific circadian phase where clock gene expression, protein level and localization are stalled. Instead of analyzing changing EGFP/DsRED ratios over a circadian cycles, ORF abundances were monitored at one time point, comparing 'clamped' cellular clocks to asynchronous cells. To 'clamp' the cellular clock, a rate limiting component either needs to be removed or added to 'arrest' the molecular feedback loops of the clock machinery. The knockdown of a gene is rarely complete and often leaves residual gene expression. Furthermore, especially the knockdown of clock components is counterbalanced by paralogs and compensatory effects[136]. In contrast, overexpression of clock proteins leads to 'gain of function' effects with compensation being unlikely. In this line, the stable expression of the repressor REV–ERBα in murine liver or in cultured cells was shown to stall circadian rhythmicity at an expected phase of attenuated *Bmal1* expression[47,137] (and appendix A.4, herein referred as hNR1D1).

Dose–dependent overexpression of positive and negative acting clock components revealed CRY1 to 'clamp' molecular circadian dynamics (see appendix A.4). The clock inhibitor CRY1 interferes with CLOCK/BMAL1 mediated transactivation of E–box driven target genes[123] (see Fig. 4.6 A and Fig. 4.7 A). Overexpression of CRY1 thus resembles the gene expression profile of BMAL1 knockout in mouse[138]. In contrast, two previous studies claimed that overexpression of CRY1 does not perturb the molecular oscillator. The two groups used tagged CRY1 protein in Rat-1 fibroblasts. In both studies tagged proteins were functional and repressed CLOCK/BMAL1 mediated transcriptional activity in co–transactivation assays. While within the study by Yamanaka *et al.*, however, reduced amplitude rhythms monitored by a *Per2* promoter-driven luciferase were shown and correlated with the amount of CRY1 overexpression, no such data were available within the study of Fan and colleagues[139,140]. Western blot results, however, showed only a very faint band of introduced CRY1[140]. Thus, in both studies doses of expressed CRY1 level were probably far below of those applied in this study and presumably beneath a critical threshold to 'clamp' the molecular oscillator. In support of the finding of this study, independently performed CRY1 overexpression in U–2 OS cells perturbed molecular circadian dynamics as monitored by a *Bmal1*–promoter luciferase reporter[117]. Altogether, CRY1 and as a control an inactive CRY1 mutant[104] were feasible to establish a 'clamped' clock model for a comparative approach determining circadian protein stability.

PER2 Validates the Comparative 'Clamped' Clock Approach to Screen for Circadian Stable Proteins

PER2 protein was previously shown to be circadian degraded[78,79]. Analysis of PER2 protein within the comparative one–time point model revealed an increased abundance in CRY1 clock 'clamped' cells (see Fig. 4.6 B and C; and Fig. 4.12 B). This is in line with the described PER2 stability profile being most stable at circadian times when CRY1 protein is expected to peak (circadian times 18–24)[78,84,80]. CRY1 protein is known to interact with and stabilize PER2[123,125]. In contrast, the CRY1 mutant was shown in a luciferase complementation assay to no longer interact with PER2 protein[124]. Thus, increased PER2 abundance could be due to an effect of direct CRY1 interaction rather than a 'clamped' circadian clock. Fujimoto *et al.*, however, demonstrated that circadian PER2 protein abundances of a constitutively expressed *Per2* coding sequence are as well found for a PER2 mutant, lacking the CRY1 binding domain[79]. Consequently, the here observed PER2 abundance changes are likely based on its circadian stability and demonstrate the applicability of the comparative 'clamped' clock model.

In conclusion, the results of this study indicate that the fluorescence based GPS reporter method is suitable to investigate global circadian protein stability of the hORFeome library in a comparative clock 'clamped' one–time point approach.

5.2 Characterization of Screen Results

To investigate the circadian 'stabilome', a comparative one–time point approach was performed in cells with an 'arrested' molecular clock representing a snapshot of a circadian phase of high CRY1 level. Observed overall protein abundances changes analyzed by the GPS method were increased, indicating increased protein stability (see Fig. 4.14). This is in accordance with bioinformatic predictions and circadian proteome data. In these studies, protein abundance of the proteome reaches its maximum around circadian times 18–24, the phase of endogenous CRY1 peak in mouse liver[41,84,80] (and S. Lück, AG Westermark, ITB Berlin; personal communication; Robles *et al.*, *Plos Genetics*, in press). In addition, protein level of arginase 1, an enzyme of the urea cycle with described circadian activity, was previously shown to decline at circadian phases when CRY1 protein rises[141,41]. This is in agreement with the CRY1 destabilizing effect on ARG1 observed in this study (see appendix Tab. A.1, position –11).

Together, these findings indicate an overall reliability of the screen results. But are proteins with abundance changes identified in the one–time point approach circadian stable proteins? What might be the biological relevance of circadian stable proteins? These questions will be addressed in the following two subsections.

5.2.1 Towards the Validation of Circadian Stability of CAAPs

In this study, protein stabilities were analyzed by the GPS method at a specific circadian phase. According to the defined selection criteria, 540 CAAPs (**C**RY1 mediated **a**ltered **a**bundant **p**roteins) were defined, representing about 9 % of detected proteins (see Fig. 4.10 C and subsection 4.3.2). To investigate whether CAAPs are proteins with circadian stability, first experimental and computational studies were performed.

First Experimental Validations of CAAPs

Several CAAPs and proteins that showed no strong alteration of their protein abundance were randomly selected for independent validation experiments (see Fig. 4.10 B). In a first approach, to test whether CRY1 overexpression (OX) alters protein abundance as observed in the high–throughput analysis, coding sequences of selected proteins were C–terminally tagged with V5 or luciferase and stably introduced into U–2 OS cells. Analyzed protein abundances showed only trends but no clear difference between CRY1 and CRY1mut OX (data not shown, detected by V5 immunoblotting or luciferase counts). In contrast to the GPS analysis, these methods detect mean protein levels of a cell population. The GPS method, however, is more sensitive and detects protein changes at the single cell level. Furthermore, although being transduced with low viral titers, the expression of coding sequences within validation experiments is probably higher than in library cells. Thus, resultant increased cellular abundances and additional OX of CRY1 or its mutant might interfere with endogenous stoichiometric relationships of components mediating (circadian) degradation.

In a second validation approach, circadian stability of proteins fused to luciferase was monitored online in oscillating cells. As positive controls, CRY1mut and CLOCK were used (see Fig. 4.12; the mutant of CRY1 was used to prevent a 'clamping' of circadian dynamics). Both clock proteins underlie circadian post–translational modifications that are believed to regulate their stability (see introduction 1.3.2). In fact, both luciferase fusion proteins showed circadian abundance profiles after synchronization of cellular clocks by temperature entrainment. This indicates an altered stability along the circadian cycle. Indeed, different stabilities of CRY1mut were determined at two times within the circadian cycle (see Fig. 4.12 D). A clear difference of CLOCK protein stability was not observed, which might be due to the faster damping of its oscillation and/ or inappropriate time points of protein stability measurements.

Analyzing of randomly selected candidates (see Fig. 4.10 B) within the luciferase reporter revealed circadian signals (although with low amplitude rhythms) for selected CAAP luciferase fusions (see Fig. 4.13). Interestingly, circadian oscillations were as well detected for two proteins that at least according to the selection criteria have not been classified as CAAPs. Only NPM1, the protein with lowest difference between CRY1 and CRY1mut OX (according to the euclidean distance, ED), did not show a circadian profile of luciferase counts. Although the minimal viral volume needed for a stable transduction was determined for each protein, mean luciferase counts of CRY1mut and CLOCK fusions were on average 5– to 80–fold below the other tested fusion proteins (data not shown). This could explain the observed low amplitude rhythm as a consequence of too strong expression of exogenous proteins impairing with endogenous processes, as mentioned above. Together, these results may indicate that the selection criteria defining CAAPs could be defined less stringent. First of all, however, the analysis of additional negative controls is recommended and further experiments validating circadian protein half–lifes are needed.

Rhythmically Abundant Proteins are Enriched among CAAPs

In addition to experimental validations, computational analysis of previous circadian large–scale studies might help to validate screen results. Proteins with circadian stability might exhibit rhythmic abundance profiles. Indeed, rhythmic proteins identified in other studies are roughly two–fold enriched among defined CAAPs (see Tab. 4.1 a). This enrichment is probably underestimated as mass spectrometry based proteome screens often miss low abundant proteins that are more likely targets of circadian protein degradation. In addition, the applied selection criteria of CAAPs give an advantage to strong abundance changes (e.g. proteins with very small r.PSIs being stabilized to highest r.PSI values or vice versa). Abundance changes of proteins with medium r.PSI values are underrated and thus are missing among CAAPs.

Altogether, experimental and bioinformatic analysis support that CAAPs are proteins with circadian stability. Nonetheless, for confirmation further experimental evidence is needed.

5.2.2 First Insights on the Biological Relevance of CAAPs

Interestingly, a slight enrichment of rhythmically abundant proteins that are either constitutively or rhythmically transcribed, was found among CAAPs (see Tab. 4.1 b and c). This might indicate an additional regulatory role of rhythmic degradation together with circadian transcriptional, post–transcriptional and/ or translational processes. In addition, CAAPs are probably not involved in the overall regulation of the core clockwork (see sub–subsection in 4.3.5). Together, these findings may indicate a role of CAAPs for the additional fine-tuning of rhythmic clock output processes, probably to sustain robust circadian rhythms of physiology.

To gain insights on the biological functions of CAAPs, Gene Ontology (GO) analysis were performed. Less than 50 % of CAAPs are enriched in GO terms (249 of 540 CAAPs, see Fig. 4.15). In addition, most p–values of GO terms are marginally significant (p–values between 0.01–0.05, see appendix Tab. A.5). This is likely not due to an inappropriate CAAPs selection, as overall p–values of GO terms did not decreased when parameters defining CAAPs were changed (data not shown). Moreover, these findings may indicate an equal distribution of CAAPs over different groups of proteins. In fact, single key regulators of divers processes will not be enrichment within a GO analysis. This is in line with the idea that circadian protein stability is important for diverse rhythmic physiological processes presumably via the control of rate limiting enzymes or hub proteins, as shown for p53[102].

Nonetheless, although not highly significant, identified GO terms are associated to interesting cellular mechanisms. Independent circadian proteome studies revealed an overrepresentation of rhythmically abundant proteins in processes related to vesicular trafficking[42,43] (and Robles *et al.*, *Plos Genetics*, in press). The enrichment of CAAPs involved in vesicular transport and secretion strengthens the importance of circadian protein level within these processes – solely or additionally mediated by rhythmic degradation (see Fig. 4.15). Indeed, pharmacological impairment of endo– and exocytosis was shown to dampen circadian oscillations in organotypic slices of the SCN, the master pacemaker of the mammalian clock[43]. Interestingly, knockout mouse models lacking different secretory vesicles were not sensitive towards light induced phase resetting or showed impaired circadian gene expression in SCN and peripheral tissues, although photoreception and clock oscillatory function were not disrupted[142,143]. In addition, circadian dynamics monitored by a clock gene promoter fused to luciferase (*Bmal1*–promoter luc) are disturbed in immortalized cells upon knockdown of genes involved in vesicular transport (internal laboratory data, provided by S. Jäschke and B. Maier). Together, this highlights the importance of vesicular trafficking for circadian rhythms. The role of rhythmic abundance and/

or stability of involved proteins, however, is unknown so far. Interestingly, HSPA8, an ATPase that has been linked to the disassembly of clathrin–coated vesicles[144], is a predicted CRY1 interactor and was found within this study as a destabilized CAAP (see appendix Tab. A.1, position -148 and A.4). HSPA8 thus represents an interesting candidate with potential circadian stability that might connect the molecular clock to vesicular trafficking.

Furthermore, processes related to mitosis represent the group where most CAAPs of analyzed proteins were enriched. Interestingly, cellular cycles of the circadian clock and cell division have been shown to be connected and may explain the role of deficient clocks in the susceptibility to cancer development[145–149]. While there are conflicting reports about gating of the cell division cycle by the circadian clock in immortalized fibroblasts[150,151], a clear role of the mammalian circadian clock to time the cell division cycle has been demonstrated *in vivo*[152,153]. In addition, previous studies showed that cell cycle related genes are circadian transcribed, resulting in rhythmically abundant active kinases (f.e. *cyclin B1*, *cdc2* and *wee1*)[152,145,154]. Furthermore, direct interactions of clock and clock-associated proteins to regulatory cell cycle components has been demonstrated[155,146,153]. The importance of rhythmically abundant and/ or stable proteins within cell cycle and mitosis itself has not been analyzed so far. Unexpectedly, another predicted CRY1 interactor among CAAPs, PPP2R1B[117], a regulatory subunit of the serine/threonine phosphatase 2A, is known to negatively control entry into mitosis[156]. In addition PP2A activity is crucial for circadian dynamics as revealed in *Drosophila melanogaster* and *Neurospora crassa*.

From a comprehensive perspective, vesicular trafficking and mitosis share a common feature. A transporting or secreting vesicle and a dividing cell are largely uncoupled from transcriptional and/ or translational events[157]. Cellular information, like the circadian time, can only be conveyed via proteins. Thus, circadian protein stability mediated by post translational modifications may represent a so far unrecognized mammalian timer[158].

In conclusion, further studies are needed to investigate the importance of circadian protein stability for biological processes. Rhythmic degradation probably represents a further mode of circadian regulation in addition to transcriptional, post–transcriptional and translational modes. Thereby, circadian protein stability may add to the robustness of circadian rhythms, without modulating them.

5.3 Perspectives

This study performed a proteome–wide high–throughput screen to identify proteins with circadian stability. A cell model where the circadian clock is 'clamped' at a specific circadian phase by CRY1 overexpression was applied. About 540 CAAPs whose protein abundance, thus stability is altered were identified. Interesting candidates were revealed that opened new questions and hypotheses. Nonetheless, first of all, slight improvements of the microarray analysis could be addressed and more experimental evidence should be accumulated to confirm that the screen results represent proteins with circadian stability.

Improvement of Microarray Data Evaluation

Selection parameters defining CAAPs (ED and Δr.PSI) underestimate the amount of proteins with altered stability. As described, the selection favors extreme changes of protein stability (e.g. very stable to very unstable and vice versa) whereas proteins with medium r.PSI values cannot reach these extreme parameters. Thus, both ED and Δr.PSI should be used relative to the initial r.PSI value. Thereby, additional proteins will be added to the list of CAAPs.

Furthermore, the GPS method could be extended to estimate protein half–lifes instead of PSI values. To this end, comparisons of protein half–lifes to ratios of EGFP/DsRED, as done in Fig. 4.3 C and appendix Fig. A.2, need to be expanded for protein with various protein half–lifes. Based on the resultant correlation fit of protein half–lifes versus EGFP/DsRED ratios, r.PSI values could be specified as 'estimated half–lifes' and Δr.PSI values as 'expected half–life differences'.

Additional Circadian Phase Specific Models of 'Clamped' Clocks

To strengthen that CAAPs represent circadian stable proteins further experiments are needed. First of all, to confirm the identified 540 CAAPs, additional cellular 'clamped' clock models could to be applied. CRY1 overexpression 'clamped' the molecular clock in the phase of the negative feedback with low E–box driven gene expression (see Fig. 4.6 A and Fig. 4.7). An 'arrested' molecular clock in anti-phase to CRY1 overexpression (e.g. *Cry1*/*Cry2* $^{-/-}$ cells) should reveal CAAPs with opposite protein abundance changes than observed by CRY1 overexpression. Secondly, a complete picture of circadian protein stability can only be obtained if 'clamped' models around the clock would be applied. Thereby, proteins would be identified that have their peak and trough of stability at intermediate phases to the CRY1 'arrested' phase.

Overexpression of components of the positive and negative arm already indicated that the negative feedback loop seems to be dominant over the activating one (see appendix Fig. A.4). Within these experiments, CLOCK and BMAL1 were both separately expressed. Acting as a heterodimer *in vivo*, both components are presumably rate–limiting and thus should be overexpressed simultaneously. Furthermore, overexpression of NR1D1 indicated an up–regulation of E–box driven genes (see appendix Fig. A.4). Increased dose of NR1D1 could be enough to 'arrest' the molecular clock in this phase. Moreover, in addition to the induction of the positive arm, the negative feedback needs to be perturbed. The combination of overexpressing positive regulators while down regulating negative clock genes as *Per*s or *Cry*s should 'clamp' the oscillator in the positive, activating phase.

Validation of Circadian Protein Stability

Additional experiments analyzing rhythmic stability of individual CAAPs with independent techniques are needed. The applied luciferase fusion reporter seems to be an appropriate tool to monitor circadian protein stability online, however needs further adaptation. For a (circadian) degradation, mediated by cellular components, the amount of ectopically expressed proteins should be close to the physiological range. This can either be achieved by the use of a less strong promoter than the applied cytomegalovirus promoter. Furthermore, selection and analysis of single clones with low overall luciferase counts could be done.

For an individual validation, the use of recombinant proteins is conceivable. Thereby, protein decay of recombinant proteins transduced into cells with low doses at different circadian times could be analyzed.

Moreover, to circumvent limitations based on ectopic expression, validation of circadian protein stability with endogenous proteins would be preferable. Still, limitations like blocking overall translation, the specificity of antibodies detecting endogenous protein and low sensitivity of the detection method are to be considered.

A Appendix

A.1 Dot Plots of Flow Cytometry data

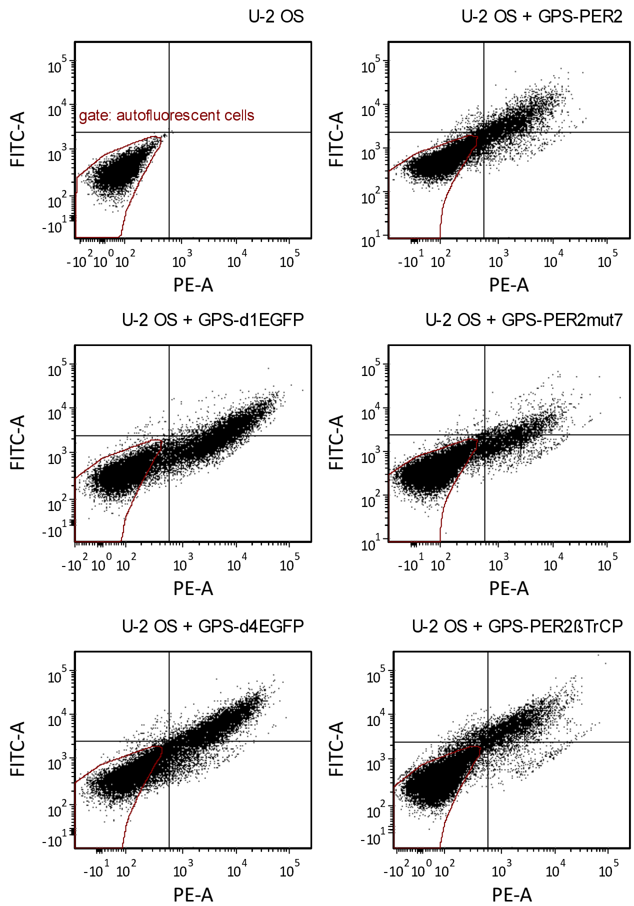


Figure A.1: Flow cytometry raw data of U–2 OS cells harboring the fluorescent reporter. Representative dot plots of flow cytometry analysis of U–2 OS cells harboring either a destabilized EGFP variant (d1/d4EGFP) or the core clock protein PERIOD2 (PER2) or a destabilizing (PER2mut7) or stabilizing (PER2βTrCP) mutant (data to Fig. 4.3 A). Cells are plotted according to their fluorescence in the green channel (FITC-A) on the y–axis versus the red channel (PE-A) on the x–axis. The red shape gates autofluorescent U–2 OS cells. According to the stability of the fused ORF green fluorescent intensity varies when comparing d1- and d4EGFP, as well as PER2 and its mutants. A–area (stands for whole cell area analyzed in flow cytometry), FITC-A–Fluorescein isothiocyanate, GPS–Global Protein Stability, PE–Phycoerythrin.

A.2 Correlation of GPS Results with Protein Half–Lifes

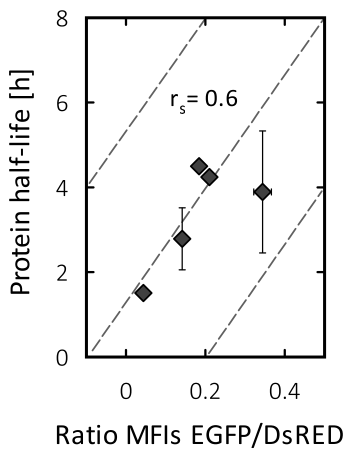


Figure A.2: Correlation of GPS results with protein half–lifes. Correlation plot of EGFP/DsRED ratios (x–axis) versus protein half–lifes obtained from C–terminally tagged luciferase fusion proteins (y–axis). Decay of luciferase activity after pharmacological block of protein synthesis relative to a solvent control was determined over six hours. Protein half–lifes were calculated from fitted exponential decay functions ($P < 0.5$, Spearman Rank Correlation test, $n = 5$). Given are mean values $\pm$ SD. For MFIs the mean of three and for protein half–lifes the mean of two measurements with each two replicates are given.

A.3 Circadian GPS Analysis

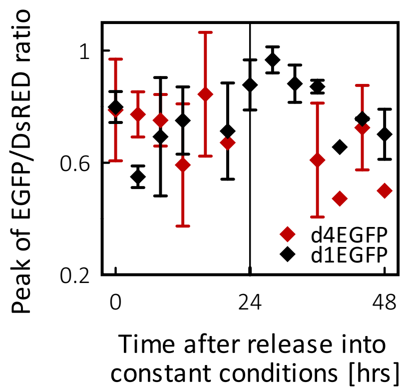
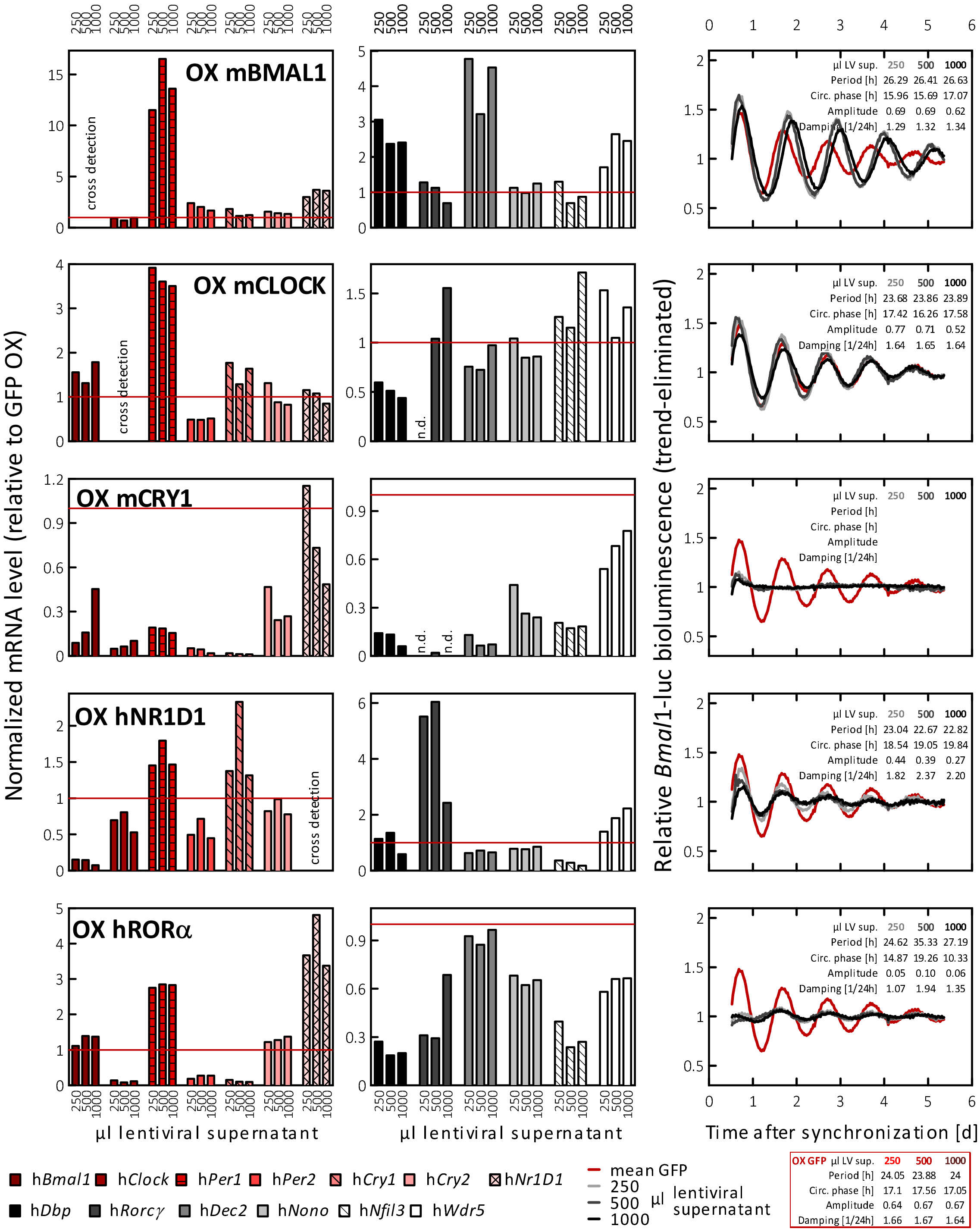


Figure A.3: Circadian GPS Analysis. U–2 OS cells stably expressing the GPS reporter either with d1EGFP (black color) or d4EGFP (red color) were synchronized by temperature entrainment (for details on temperature entrainment see subsection 4.3.4). After release into constant conditions, cells were harvested in 4–hrs intervals over two circadian cycles. Cells were fix with paraformaldehyde for a subsequent flow cytometry analysis. Given is one exemplary experiment over two circadian cycles of additional two experiments over one circadian cycle. Values where the EGFP/DsRED ratio distributions peaks were determined and set relative to the maximum value of the time course. Depicted are mean values of two replicates. Error bars represent the maximum and minimum value. *Note:* for d4EGFP three time points of the second day were not evaluable.

A.4 Overexpression Analysis to 'Arrest' the Molecular Clock

Figure A.4 *(following page)*: Overexpression of clock proteins to 'arrest' the molecular clock. Stable dose–dependent overexpression (OX) of lentivirally delivered core clock components BMAL1, CLOCK, CRY1, NR1D1 and RORα or a GFP control in U–2 OS cells. *(Left and middle panels)* Expression of core clock genes (*left panels:* h*Bmal1*, h*Clock*, h*Per1*, h*Per2*, h*Cry1*, h*Cry2*, h*Nr1d1*; for red color code see legend) and clock–controlled genes (*middle panels:* h*Dbp*, h*Ror*γ, h*Dec2*, h*Nono*, h*Nfil3*, h*Wdr5*; for gray color code see legend) were analyzed after seven days in constant conditions. Depicted are h*Gapdh*–normalized data relative to the corresponding level of GFP OX in cells (n = 3 technical replicates per OX dose). Horizontal red lines indicate where gene expression compared to GFP control is unchanged. Results, in which the qRT–PCR primer cross-reacted with the overexpressed ORF are not shown (marked as 'cross detection'). Data of hNR1D1 OX were published in [137]. *(Right panels)* Bioluminescence recordings of the *Bmal1*–promoter luciferase reporter (stably integrated clock gene promoter fused to luciferase) over six days after synchronization. The red curve represents the mean bioluminescence of all GFP OX doses. Depicted are 24–hrs running average trend–eliminated data (n = 4 technical replicates per OX dose). Circadian parameter are depicted in the upper right corner of each plot, for OX of GFP on the bottom right (red framed box). OX of one candidate, CRY1 showed strong repressive action on circadian dynamics and was independently confirmed (see Fig. 4.7). LV–lentiviral, n.d.–not detected, OX–overexpression.



A.5 Establishment of Nested PCR Protocol for Quantitative Amplification

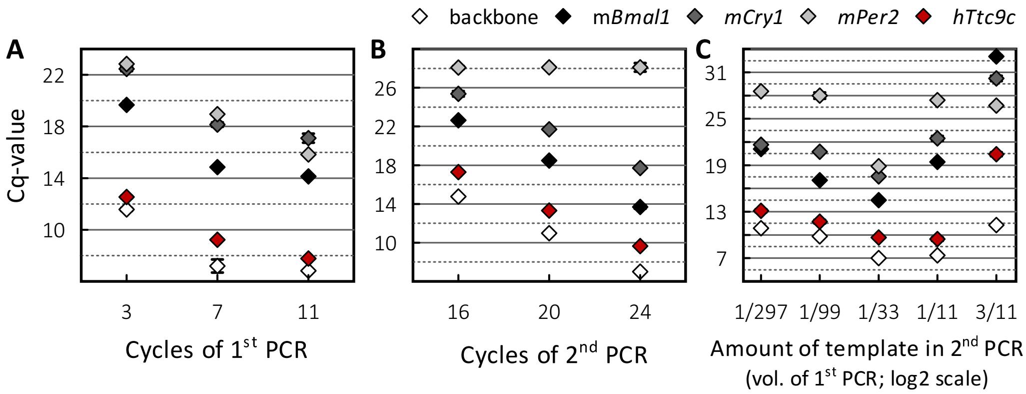


Figure A.5: Optimization of nested PCR protocol. Optimization steps of the nested PCR protocol were done to guarantee for a quantitative amplification of ORFs in the library. Depicted are Cq–values of indicated targets obtained by qRT–PCR after nested PCR at the specified settings. Nested PCR was done using 2 μg genomic DNA of a cell mixture consisting of 16,000 reporter library cells, 9 copies of m*Bmal1*, 27 copies of m*Cry1*, 3 copies of m*Per2* and 81 cell copies of h*Ttc9c* within the GPS reporter construct (mixture B in Fig. 4.8 A). (A) Test for sufficient cycle number of the 1^{st} PCR. 2^{nd} PCR was performed with 24 cycles and $1/33$ volume of the first PCR as template (corresponds to 34.8 μl). Between three and seven cycles linear amplification was observed. (B) Test for number of cycles in 2^{nd} PCR. 1^{st} PCR was performed with seven cycles and $1/33$ template was used in the second PCR. A linear amplification was observed in all tested conditions. Only m*Per2* was not linearly amplified which might be due to the low copy number in the mixture. (C) Test for amount of template used in the 2^{nd} PCR. For the 1^{st} and 2^{nd} PCR seven and 24 cycles were performed, respectively. Up to $1/33$ template of the first PCR resulted in a linear amplification as expected. Based on this, the final settings in the nested PCR protocol were set to five cycles 1^{st} PCR followed by 20 cycles of the 2^{nd}. Due to the low amount of cycles chosen, $1/33$ of the first PCR was used as template for the second amplification. These settings were tested for a quantitative amplification of ORFs in a test setup as shown in Fig. 4.8 B. Depicted are mean values $\pm$ SD of three technical replicates of the qRT–PCR after nested PCR. Vol.–volume.

A.6 Microarray Pre–Processing

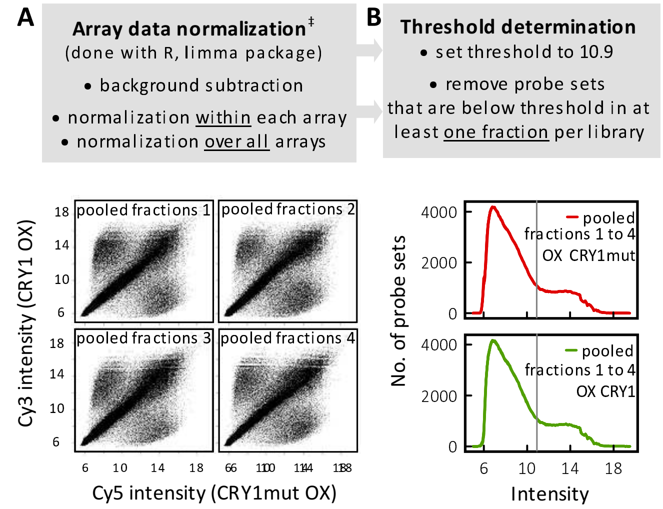


Figure A.6: Microarray raw data. Content of the gray boxes describes the steps of microarray data pre–processing. (A) Dot plots depict background corrected and normalized (within and over all arrays) intensities of competitive hybridized Cy5 labeled samples (library with OX of CRY1mut, x–axis) versus Cy3 labeled samples (library with OX of CRY1, y–axis) for pooled fractions one to four. ‡ Microarray data normalization was done in cooperation with K. Jürchott. (B) Histograms show the distribution of intensities for all four fractions of samples overexpressing CRY1mut (upper panel, red data) or CRY1 (lower panel, green data). Based on intensities of denoted library ORFs, threshold was set to 10.9 for all samples (gray vertical lines, see methods sub–subsection in 3.4.1). R–statistical software program.

A.7 Spike–in Controls of Microarray Hybridization

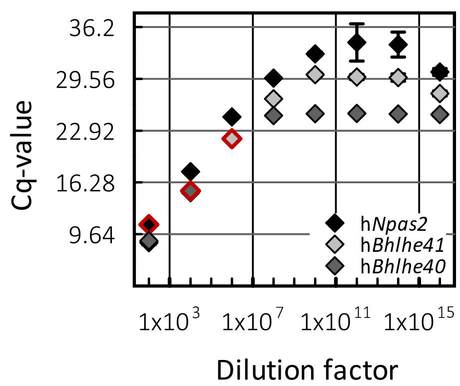
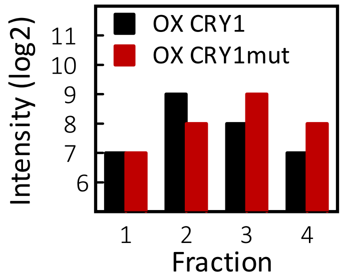


Figure A.7: Dilution series of microarray spike–in controls. In order to control for effects of sample labeling and hybridization, spike–in controls were added after the nested PCR to each sample. To this end, PCR products of full length coding sequences of h*Bhlhe40*, h*Bhlhe41* and h*Npas2* were added in different concentrations. PCR products were diluted in serial 100–fold steps. Depicted are Cq–values of qRT–PCR analysis of their dilutions. Red framed symbols indicate the dilution for each PCR product that was used as spike–in control in the screen. Given are representative results with mean values ± SD of three technical replicates. Y–axis is spaced in expected Cq–value differences of a 100–fold dilution series.

A.8 Calculation of Δr.PSI (Protein Stability Index) and ED (Euclidean Distance)



	intensities (log2)			
fraction	1	2	3	4
OX CRY1	7	9	8	7
OX CRY1mut	7	8	9	8

Calculation of Δr.PSI

$$\mathrm{r.PSI} = \sum_{i=1}^{4}\left(\frac{2^{int.frac.i}}{\sum_{i=1}^{4} 2^{int.frac.i}}\right)\times i \qquad \Delta\mathrm{r.PSI} = \mathrm{r.PSI}_{CRY1} - \mathrm{r.PSI}_{CRY1mut}$$

$$\mathrm{r.PSI}_{CRY1} = \left(\frac{2^7}{2^7+2^9+2^8+2^7}\right)\times 1 + \left(\frac{2^9}{2^7+2^9+2^8+2^7}\right)\times 2 + \left(\frac{2^8}{2^7+2^9+2^8+2^7}\right)\times 3 + \left(\frac{2^7}{2^7+2^9+2^8+2^7}\right)\times 4 = 2.38$$

$$\mathrm{r.PSI}_{CRY1mut} = \left(\frac{2^7}{2^7+2^8+2^9+2^8}\right)\times 1 + \left(\frac{2^8}{2^7+2^8+2^9+2^8}\right)\times 2 + \left(\frac{2^9}{2^7+2^8+2^9+2^8}\right)\times 3 + \left(\frac{2^8}{2^7+2^8+2^9+2^8}\right)\times 4 = 2.78$$

$$\Delta\mathrm{r.PSI} = 2.38 - 2.78 = -0.4$$

Calculation of ED

$$\mathrm{cent.int.}\,i = int.frac.i - \frac{\sum_{i=1}^{4} int.frac.i}{4} \qquad \mathrm{ED} = \sqrt{\sum_{i=1}^{4}(\mathrm{cent.int.}\,i_{CRY1mut} - \mathrm{cent.int.}\,i_{CRY1})^2}$$

$$\mathrm{cent.int.}\,1_{CRY1} = 7 - \left(\frac{7+9+8+7}{4}\right) = -0.75 \qquad \mathrm{cent.int.}\,1_{CRY1mut} = 7 - \left(\frac{7+8+9+8}{4}\right) = -1$$

$$\mathrm{cent.int.}\,2_{CRY1} = 9 - \left(\frac{7+9+8+7}{4}\right) = 1.25 \qquad \mathrm{cent.int.}\,2_{CRY1mut} = 8 - \left(\frac{7+8+9+8}{4}\right) = 0$$

$$\mathrm{cent.int.}\,3_{CRY1} = 8 - \left(\frac{7+9+8+7}{4}\right) = 0.25 \qquad \mathrm{cent.int.}\,3_{CRY1mut} = 9 - \left(\frac{7+8+9+8}{4}\right) = 1$$

$$\mathrm{cent.int.}\,4_{CRY1} = 7 - \left(\frac{7+9+8+7}{4}\right) = -0.75 \qquad \mathrm{cent.int.}\,4_{CRY1mut} = 8 - \left(\frac{7+8+9+8}{4}\right) = 0$$

$$\mathrm{ED} = \sqrt{(-1+0.75)^2 + (0-1.25)^2 + (1-0.25)^2 + (0+0.75)^2} = 1.66$$

Figure A.8: Example Calculations of Δr.PSI and ED. Δr.PSI and Euclidean Distance (ED) are calculated for the depicted example of intensity profiles according to the formulas in the gray boxes (see Fig. 4.10 A). Δ–difference, cent.–centered, frac.–fraction, i–counter for fraction number, int.–intensity.

A.9 Selected CAAPs

The following table lists selected CAAPs (**C**RY1 mediated **a**ltered **a**bundant **p**roteins), ranked from destabilized to stabilized. Given z–scores belong to Δr.PSI values. *Selected probe sets that fulfilled thresholds of detection, z–score$_{\Delta r.PSI}$ and Euclidean Distance (ED). For details see results subsection 4.3.2. Sel.–selected, pres.–present.

Table A.1: Selected CAAPs

Position	gs	No. of sel. ps*	No. of ps pres.	Directed ED	Mean ED	SD ED	Mean Δr.PSI	SD Δr.PSI	Mean z–score	SD z–score
-1	IL18RAP	8	10	-12.489	12.489	1.289	-1.562	0.102	-2.437	0.147
-2	HSPA14	8	13	-12.198	12.198	1.290	-1.065	0.049	-1.721	0.071
-3	TACR3	4	5	-12.064	12.064	0.410	-1.592	0.054	-2.480	0.077
-4	FAM172A	9	11	-11.565	11.565	1.271	-1.514	0.078	-2.368	0.112
-5	FSIP1	10	11	-11.001	11.001	2.863	-1.404	0.121	-2.210	0.173
-6	ELOVL5	5	7	-10.936	10.936	1.787	-2.097	0.200	-3.207	0.288
-7	VNN1	4	7	-10.887	10.887	0.987	-1.203	0.037	-1.920	0.053
-8	TSPAN12	5	7	-10.887	10.887	0.529	-2.141	0.082	-3.270	0.118
-9	PAH	10	12	-10.815	10.815	1.095	-1.412	0.061	-2.221	0.088
-10	CCNJ	3	5	-10.764	10.764	0.298	-1.174	0.015	-1.878	0.022
-11	ARG1	8	8	-10.667	10.667	2.060	-1.631	0.120	-2.536	0.173
-12	ST6GAL1	4	5	-10.553	10.553	2.751	-1.579	0.223	-2.462	0.320
-13	SLC39A9	5	7	-10.490	10.490	1.069	-1.718	0.045	-2.662	0.064
-14	TRAF5	7	9	-10.462	10.462	0.939	-1.516	0.044	-2.371	0.064
-15	C19orf44	4	7	-10.385	10.385	1.152	-1.195	0.038	-1.909	0.055
-16	MAPKAPK3	6	9	-10.346	10.346	1.904	-1.850	0.179	-2.851	0.257
-17	FNIP1	11	18	-10.296	10.296	1.495	-1.419	0.148	-2.231	0.213
-18	C3orf17	5	9	-10.266	10.266	0.392	-0.957	0.010	-1.566	0.015
-19	B4GALT4	5	6	-10.207	10.207	0.957	-1.338	0.129	-2.114	0.185
-20	HNMT	4	6	-10.183	10.183	0.975	-1.247	0.021	-1.984	0.030
-21	CPSF6	7	9	-10.178	10.178	1.007	-1.303	0.043	-2.064	0.062
-22	GPC5	6	8	-10.100	10.100	1.304	-2.322	0.147	-3.531	0.212
-23	SUPV3L1	13	15	-10.096	10.096	1.015	-1.463	0.090	-2.295	0.130
-24	HLX	3	4	-10.074	10.074	0.918	-1.660	0.050	-2.578	0.072
-25	SLC1A1	8	11	-10.053	10.053	1.370	-0.975	0.021	-1.593	0.031
-26	CAPZA2	6	8	-10.053	10.053	0.763	-1.798	0.119	-2.777	0.171
-27	BTK	13	17	-9.948	9.948	0.862	-0.963	0.018	-1.574	0.025
-28	FGF14	4	6	-9.931	9.931	1.077	-1.944	0.014	-2.987	0.020
-29	XRN2	13	25	-9.924	9.924	1.288	-1.641	0.107	-2.551	0.155
-30	METAP1D	5	7	-9.886	9.886	1.226	-1.055	0.042	-1.707	0.060
-31	PIGK	6	10	-9.856	9.856	0.564	-1.542	0.122	-2.408	0.175
-32	NSMCE1	6	7	-9.755	9.755	1.406	-1.224	0.120	-1.951	0.173
-33	RBBP7	8	11	-9.753	9.753	1.518	-1.826	0.090	-2.816	0.130
-34	AGFG1	8	14	-9.689	9.689	1.293	-1.464	0.039	-2.295	0.055
-35	RFC4	7	9	-9.628	9.628	0.370	-0.963	0.038	-1.575	0.054
-36	FANCC	11	13	-9.615	9.615	2.157	-1.368	0.128	-2.158	0.184
-37	C3orf15	13	16	-9.584	9.584	0.994	-1.343	0.188	-2.122	0.270
-38	KIAA1467	11	13	-9.549	9.549	2.244	-2.000	0.248	-3.067	0.357
-39	NIM1	2	3	-9.549	9.549	1.425	-1.106	0.113	-1.781	0.163
-40	CRBN	9	11	-9.547	9.547	1.168	-1.532	0.135	-2.394	0.194
-41	TGM4	13	14	-9.532	9.532	1.699	-1.454	0.151	-2.281	0.218
-42	CTNND1	12	18	-9.512	9.512	1.250	-1.122	0.032	-1.803	0.046
-43	PKN2	17	22	-9.486	9.486	1.114	-1.337	0.210	-2.113	0.302
-44	MAP2	9	14	-9.479	9.479	1.635	-2.109	0.180	-3.224	0.259
-45	ZNF215	3	5	-9.466	9.466	1.092	-0.975	0.012	-1.592	0.018
-46	HMG20A	5	7	-9.465	9.465	1.265	-1.388	0.079	-2.187	0.113
-47	ENPP5	2	3	-9.430	9.430	0.337	-0.971	0.011	-1.586	0.016
-48	LHX8	6	9	-9.428	9.428	1.239	-1.862	0.122	-2.869	0.175
-49	SCCPDH	8	10	-9.427	9.427	0.839	-1.076	0.056	-1.738	0.080
-50	ZNF192	4	5	-9.415	9.415	0.941	-2.460	0.119	-3.729	0.171
-51	C11orf16	3	5	-9.394	9.394	0.686	-1.255	0.067	-1.996	0.097
-52	DYNC1I2	10	16	-9.314	9.314	1.101	-1.426	0.213	-2.241	0.307
-53	RTN4IP1	7	8	-9.304	9.304	0.853	-2.222	0.162	-3.387	0.232
-54	LARGE	9	14	-9.300	9.300	1.776	-1.025	0.067	-1.664	0.097
-55	GDF9	2	2	-9.248	9.248	0.832	-1.187	0.013	-1.897	0.019
-56	ELOVL7	4	6	-9.240	9.240	0.879	-1.094	0.020	-1.764	0.029
-57	MCMBP	10	15	-9.191	9.191	0.695	-1.667	0.111	-2.588	0.160
-58	SSX2IP	10	13	-9.184	9.184	0.738	-1.628	0.063	-2.531	0.090
-59	CHEK2	8	11	-9.179	9.179	0.871	-0.969	0.019	-1.584	0.028
-60	SGMS1	4	5	-9.160	9.160	2.917	-1.160	0.093	-1.859	0.133
-61	CFHR5	7	9	-9.157	9.157	1.114	-1.321	0.191	-2.090	0.275
-62	TBC1D7	5	7	-9.146	9.146	1.369	-1.294	0.060	-2.051	0.086
-63	ST3GAL3	6	9	-9.141	9.141	1.919	-1.041	0.048	-1.687	0.069
-64	MLH1	17	18	-9.110	9.110	1.255	-1.541	0.245	-2.407	0.352
-65	MFN2	12	17	-9.090	9.090	1.436	-1.360	0.138	-2.147	0.199
-66	CPVL	10	11	-9.085	9.085	1.073	-1.507	0.251	-2.357	0.361
-67	ADAD1	11	11	-9.084	9.084	1.511	-1.618	0.266	-2.518	0.383
-68	NASP	9	12	-9.004	9.004	1.016	-1.472	0.054	-2.307	0.078
-69	WBSCR17	8	11	-8.991	8.991	0.953	-0.986	0.046	-1.608	0.066
-70	PIK3R3	9	10	-8.985	8.985	1.859	-1.913	0.232	-2.942	0.334

Position	gs	No. of sel. ps*	No. of ps pres.	Directed ED	Mean ED	SD ED	Mean Δr.PSI	SD Δr.PSI	Mean z-score	SD z-score
-71	C12orf42	4	5	-8.974	8.974	1.931	-1.234	0.184	-1.965	0.264
-72	KLK13	3	5	-8.971	8.971	1.819	-1.154	0.051	-1.850	0.074
-73	DDX17	8	14	-8.969	8.969	0.983	-0.995	0.047	-1.621	0.067
-74	GRAMD3	10	14	-8.953	8.953	1.233	-1.380	0.197	-2.175	0.283
-75	CTSG	3	5	-8.942	8.942	1.542	-1.146	0.032	-1.838	0.046
-76	YTHDC1	11	14	-8.860	8.860	1.209	-2.606	0.254	-3.939	0.365
-77	PDZK1	6	8	-8.855	8.855	0.611	-1.701	0.137	-2.636	0.197
-78	CD44	11	17	-8.846	8.846	0.999	-0.958	0.024	-1.568	0.034
-79	RNF175	6	8	-8.831	8.831	1.084	-1.862	0.081	-2.869	0.116
-80	RSL1D1	6	9	-8.830	8.830	0.717	-1.577	0.048	-2.458	0.069
-81	PGM2L1	12	14	-8.796	8.796	1.460	-1.432	0.092	-2.250	0.133
-82	TGFB1	5	7	-8.733	8.733	2.506	-1.316	0.112	-2.083	0.161
-83	ARFGAP3	14	15	-8.729	8.729	2.110	-1.686	0.211	-2.615	0.304
-84	SGCE	6	9	-8.727	8.727	0.913	-0.960	0.016	-1.571	0.024
-85	C3orf37	5	6	-8.704	8.704	1.371	-1.183	0.071	-1.891	0.101
-86	FOXN2	3	5	-8.666	8.666	0.351	-1.013	0.017	-1.647	0.024
-87	BCAP29	6	9	-8.616	8.616	1.326	-1.069	0.066	-1.727	0.094
-88	VPS52	16	18	-8.615	8.615	1.805	-1.552	0.159	-2.422	0.228
-89	TTLL1	7	12	-8.594	8.594	1.555	-1.124	0.043	-1.807	0.062
-90	ANKRD40	3	5	-8.573	8.573	0.128	-1.136	0.011	-1.824	0.016
-91	DCP1A	8	9	-8.552	8.552	2.195	-1.765	0.285	-2.729	0.411
-92	SRPR	9	14	-8.547	8.547	1.409	-1.466	0.240	-2.299	0.346
-93	DAB2	8	11	-8.547	8.547	1.090	-2.115	0.152	-3.232	0.219
-94	SLC38A4	13	14	-8.505	8.505	1.360	-1.845	0.141	-2.844	0.203
-95	CFHR1	2	3	-8.487	8.487	0.697	-1.003	0.023	-1.632	0.033
-96	PREPL	12	14	-8.447	8.447	0.945	-1.836	0.069	-2.831	0.100
-97	QRFPR	6	6	-8.423	8.423	2.515	-1.199	0.073	-1.915	0.106
-98	PI4KB	7	11	-8.418	8.418	1.086	-2.452	0.203	-3.718	0.292
-99	SEC63	11	19	-8.395	8.395	0.547	-0.951	0.016	-1.557	0.024
-100	TFDP2	7	9	-8.387	8.387	1.044	-2.626	0.162	-3.968	0.233
-101	ODF2	13	24	-8.342	8.342	1.311	-2.101	0.235	-3.213	0.338
-102	GMCL1	12	14	-8.306	8.306	0.624	-1.250	0.117	-1.987	0.168
-103	PIH1D1	7	9	-8.278	8.278	1.098	-1.476	0.162	-2.313	0.233
-104	IL20	4	5	-8.260	8.260	1.102	-1.111	0.070	-1.788	0.100
-105	CHMP5	4	7	-8.238	8.238	0.807	-1.138	0.158	-1.827	0.227
-106	CERS2	5	8	-8.230	8.230	1.881	-1.414	0.242	-2.224	0.348
-107	ATP6V0A4	15	20	-8.206	8.206	0.929	-0.951	0.021	-1.558	0.031
-108	P2RX7	8	13	-8.180	8.180	0.761	-0.960	0.027	-1.570	0.038
-109	NDST2	11	13	-8.148	8.148	0.969	-1.860	0.115	-2.866	0.166
-110	TMOD2	7	9	-8.122	8.122	1.238	-1.560	0.274	-2.434	0.395
-111	TFG	4	7	-8.118	8.118	2.019	-1.422	0.231	-2.235	0.332
-112	EIF3D	7	13	-8.078	8.078	1.039	-0.950	0.023	-1.556	0.033
-113	PDHA1	6	11	-8.067	8.067	0.697	-1.038	0.068	-1.683	0.098
-114	STAP1	8	8	-8.057	8.057	1.225	-1.330	0.141	-2.103	0.203
-115	RPRD1A	4	6	-8.032	8.032	1.494	-1.521	0.286	-2.377	0.412
-116	NUP43	6	8	-8.029	8.029	0.813	-1.028	0.042	-1.668	0.061
-117	SNAP91	15	26	-8.022	8.022	1.911	-1.280	0.115	-2.032	0.165
-118	SDC2	3	5	-8.004	8.004	0.571	-0.949	0.009	-1.554	0.013
-119	C17orf101	5	8	-7.995	7.995	1.945	-1.110	0.105	-1.786	0.151
-120	DCN	6	7	-7.991	7.991	1.307	-1.005	0.072	-1.635	0.103
-121	RPS6KA3	16	20	-7.961	7.961	1.021	-1.716	0.113	-2.659	0.163
-122	S100PBP	3	5	-7.956	7.956	0.703	-0.951	0.015	-1.557	0.022
-123	RNF32	4	7	-7.942	7.942	1.982	-1.267	0.133	-2.012	0.192
-124	TM4SF19	3	4	-7.882	7.882	1.530	-1.345	0.145	-2.125	0.209
-125	LPCAT2	8	11	-7.857	7.857	1.827	-1.709	0.232	-2.648	0.334
-126	PLK1S1	7	11	-7.842	7.842	0.786	-1.017	0.059	-1.653	0.085
-127	STRAP	7	10	-7.826	7.826	1.208	-1.268	0.067	-2.014	0.097
-128	IRF8	6	7	-7.808	7.808	1.776	-1.433	0.211	-2.251	0.303
-129	RPIA	5	8	-7.792	7.792	0.785	-1.070	0.097	-1.729	0.140
-130	ALOX15B	8	14	-7.792	7.792	1.182	-1.765	0.176	-2.729	0.253
-131	CDR2	3	5	-7.786	7.786	1.424	-1.695	0.120	-2.629	0.173
-132	NAT10	19	27	-7.759	7.759	0.956	-1.473	0.199	-2.309	0.286
-133	CCDC99	6	11	-7.718	7.718	0.643	-0.935	0.016	-1.535	0.024
-134	SMOC1	8	10	-7.708	7.708	2.037	-1.479	0.241	-2.317	0.346
-135	PLA2G2D	2	3	-7.698	7.698	0.197	-1.155	0.016	-1.851	0.024
-136	SYTL5	13	16	-7.691	7.691	1.742	-1.719	0.293	-2.662	0.422
-137	ZCCHC7	4	6	-7.684	7.684	0.570	-0.960	0.022	-1.570	0.032
-138	CCDC47	7	11	-7.670	7.670	0.507	-0.941	0.016	-1.544	0.023
-139	SLC24A6	7	13	-7.644	7.644	0.888	-1.301	0.205	-2.061	0.295
-140	RNF2	5	6	-7.636	7.636	0.705	-1.117	0.054	-1.796	0.078
-141	KPTN	8	11	-7.605	7.605	1.060	-1.042	0.056	-1.689	0.081
-142	IDE	19	25	-7.583	7.583	1.401	-1.897	0.216	-2.919	0.311
-143	LMOD3	2	3	-7.574	7.574	1.929	-2.295	0.479	-3.492	0.690
-144	HENMT1	5	7	-7.541	7.541	1.495	-1.758	0.208	-2.718	0.299
-145	GTPBP1	9	12	-7.518	7.518	1.032	-2.435	0.243	-3.693	0.349
-146	ASH2L	13	17	-7.496	7.496	1.054	-1.359	0.157	-2.145	0.225
-147	C15orf57	2	3	-7.465	7.465	0.456	-1.061	0.065	-1.716	0.094
-148	HSPA8	8	10	-7.453	7.453	1.490	-1.357	0.134	-2.142	0.193
-149	PIP5K1B	8	12	-7.387	7.387	1.103	-2.348	0.270	-3.569	0.388
-150	ADAM32	12	20	-7.379	7.379	0.607	-2.111	0.208	-3.227	0.299
-151	IRF3	7	7	-7.311	7.311	1.268	-1.933	0.167	-2.971	0.240
-152	TARDBP	4	5	-7.306	7.306	0.170	-1.492	0.094	-2.336	0.135
-153	PMEL	6	10	-7.269	7.269	2.493	-1.214	0.125	-1.936	0.180
-154	GCDH	8	11	-7.250	7.250	1.285	-1.165	0.092	-1.865	0.132
-155	SLC7A8	7	12	-7.240	7.240	0.514	-1.115	0.030	-1.793	0.043

A.9 Selected CAAPs

Position	gs	No. of sel. ps*	No. of ps pres.	Directed ED	Mean ED	SD ED	Mean Δr.PSI	SD Δr.PSI	Mean z−score	SD z−score
-156	TMEM38A	15	15	-7.235	7.235	1.234	-1.717	0.130	-2.660	0.187
-157	STAR	5	7	-7.232	7.232	0.901	-1.185	0.087	-1.894	0.126
-158	RAB27B	4	5	-7.222	7.222	0.463	-1.432	0.058	-2.250	0.084
-159	PSAP	10	13	-7.195	7.195	0.944	-2.322	0.208	-3.530	0.299
-160	CHST10	3	5	-7.166	7.166	1.030	-1.005	0.027	-1.635	0.039
-161	AP4M1	8	12	-7.164	7.164	0.875	-1.454	0.082	-2.281	0.118
-162	NBR1	12	17	-7.151	7.151	1.060	-1.159	0.092	-1.857	0.132
-163	SCNN1B	8	11	-7.145	7.145	0.523	-1.839	0.143	-2.836	0.205
-164	MSN	7	12	-7.113	7.113	0.780	-1.047	0.079	-1.695	0.114
-165	PACRG	5	5	-7.073	7.073	1.016	-1.179	0.059	-1.886	0.085
-166	CYLC1	2	3	-7.058	7.058	0.691	-2.187	0.246	-3.336	0.354
-167	CAMKK1	6	11	-7.044	7.044	1.417	-1.526	0.134	-2.385	0.193
-168	PACSIN2	8	10	-7.039	7.039	1.522	-1.755	0.276	-2.715	0.397
-169	SLC25A11	4	7	-6.983	6.983	0.924	-1.148	0.103	-1.841	0.148
-170	MRPL1	5	9	-6.973	6.973	0.646	-1.861	0.065	-2.867	0.094
-171	O3FAR1	2	3	-6.935	6.935	0.402	-0.927	0.002	-1.523	0.002
-172	VDAC2	4	5	-6.935	6.935	1.601	-1.051	0.087	-1.701	0.125
-173	CD3D	3	4	-6.837	6.837	2.278	-1.186	0.169	-1.896	0.243
-174	TRIM16L	4	6	-6.809	6.809	1.931	-1.242	0.270	-1.976	0.389
-175	AFP	10	12	-6.778	6.778	1.149	-1.489	0.186	-2.331	0.268
-176	OXR1	11	16	-6.742	6.742	0.815	-2.112	0.185	-3.228	0.266
-177	SLC22A14	6	10	-6.739	6.739	1.234	-1.044	0.112	-1.692	0.161
-178	SLC46A2	3	4	-6.697	6.697	1.594	-1.529	0.316	-2.390	0.455
-179	MKI67IP	4	7	-6.655	6.655	0.471	-0.941	0.021	-1.543	0.030
-180	SLC25A38	5	7	-6.605	6.605	0.346	-1.088	0.061	-1.755	0.087
-181	HDAC8	7	10	-6.562	6.562	1.164	-1.203	0.092	-1.920	0.133
-182	OTUD6B	4	7	-6.457	6.457	1.757	-1.497	0.258	-2.343	0.371
-183	SELRC1	2	3	-6.449	6.449	1.803	-1.747	0.261	-2.703	0.375
-184	LTA4H	15	16	-6.447	6.447	1.037	-1.674	0.134	-2.598	0.192
-185	UCHL3	4	6	-6.409	6.409	0.295	-1.112	0.020	-1.789	0.029
-186	EIF3CL	10	12	-6.405	6.405	1.137	-1.068	0.091	-1.726	0.131
-187	ALB	9	13	-6.380	6.380	1.221	-1.538	0.214	-2.402	0.307
-188	TSSK3	2	2	-6.374	6.374	0.499	-1.023	0.048	-1.662	0.069
-189	SLC7A13	3	4	-6.352	6.352	0.606	-1.812	0.190	-2.796	0.274
-190	GCNT2	3	5	-6.308	6.308	1.600	-1.652	0.274	-2.566	0.394
-191	COCH	6	8	-6.291	6.291	1.097	-1.604	0.189	-2.498	0.272
-192	RAN	4	5	-6.284	6.284	0.811	-1.786	0.133	-2.760	0.191
-193	MRPL48	4	6	-6.250	6.250	1.218	-1.072	0.096	-1.731	0.138
-194	GPBP1	6	11	-6.241	6.241	0.842	-1.288	0.103	-2.042	0.149
-195	TRUB1	4	6	-6.232	6.232	0.908	-1.139	0.037	-1.828	0.053
-196	CPOX	5	7	-6.203	6.203	0.508	-1.136	0.087	-1.824	0.126
-197	DTD1	3	5	-6.203	6.203	0.235	-1.128	0.088	-1.812	0.126
-198	SNX9	11	18	-6.161	6.161	0.801	-1.403	0.170	-2.208	0.245
-199	TTC25	8	11	-6.128	6.128	0.985	-1.394	0.097	-2.195	0.140
-200	BEST1	14	20	-6.041	6.041	1.897	-1.119	0.101	-1.799	0.145
-201	RENBP	6	11	-6.015	6.015	0.932	-1.558	0.162	-2.431	0.234
-202	IMMT	10	14	-5.957	5.957	1.068	-1.811	0.261	-2.795	0.375
-203	ST6GALNAC6	3	5	-5.932	5.932	0.900	-1.046	0.096	-1.694	0.139
-204	GFM2	15	20	-5.928	5.928	0.686	-1.346	0.131	-2.126	0.189
-205	SARS	9	11	-5.914	5.914	0.575	-1.040	0.040	-1.685	0.057
-206	TAGLN2	4	4	-5.879	5.879	1.132	-1.352	0.110	-2.135	0.159
-207	PDIA3	11	12	-5.878	5.878	0.842	-1.844	0.152	-2.843	0.219
-208	EXOC5	12	16	-5.865	5.865	0.504	-1.256	0.180	-1.996	0.259
-209	FAM55C	3	5	-5.845	5.845	0.477	-1.127	0.136	-1.811	0.196
-210	NDFIP1	4	7	-5.814	5.814	0.892	-1.084	0.083	-1.750	0.120
-211	REEP4	5	8	-5.789	5.789	1.691	-1.474	0.248	-2.311	0.357
-212	UBXN2B	6	8	-5.788	5.788	0.528	-1.203	0.055	-1.921	0.079
-213	PARK2	8	12	-5.733	5.733	1.256	-2.463	0.344	-3.733	0.495
-214	ZNF593	2	3	-5.664	5.664	1.276	-1.493	0.262	-2.338	0.377
-215	ARSF	8	9	-5.662	5.662	0.570	-2.276	0.186	-3.465	0.268
-216	ST5	12	16	-5.656	5.656	0.931	-1.096	0.102	-1.766	0.147
-217	EIF2S3	8	11	-5.551	5.551	0.903	-0.973	0.049	-1.589	0.070
-218	SPATA6	8	10	-5.545	5.545	0.638	-1.645	0.075	-2.556	0.108
-219	FMO2	6	8	-5.539	5.539	0.606	-1.878	0.131	-2.892	0.189
-220	SERPINB3	3	5	-5.469	5.469	0.871	-1.024	0.088	-1.662	0.126
-221	ARHGAP8	7	9	-5.462	5.462	0.810	-1.230	0.115	-1.960	0.165
-222	EPHX2	10	15	-5.419	5.419	0.663	-1.504	0.078	-2.353	0.113
-223	GOT2	6	10	-5.299	5.299	0.912	-1.228	0.143	-1.957	0.206
-224	DTWD2	4	6	-5.269	5.269	0.675	-1.586	0.087	-2.472	0.125
-225	ZAP70	8	11	-5.232	5.232	1.493	-1.357	0.283	-2.142	0.407
-226	ANKRD45	4	5	-5.218	5.218	0.623	-1.214	0.094	-1.936	0.135
-227	DCTN2	9	13	-5.188	5.188	0.606	-1.461	0.067	-2.291	0.097
-228	UQCRC2	8	12	-5.183	5.183	0.764	-1.116	0.085	-1.795	0.122
-229	FAM20B	6	7	-5.164	5.164	0.334	-1.428	0.042	-2.244	0.060
-230	TMBIM1	3	5	-5.160	5.160	0.136	-0.933	0.020	-1.532	0.028
-231	UROD	6	10	-5.147	5.147	0.306	-1.152	0.061	-1.847	0.088
-232	SSRP1	9	15	-5.126	5.126	0.984	-1.552	0.303	-2.423	0.437
-233	CORO1C	8	9	-5.092	5.092	0.551	-1.765	0.079	-2.729	0.114
-234	FAM114A2	9	11	-5.091	5.091	0.633	-1.457	0.085	-2.286	0.123
-235	SLC43A3	7	11	-5.072	5.072	0.442	-1.375	0.080	-2.168	0.115
-236	SLC35B4	5	9	-5.068	5.068	0.629	-1.103	0.096	-1.776	0.138
-237	NDUFV1	6	10	-5.063	5.063	0.671	-1.006	0.035	-1.636	0.051
-238	DET1	4	7	-5.052	5.052	0.393	-1.287	0.066	-2.042	0.095
-239	LONRF1	7	12	-5.033	5.033	0.269	-1.039	0.090	-1.685	0.130
-240	FKBP4	6	10	-4.979	4.979	0.452	-0.960	0.022	-1.571	0.031

Position	gs	No. of sel. ps*	No. of ps pres.	Directed ED	Mean ED	SD ED	Mean Δr.PSI	SD Δr.PSI	Mean z–score	SD z–score
-241	RDX	8	13	-4.910	4.910	0.388	-2.131	0.185	-3.255	0.267
-242	RFTN2	8	9	-4.884	4.884	0.578	-1.459	0.087	-2.288	0.126
-243	MYNN	7	7	-4.873	4.873	0.571	-2.139	0.160	-3.267	0.230
-244	RPS6KL1	5	9	-4.868	4.868	0.699	-1.282	0.259	-2.033	0.372
-245	UBE2E3	3	4	-4.858	4.858	0.458	-1.215	0.070	-1.938	0.100
-246	GLA	5	7	-4.809	4.809	0.385	-1.880	0.096	-2.894	0.139
-247	CASC4	5	9	-4.791	4.791	0.430	-1.707	0.111	-2.645	0.160
-248	THPO	3	5	-4.757	4.757	0.971	-1.008	0.054	-1.640	0.078
-249	C7orf36	2	3	-4.734	4.734	0.397	-1.471	0.066	-2.306	0.096
-250	NOXRED1	5	7	-4.605	4.605	0.392	-1.828	0.098	-2.819	0.142
-251	PRRC1	6	8	-4.601	4.601	0.492	-2.149	0.185	-3.282	0.266
-252	GMEB1	8	9	-4.597	4.597	0.828	-1.294	0.158	-2.051	0.227
-253	ELMOD2	5	7	-4.517	4.517	0.484	-2.235	0.140	-3.405	0.202
-254	PREB	6	9	-4.511	4.511	0.828	-1.275	0.173	-2.024	0.249
-255	FZR1	11	13	-4.297	4.297	0.577	-1.087	0.069	-1.753	0.099
-256	RAB5A	4	5	-4.273	4.273	0.422	-1.419	0.071	-2.231	0.103
-257	SDSL	4	6	-4.259	4.259	0.686	-1.314	0.088	-2.080	0.127
-258	NUDT22	3	5	-4.121	4.121	0.465	-1.103	0.131	-1.776	0.188
-259	RPL24	4	5	-4.046	4.046	0.230	-1.310	0.062	-2.074	0.090
281	FUNDC2	3	5	4.270	4.270	0.645	1.402	0.153	1.829	0.220
280	CA11	5	9	4.349	4.349	0.162	1.238	0.044	1.592	0.063
279	EIF4A1	6	9	4.426	4.426	0.645	1.375	0.110	1.789	0.158
278	ADSSL1	7	13	4.491	4.491	0.686	1.285	0.070	1.661	0.100
277	GALT	6	9	4.527	4.527	0.778	1.297	0.085	1.677	0.123
276	PSMA3	6	9	4.641	4.641	0.685	1.427	0.191	1.865	0.275
275	PEX5	8	14	4.800	4.800	0.749	1.685	0.203	2.235	0.292
274	AQP8	4	6	4.880	4.880	0.571	1.493	0.103	1.960	0.149
273	CPEB4	5	8	4.891	4.891	0.851	1.719	0.218	2.284	0.313
272	KLRC2	2	3	4.895	4.895	1.154	1.394	0.064	1.816	0.093
271	TCF25	13	17	5.173	5.173	0.889	1.443	0.149	1.888	0.214
270	MTMR14	9	17	5.245	5.245	0.812	1.495	0.075	1.962	0.108
269	MAPK8	9	10	5.257	5.257	0.687	1.485	0.084	1.949	0.121
268	LZTR1	10	19	5.324	5.324	0.997	1.447	0.112	1.894	0.161
267	SNAP23	4	4	5.369	5.369	1.176	1.530	0.197	2.013	0.284
266	LYZL4	3	4	5.369	5.369	0.097	1.418	0.037	1.852	0.054
265	FAM86A	3	5	5.375	5.375	0.560	1.276	0.095	1.647	0.137
264	CYP4F22	6	11	5.481	5.481	0.394	1.351	0.050	1.755	0.073
263	STAU1	11	12	5.504	5.504	0.829	1.581	0.104	2.086	0.149
262	TAB1	7	12	5.530	5.530	0.874	1.438	0.122	1.880	0.176
261	FABP3	3	4	5.563	5.563	0.225	1.249	0.071	1.609	0.102
260	NUDT21	5	7	5.592	5.592	0.701	1.530	0.067	2.013	0.097
259	MACROD1	4	7	5.611	5.611	0.545	1.338	0.112	1.737	0.162
258	PSMB10	5	7	5.652	5.652	0.932	1.285	0.081	1.660	0.116
257	NACA	4	7	5.669	5.669	0.661	1.389	0.147	1.810	0.212
256	CORO2A	10	11	5.686	5.686	1.057	1.486	0.093	1.950	0.134
255	FAM82A2	7	11	5.701	5.701	0.554	1.247	0.057	1.605	0.082
254	NELF	9	13	5.705	5.705	1.499	1.602	0.200	2.117	0.288
253	RCN1	4	6	5.713	5.713	0.398	1.564	0.068	2.061	0.097
252	CHAF1B	11	13	5.721	5.721	0.376	1.408	0.050	1.838	0.072
251	GTF2B	6	7	5.771	5.771	0.712	1.645	0.097	2.178	0.139
250	CLK2	7	12	5.792	5.792	0.304	1.244	0.031	1.601	0.045
249	LRRC71	9	15	5.808	5.808	0.509	1.333	0.071	1.729	0.102
248	RABGGTB	6	8	5.816	5.816	0.426	1.430	0.067	1.870	0.097
247	CAMK1D	10	12	5.871	5.871	0.923	1.346	0.115	1.749	0.165
246	TMEM45A	4	5	5.890	5.890	0.234	1.486	0.050	1.950	0.072
245	FOXP3	9	11	5.911	5.911	0.763	1.784	0.122	2.378	0.175
244	CD4	7	9	5.941	5.941	1.149	1.652	0.137	2.188	0.197
243	SH3GLB1	7	9	5.956	5.956	0.582	1.283	0.068	1.657	0.097
242	PRKAB2	4	7	5.963	5.963	1.211	1.745	0.144	2.322	0.207
241	KIAA0090	13	23	5.979	5.979	0.581	1.933	0.199	2.593	0.287
240	SEMA4A	12	14	6.024	6.024	0.764	1.519	0.154	1.997	0.222
239	COQ4	4	7	6.068	6.068	0.778	1.268	0.041	1.635	0.059
238	PPIG	7	10	6.115	6.115	0.524	1.453	0.086	1.902	0.124
237	MTA1	11	16	6.118	6.118	0.885	2.321	0.256	3.151	0.369
236	ALDH1L2	12	22	6.129	6.129	0.655	1.414	0.091	1.846	0.132
235	C9orf91	5	8	6.145	6.145	1.061	1.236	0.060	1.589	0.086
234	MRPS22	7	8	6.153	6.153	0.426	1.463	0.135	1.916	0.194
233	DCP1B	5	8	6.178	6.178	0.984	2.408	0.150	3.277	0.216
232	GABPB1	6	9	6.286	6.286	0.891	2.490	0.166	3.395	0.238
231	KLHL1	9	11	6.320	6.320	0.218	2.313	0.046	3.139	0.066
230	HPD	9	11	6.364	6.364	0.952	1.461	0.095	1.913	0.136
229	PARP3	10	12	6.368	6.368	0.783	1.294	0.056	1.673	0.080
228	KEAP1	3	5	6.409	6.409	0.967	1.909	0.205	2.558	0.295
227	PEPD	9	12	6.413	6.413	1.361	1.873	0.141	2.507	0.204
226	GRAMD1B	15	18	6.421	6.421	1.126	1.646	0.116	2.180	0.167
225	GNA15	4	7	6.450	6.450	1.167	1.495	0.137	1.963	0.198
224	KRT20	5	7	6.491	6.491	1.322	1.480	0.130	1.941	0.186
223	NDE1	7	8	6.533	6.533	0.606	2.513	0.131	3.427	0.189
222	RASL10B	2	3	6.555	6.555	0.609	1.257	0.025	1.621	0.036
221	AIDA	5	8	6.556	6.556	0.770	1.657	0.103	2.196	0.148
220	SERPINA3	3	4	6.564	6.564	1.246	1.494	0.108	1.961	0.155
219	GTF3C5	10	11	6.614	6.614	1.218	1.700	0.162	2.257	0.233
218	PDCD5	3	5	6.636	6.636	0.199	1.657	0.132	2.196	0.189
217	QTRT1	8	10	6.652	6.652	0.728	1.446	0.065	1.892	0.094
216	KLHL6	6	7	6.765	6.765	0.433	1.733	0.033	2.305	0.048

Position	gs	No. of sel. ps*	No. of ps pres.	Directed ED	Mean ED	SD ED	Mean Δr.PSI	SD Δr.PSI	Mean z−score	SD z−score
215	XPA	4	6	6.768	6.768	0.559	2.263	0.202	3.068	0.291
214	ARIH2	9	12	6.788	6.788	0.522	1.334	0.054	1.731	0.077
213	KRT23	6	7	6.807	6.807	0.880	2.304	0.149	3.126	0.214
212	FKBP7	3	4	6.815	6.815	0.472	1.256	0.023	1.619	0.034
211	PPM1G	8	10	6.824	6.824	1.060	1.485	0.133	1.948	0.191
210	CRISPLD2	10	12	6.828	6.828	0.807	1.291	0.072	1.669	0.104
209	NUDT6	2	3	6.829	6.829	0.616	1.781	0.025	2.374	0.036
208	L3MBTL3	16	20	6.833	6.833	0.688	2.042	0.225	2.750	0.324
207	PRKCE	7	13	6.836	6.836	0.700	1.359	0.105	1.767	0.150
206	TBL1XR1	10	14	6.874	6.874	1.142	1.989	0.211	2.674	0.303
205	ILKAP	7	11	6.885	6.885	0.615	1.273	0.045	1.643	0.065
204	PRDM14	5	6	6.907	6.907	0.550	1.406	0.044	1.834	0.063
203	VTI1B	3	5	6.928	6.928	0.320	1.675	0.046	2.222	0.066
202	CBR3	2	3	6.959	6.959	0.297	1.250	0.025	1.610	0.036
201	C17orf75	8	9	6.966	6.966	1.132	1.856	0.106	2.482	0.153
200	IRAK3	7	10	6.973	6.973	0.371	2.571	0.077	3.511	0.110
199	RUFY1	14	19	6.994	6.994	1.138	1.886	0.141	2.525	0.203
198	ANKRD44	15	25	7.003	7.003	0.978	1.996	0.485	2.683	0.698
197	CFHR3	9	12	7.010	7.010	0.168	2.151	0.056	2.906	0.081
196	PFDN5	3	4	7.030	7.030	0.829	1.385	0.022	1.804	0.031
195	TEX10	14	14	7.034	7.034	0.575	2.030	0.077	2.733	0.111
194	AHCYL1	15	17	7.036	7.036	0.972	1.412	0.097	1.843	0.139
193	CDK15	10	13	7.097	7.097	1.099	1.684	0.123	2.235	0.177
192	BFSP1	5	8	7.105	7.105	2.675	1.895	0.369	2.539	0.531
191	DCTD	4	5	7.112	7.112	0.670	1.454	0.047	1.903	0.068
190	TBC1D5	15	19	7.123	7.123	1.417	1.383	0.073	1.801	0.105
189	TMPRSS11E	6	8	7.140	7.140	0.759	1.744	0.073	2.321	0.105
188	KLRC1	2	3	7.168	7.168	1.784	1.811	0.112	2.417	0.161
187	HPSE2	8	11	7.169	7.169	0.847	1.659	0.058	2.198	0.083
186	AKR1B15	6	6	7.178	7.178	1.559	1.881	0.286	2.518	0.412
185	BRD9	10	16	7.194	7.194	1.229	1.261	0.058	1.626	0.083
184	ERLIN1	7	9	7.216	7.216	1.285	2.352	0.245	3.196	0.353
183	DGCR8	10	13	7.229	7.229	0.914	1.815	0.058	2.423	0.084
182	C15orf23	8	9	7.236	7.236	1.143	1.499	0.154	1.968	0.222
181	ELAVL2	3	5	7.269	7.269	0.696	2.451	0.059	3.338	0.085
180	NDEL1	6	8	7.292	7.292	1.628	1.399	0.063	1.824	0.091
179	HP	3	4	7.312	7.312	1.456	1.375	0.045	1.790	0.065
178	PINX1	4	6	7.312	7.312	0.310	1.230	0.021	1.581	0.030
177	IFT81	10	17	7.315	7.315	0.845	1.483	0.067	1.946	0.096
176	CBR4	3	5	7.363	7.363	0.350	1.887	0.083	2.526	0.120
175	LECT1	4	7	7.397	7.397	0.185	1.309	0.114	1.695	0.163
174	MAB21L3	5	6	7.407	7.407	0.730	1.350	0.026	1.753	0.037
173	CECR1	7	10	7.408	7.408	1.113	1.708	0.154	2.270	0.222
172	GNAO1	6	10	7.439	7.439	1.882	2.389	0.406	3.249	0.584
171	TINF2	9	12	7.440	7.440	1.082	1.902	0.120	2.548	0.173
170	HPR	2	3	7.452	7.452	1.331	1.389	0.033	1.809	0.048
169	LYVE1	4	6	7.457	7.457	1.370	1.250	0.066	1.611	0.095
168	ATG16L1	11	16	7.476	7.476	1.271	2.005	0.165	2.697	0.237
167	GABRB2	5	9	7.483	7.483	0.345	1.415	0.014	1.848	0.020
166	MICA	4	5	7.562	7.562	1.593	1.748	0.106	2.327	0.152
165	FAHD2A	7	8	7.589	7.589	1.270	1.414	0.079	1.845	0.114
164	SERINC1	7	10	7.593	7.593	0.513	1.362	0.106	1.771	0.153
163	CALD1	9	14	7.640	7.640	1.481	1.692	0.165	2.246	0.238
162	FAHD2B	2	2	7.650	7.650	0.393	1.438	0.064	1.880	0.092
161	MS4A4A	3	5	7.652	7.652	0.371	1.615	0.027	2.135	0.039
160	IKZF2	5	7	7.655	7.655	1.335	1.653	0.097	2.191	0.140
159	CALCOCO2	10	11	7.691	7.691	0.406	1.513	0.046	1.988	0.067
158	GUCY1B3	12	14	7.771	7.771	1.608	1.758	0.183	2.341	0.264
157	RNF19B	7	9	7.772	7.772	0.280	1.639	0.039	2.170	0.056
156	ABCB6	14	18	7.813	7.813	2.085	1.556	0.190	2.051	0.273
155	SMYD1	8	9	7.814	7.814	1.205	1.534	0.056	2.018	0.080
154	RAB8B	4	7	7.839	7.839	0.431	1.770	0.150	2.358	0.215
153	ACP6	10	10	7.854	7.854	1.554	1.274	0.039	1.644	0.056
152	SSSCA1	2	3	7.860	7.860	2.925	1.346	0.198	1.748	0.285
151	PSMC2	9	11	7.884	7.884	0.825	2.329	0.180	3.162	0.259
150	CDKN1B	2	2	7.902	7.902	0.072	1.346	0.000	1.748	0.000
149	TSPAN6	4	7	7.979	7.979	0.623	1.394	0.054	1.817	0.077
148	KLHDC2	6	11	7.991	7.991	0.663	1.306	0.093	1.690	0.133
147	CCT5	6	11	8.028	8.028	0.801	2.417	0.193	3.290	0.278
146	FETUB	4	7	8.065	8.065	0.786	1.593	0.154	2.104	0.222
145	NDUFA10	7	9	8.091	8.091	0.584	2.518	0.075	3.435	0.108
144	C16orf71	6	8	8.093	8.093	1.976	1.818	0.293	2.428	0.421
143	DYX1C1	5	8	8.130	8.130	0.821	1.223	0.042	1.571	0.060
142	WDR45L	7	10	8.131	8.131	0.935	1.438	0.058	1.881	0.083
141	AKR1B10	6	8	8.162	8.162	2.164	1.951	0.327	2.619	0.471
140	ETS2	8	9	8.263	8.263	0.904	1.565	0.083	2.064	0.119
139	IVD	6	10	8.263	8.263	1.606	1.513	0.263	1.988	0.378
138	DLST	9	11	8.293	8.293	1.921	1.903	0.278	2.549	0.400
137	FADS1	9	13	8.302	8.302	1.727	2.334	0.439	3.170	0.632
136	FAM63A	6	9	8.359	8.359	1.495	1.708	0.217	2.269	0.312
135	TSPAN8	7	8	8.363	8.363	1.696	1.470	0.120	1.927	0.173
134	FBXO24	9	10	8.364	8.364	2.419	1.661	0.200	2.202	0.288
133	RWDD4	3	5	8.366	8.366	1.005	2.254	0.159	3.055	0.229
132	USP18	5	9	8.387	8.387	1.053	1.291	0.067	1.668	0.097
131	RANGAP1	14	15	8.395	8.395	1.452	1.362	0.060	1.772	0.086

Position	gs	No. of sel. ps*	No. of ps pres.	Directed ED	Mean ED	SD ED	Mean Δr.PSI	SD Δr.PSI	Mean z–score	SD z–score
130	LAP3	10	12	8.402	8.402	1.424	1.713	0.195	2.276	0.280
129	C12orf26	9	11	8.461	8.461	1.149	1.707	0.141	2.267	0.203
128	GOLM1	5	8	8.478	8.478	2.050	1.644	0.167	2.177	0.240
127	DUSP11	5	9	8.508	8.508	0.963	1.632	0.160	2.160	0.230
126	DAXX	6	10	8.536	8.536	1.624	1.945	0.252	2.610	0.363
125	ATG3	6	8	8.540	8.540	1.615	1.836	0.089	2.454	0.128
124	KIAA1598	11	15	8.547	8.547	0.980	1.705	0.131	2.265	0.188
123	SH3BP5	5	9	8.551	8.551	2.459	1.738	0.196	2.311	0.282
122	GSTK1	7	8	8.552	8.552	1.837	1.426	0.254	1.863	0.366
121	SLC12A4	13	23	8.560	8.560	1.471	1.574	0.222	2.076	0.320
120	ACVR1	8	9	8.574	8.574	2.334	2.002	0.287	2.692	0.413
119	TXNL1	6	9	8.632	8.632	0.346	1.696	0.034	2.252	0.049
118	SERPINE1	7	8	8.682	8.682	1.665	1.672	0.114	2.217	0.165
117	PHB2	6	8	8.771	8.771	0.870	1.328	0.116	1.723	0.168
116	C15orf55	5	7	8.781	8.781	1.245	1.243	0.059	1.599	0.085
115	NME7	9	11	8.787	8.787	1.198	1.845	0.096	2.466	0.138
114	MAPT	7	12	8.797	8.797	2.752	1.810	0.245	2.416	0.353
113	CYB5R4	8	13	8.822	8.822	1.088	2.574	0.182	3.516	0.262
112	STK38L	9	12	8.829	8.829	1.432	1.723	0.169	2.290	0.243
111	WIF1	6	10	8.845	8.845	1.547	1.693	0.102	2.248	0.147
110	UPP2	5	9	8.859	8.859	1.071	2.213	0.109	2.996	0.157
109	ODC1	9	10	8.922	8.922	1.183	1.808	0.116	2.412	0.167
108	FAM161A	4	6	8.930	8.930	1.885	1.322	0.067	1.713	0.097
107	DNAI2	9	12	8.933	8.933	1.676	1.580	0.159	2.085	0.228
106	C9orf7	3	5	8.961	8.961	1.388	1.499	0.138	1.969	0.199
105	HARS2	10	13	8.964	8.964	2.054	1.351	0.080	1.755	0.115
104	STOML3	5	8	8.977	8.977	0.610	1.681	0.072	2.231	0.104
103	FLI1	6	8	9.002	9.002	1.840	2.433	0.251	3.312	0.361
102	ANXA7	9	12	9.028	9.028	1.211	1.217	0.034	1.563	0.049
101	SF3A3	10	14	9.036	9.036	0.717	1.268	0.034	1.635	0.049
100	MS4A3	3	5	9.050	9.050	0.889	1.290	0.030	1.668	0.043
99	PXK	11	15	9.075	9.075	2.461	1.576	0.216	2.079	0.311
98	PAPSS2	9	12	9.075	9.075	1.797	1.436	0.078	1.878	0.112
97	SPCS3	4	5	9.088	9.088	0.396	2.162	0.101	2.922	0.146
96	SELE	8	11	9.089	9.089	1.370	1.789	0.154	2.385	0.222
95	TXNDC5	8	10	9.101	9.101	1.254	2.335	0.205	3.171	0.295
94	ATL1	8	10	9.165	9.165	1.009	1.829	0.074	2.444	0.107
93	DPP8	17	19	9.167	9.167	0.847	2.610	0.135	3.567	0.195
92	PLK1	8	10	9.170	9.170	1.206	2.523	0.175	3.442	0.252
91	TTC33	3	4	9.174	9.174	0.673	1.206	0.008	1.547	0.012
90	MAP2K6	7	9	9.193	9.193	1.863	1.457	0.051	1.907	0.073
89	ABHD10	3	5	9.202	9.202	0.337	2.465	0.027	3.358	0.038
88	LCORL	5	7	9.219	9.219	1.392	2.603	0.267	3.558	0.384
87	KNG1	8	12	9.245	9.245	0.979	1.511	0.112	1.985	0.161
86	TIMM23	4	5	9.245	9.245	0.578	1.959	0.264	2.630	0.380
85	TMEM161B	10	12	9.252	9.252	0.656	1.832	0.056	2.448	0.080
84	CFLAR	7	8	9.255	9.255	1.254	1.489	0.059	1.955	0.084
83	NOB1	7	9	9.257	9.257	1.488	1.442	0.101	1.886	0.145
82	YKT6	5	7	9.260	9.260	0.677	1.290	0.055	1.667	0.079
81	ITGB1	15	26	9.266	9.266	0.415	2.265	0.125	3.070	0.180
80	RASSF2	6	9	9.273	9.273	1.774	1.412	0.094	1.843	0.136
79	AASDHPPT	4	6	9.280	9.280	0.554	1.842	0.046	2.462	0.067
78	MATN3	6	7	9.283	9.283	0.942	2.276	0.134	3.087	0.192
77	PCOLCE2	5	9	9.331	9.331	0.593	1.920	0.020	2.573	0.029
76	GTF2E2	6	7	9.357	9.357	0.891	1.801	0.108	2.403	0.155
75	TMEM209	11	14	9.380	9.380	1.056	1.821	0.111	2.432	0.160
74	AGPAT6	10	11	9.389	9.389	2.142	1.728	0.191	2.298	0.275
73	QRSL1	9	11	9.389	9.389	0.913	2.592	0.182	3.541	0.262
72	SLC9A6	12	16	9.396	9.396	0.921	1.867	0.059	2.498	0.084
71	ADPGK	6	9	9.417	9.417	1.168	1.491	0.088	1.956	0.126
70	NXF2	7	10	9.428	9.428	1.170	1.425	0.098	1.861	0.141
69	CD99L2	6	10	9.448	9.448	1.660	1.839	0.120	2.457	0.173
68	PDLIM1	5	7	9.462	9.462	0.672	1.619	0.052	2.141	0.075
67	NQO2	4	6	9.473	9.473	1.293	2.217	0.076	3.001	0.109
66	ABCB7	15	16	9.537	9.537	1.036	2.715	0.149	3.718	0.214
65	AKR1E2	8	10	9.548	9.548	1.112	1.396	0.054	1.820	0.077
64	TMPRSS2	10	12	9.555	9.555	2.077	1.873	0.195	2.507	0.281
63	BAG1	5	7	9.573	9.573	0.555	2.125	0.070	2.869	0.101
62	GMDS	8	10	9.607	9.607	1.424	1.361	0.017	1.769	0.025
61	ALDH1A2	8	13	9.626	9.626	1.231	1.344	0.127	1.745	0.183
60	CA5B	6	7	9.641	9.641	1.688	1.909	0.061	2.559	0.087
59	CPSF7	8	11	9.673	9.673	2.677	2.002	0.274	2.692	0.394
58	MYL2	4	5	9.694	9.694	1.557	1.294	0.045	1.674	0.065
57	DBC1	5	7	9.698	9.698	2.723	1.582	0.223	2.087	0.320
56	GPC6	7	9	9.725	9.725	0.614	1.841	0.066	2.460	0.096
55	M6PR	5	6	9.734	9.734	0.452	2.337	0.046	3.175	0.066
54	ITM2A	4	6	9.855	9.855	0.788	1.376	0.025	1.792	0.035
53	CHIA	5	9	9.882	9.882	1.582	2.193	0.085	2.966	0.123
52	NOL4	7	11	9.883	9.883	1.551	1.260	0.027	1.625	0.039
51	PCYOX1	4	6	9.938	9.938	0.594	1.368	0.025	1.780	0.036
50	PPP3CC	10	13	9.964	9.964	1.444	1.677	0.072	2.224	0.104
49	UBE2U	5	7	9.970	9.970	1.148	2.238	0.223	3.032	0.321
48	FAR1	8	10	9.989	9.989	2.914	2.014	0.301	2.709	0.433
47	RBM45	8	9	10.060	10.060	1.715	1.880	0.183	2.517	0.263
46	CLEC4D	4	6	10.099	10.099	0.839	2.139	0.060	2.890	0.087

A.9 Selected CAAPs

Position	gs	No. of sel. ps*	No. of ps pres.	Directed ED	Mean ED	SD ED	Mean Δr.PSI	SD Δr.PSI	Mean z−score	SD z−score
45	FAM122B	5	8	10.114	10.114	1.862	1.277	0.067	1.649	0.096
44	EXT1	9	11	10.141	10.141	1.673	1.342	0.064	1.743	0.092
43	NDRG1	5	8	10.187	10.187	2.050	2.122	0.143	2.865	0.205
42	USO1	17	22	10.188	10.188	0.656	2.713	0.083	3.715	0.120
41	PRKCB	10	18	10.276	10.276	2.214	1.961	0.276	2.633	0.397
40	YARS	11	13	10.315	10.315	1.611	1.441	0.098	1.885	0.142
39	FAM116A	14	16	10.398	10.398	1.818	2.104	0.166	2.839	0.239
38	C10orf88	4	6	10.402	10.402	1.367	1.553	0.028	2.046	0.041
37	IL18	12	13	10.407	10.407	1.163	1.931	0.078	2.591	0.113
36	ERAL1	5	9	10.430	10.430	0.302	1.287	0.054	1.664	0.078
35	SEC61A2	9	16	10.435	10.435	2.018	1.703	0.115	2.262	0.166
34	YARS2	3	5	10.455	10.455	0.427	1.250	0.019	1.610	0.027
33	SHQ1	10	10	10.504	10.504	1.337	2.301	0.102	3.122	0.146
32	HTR2B	3	3	10.514	10.514	2.272	1.394	0.197	1.818	0.283
31	PPP2R1B	16	17	10.520	10.520	1.707	2.823	0.169	3.873	0.243
30	C2orf49	3	4	10.601	10.601	0.933	1.538	0.077	2.025	0.111
29	KIAA1279	5	7	10.612	10.612	0.799	1.548	0.120	2.038	0.173
28	ST3GAL4	6	10	10.633	10.633	1.487	1.222	0.037	1.570	0.053
27	COIL	4	6	10.724	10.724	1.113	1.365	0.026	1.775	0.038
26	SEC23IP	16	18	10.733	10.733	1.561	1.454	0.096	1.904	0.138
25	GABRA4	5	9	10.735	10.735	0.937	1.209	0.022	1.552	0.032
24	RPUSD4	7	8	10.740	10.740	1.927	1.438	0.059	1.880	0.085
23	VTA1	7	8	10.758	10.758	1.694	1.449	0.046	1.896	0.066
22	PROSC	5	7	10.765	10.765	1.188	1.850	0.054	2.473	0.078
21	ZDHHC13	10	17	10.831	10.831	1.682	1.734	0.250	2.306	0.359
20	EXOSC7	6	10	10.849	10.849	2.132	1.261	0.031	1.625	0.044
19	ACSM1	7	13	10.897	10.897	1.119	2.488	0.083	3.391	0.120
18	AIFM1	11	20	10.935	10.935	0.625	1.365	0.084	1.776	0.121
17	PSEN1	8	10	11.032	11.032	1.241	2.147	0.060	2.901	0.086
16	MBIP	6	8	11.145	11.145	1.681	1.852	0.059	2.476	0.085
15	LPL	8	10	11.188	11.188	1.871	2.318	0.148	3.146	0.213
14	MINA	8	9	11.204	11.204	1.386	2.092	0.072	2.821	0.104
13	RHOT1	13	17	11.234	11.234	1.972	2.125	0.134	2.869	0.193
12	SAMD7	6	7	11.240	11.240	1.041	1.501	0.042	1.972	0.060
11	SLC7A14	5	7	11.257	11.257	1.201	1.294	0.050	1.673	0.071
10	IL31RA	11	14	11.414	11.414	1.760	1.463	0.053	1.916	0.076
9	TGFBRAP1	6	10	11.426	11.426	1.181	1.227	0.044	1.576	0.063
8	DCT	6	8	11.698	11.698	0.955	1.251	0.015	1.611	0.022
7	SEC22C	5	7	11.709	11.709	0.449	1.576	0.135	2.079	0.194
6	FBXW12	6	10	11.724	11.724	1.169	2.655	0.056	3.632	0.081
5	ATP1B4	6	7	11.899	11.899	2.175	2.184	0.145	2.954	0.208
4	LRRC28	6	7	11.916	11.916	1.536	1.746	0.058	2.324	0.084
3	ANGPT2	8	9	12.154	12.154	1.663	2.405	0.128	3.273	0.184
2	LANCL1	8	9	13.014	13.014	1.653	1.428	0.042	1.866	0.060
1	FAM151B	4	6	13.643	13.643	0.892	1.400	0.157	1.826	0.227

A.10 Selected Candidates for Validation Studies

The following table lists candidates that were selected for validation studies. *Selected probe sets that fulfilled thresholds of detection, z–score$_{\Delta r.PSI}$ and Euclidean Distance (ED). For details see results subsection 4.3.2. For ORFs in this list that are not defined as CAAPs, 'no. of selected ps' represents the number of ps above the detection threshold. *Note:* Although the mean values of ED and z–score$_{\Delta r.PSI}$ fulfill the defined thresholds, RANBP3 was dismissed as CAAP because more than 50 % of its probe sets did not passed the selection criteria.

Table A.2: Selected candidates for validation studies

Position in CAAPs	**gs**	**No. of selected ps***	**No. of ps present**	**Directed ED**	**Mean cent.ED**	**SD ED**	**Mean Δr.PSI**	**SD Δr.PSI**	**Mean z–score$_{\Delta r.PSI}$**	**SD z–score$_{\Delta r.PSI}$**	**Mean PSI$_{CRY1mut}$**	**SD PSI$_{CRY1mut}$**	**Protein half–life [hrs]**[63]
-11	ARG1	8	8	-10.667	10.667	2.060	-1.631	0.120	-2.536	0.173	3.077	0.079	n.d.
-83	ARFGAP3	14	15	-8.729	8.729	2.110	-1.686	0.211	-2.615	0.304	2.832	0.153	28.01
-103	PIH1D1	7	9	-8.278	8.278	1.098	-1.476	0.162	-2.313	0.233	3.039	0.154	50.31
-161	AP4M1	8	12	-7.164	7.164	0.875	-1.454	0.082	-2.281	0.118	3.224	0.065	n.d.
	NUTF2	4	4		3.933	0.506	0.309	0.039	0.344	0.085	1.565	0.043	139.71
	NPM1	8	10		3.170	1.274	-0.542	0.287	-0.969	0.414	2.215	0.219	150.84
	THOC6	9	10		6.109	2.842	0.809	0.201	0.976	0.289	1.606	0.033	149.15
	RANBP3	12	15		6.339	3.478	1.304	0.753	1.688	1.084	1.442	0.127	48.42
	MR1	5	6		7.810	1.865	0.699	0.145	0.817	0.209	1.325	0.061	0.93
223	NDE1	7	8	6.533	6.533	0.606	2.513	0.131	3.427	0.189	1.182	0.027	6.8
137	FADS1	9	13	8.302	8.302	1.727	2.334	0.439	3.170	0.632	1.044	0.053	n.d.
92	PLK1	8	10	9.170	9.170	1.206	2.523	0.175	3.442	0.252	1.017	0.007	5.05
35	SEC61A2	9	16	10.435	10.435	2.018	1.703	0.115	2.262	0.166	1.819	0.074	51.43

A.11 Rhythmically Abundant Proteins

The following table shows rhythmically abundant proteins and the study they were identified for gene symbols (gs) that have been detected with the microarray study performed in here (see sub–subsection in 4.3.5). Gray colored lines highlight gs that where found to be rhythmically abundant in the circadian liver transcriptome analysis by Hughes *et al.*[40]. *Note:* 45 rhythmically abundant proteins identified by Robles *et al.* (*Plos Genetics*, in press) are not published so far and thus not included here.

Table A.3: Rhythmically abundant proteins

gs	Publication(s)		
AHSG	Møller *et al.*[45]		
AKR1B1	Møller *et al.*[45]		
ALB	Reddy *et al.*[41]	Møller *et al.*[45]	
ALDH2	Reddy *et al.*[41]		
ALDOA	Deery *et al.*[43]		
ANXA2	Møller *et al.*[45]		
ANXA5	Møller *et al.*[45]		
APOA1	Martino *et al.*[44]		
APOA4	Reddy *et al.*[41]		
ARG1	Reddy *et al.*[41]		
ASS1	Reddy *et al.*[41]		
ATP5A1	Reddy *et al.*[41]		
BHMT	Reddy *et al.*[41]		
CALR	Reddy *et al.*[41]		
CCT5	Deery *et al.*[43]		
CCT7	Møller *et al.*[45]		
CKB	Møller *et al.*[45]	Deery *et al.*[43]	
CLU	Martino *et al.*[44]		
CS	Møller *et al.*[45]		
DPYSL2	Deery *et al.*[43]		
EEF1D	Reddy *et al.*[41]		
ENO1	Reddy *et al.*[41]	Møller *et al.*[45]	Deery *et al.*[43]
ENO2	Møller *et al.*[45]		
ETFA	Deery *et al.*[43]		
GLUD1	Deery *et al.*[43]		
GSTM1	Reddy *et al.*[41]		
HAO1	Reddy *et al.*[41]		
HSP90AB1	Deery *et al.*[43]		
HSPA8	Deery *et al.*[43]		
HSPD1	Reddy *et al.*[41]		
KHK	Reddy *et al.*[41]		
LTA4H	Tsuji *et al.*[42]		
MAT2A	Deery *et al.*[43]		
NAPB	Deery *et al.*[43]		
PDIA3	Deery *et al.*[43]		
PGK1	Tsuji *et al.*[42]		
PPIA	Møller *et al.*[45]	Deery *et al.*[43]	
PSMA1	Tsuji *et al.*[42]		
PSMA6	Deery *et al.*[43]		
PSPH	Møller *et al.*[45]		
SELENBP1	Reddy *et al.*[41]		
TPI1	Møller *et al.*[45]	Deery *et al.*[43]	
TTR	Martino *et al.*[44]		
TUFM	Deery *et al.*[43]		
VCP	Møller *et al.*[45]		

A.12 CRY1 Interactors from UniHI Database

The following table lists proteins that are CRY1 interactors according to UniHI database and have been detected within the microarray screen. Highlighted proteins were found within CAAPs.

Table A.4: CRY1 interactors from UniHI database, date 01-11-13

gs	Name
MDFI	MyoD family inhibitor
PER1	period homolog 1 (Drosophila)
TRAF2	TNF receptor-associated factor 2
ARNTL	aryl hydrocarbon receptor nuclear translocator-like
CRY2	cryptochrome 2 (photolyase-like)
CSNK1E	casein kinase 1, epsilon
FBXL3	F-box and leucine-rich repeat protein 3
DARS2	aspartyl-tRNA synthetase 2, mitochondrial
HSPA8	heat shock 70kDa protein 8
ATP6V1B2	ATPase, H+ transporting, lysosomal 56/58kDa, V1 subunit B2
EEF1A1	eukaryotic translation elongation factor 1 alpha 1
TUBA4A	tubulin, alpha 4a
GMPPB	GDP-mannose pyrophosphorylase B
PPP2CB	protein phosphatase 2, catalytic subunit, beta isozyme
PPP2R1B	protein phosphatase 2, regulatory subunit A, beta
GRN	granulin
RPS27A	ribosomal protein S27a

A.13 Gene Ontology Enrichment Analysis

Gene Ontology analysis was performed using GO-Elite software[108]. GO terms were manually classified into major terms (separated by horizontal lines; see sub–subsection in 4.3.5). Gray highlighted GO terms are high in the GO hierarchy and contain > 200 gs and were thus excluded from visualization in major terms (see Fig. 4.15). BP–Biological process, CC–cellular component, MF–Molecular function. *[1]–Development and growth, *[2]–RNA processing, *[3]–Lipid biosynthesis

Table A.5: Gene Ontology terms of CAAPs

Major term	GO–ID	Name	Type	CAAPs	gs analyzed	gs in GO	% Changed	% Present	Z–Score	P–value
Mitosis	GO:0051310	metaphase plate congression	BP	3	4	14	75.00	28.57	4.6004	0.0027
	GO:0000776	kinetochore	CC	9	26	84	34.62	30.95	4.5560	0.0003
	GO:0034502	protein localization to chromosome	BP	3	5	20	60.00	25.00	3.9740	0.0064
	GO:0031063	regulation of histone deacetylation	BP	3	5	9	60.00	55.56	3.9740	0.0064
	GO:0007052	mitotic spindle organization	BP	3	6	22	50.00	27.27	3.4992	0.0119
	GO:0000781	chromosome, telomeric region	CC	3	7	37	42.86	18.92	3.1207	0.0195
	GO:0005694	chromosome	CC	11	56	262	19.64	21.37	2.7787	0.0152
	GO:0000236	mitotic prometaphase	BP	6	25	84	24.00	29.76	2.6127	0.0212
	GO:0032155	cell division site part	CC	4	14	38	28.57	36.84	2.5503	0.0318
	GO:0000278	mitotic cell cycle	BP	17	113	321	15.04	35.20	2.2460	0.0309
	GO:0070507	regulation of microtubule cytoskeleton organization	BP	5	22	74	22.73	29.73	2.2418	0.0429
	GO:0000087	M phase of mitotic cell cycle	BP	6	29	92	20.69	31.52	2.1919	0.0419
	GO:0045910	negative regulation of DNA recombination	BP	3	3	11	100	27.27	5.4937	0.0007
	GO:0005875	microtubule associated complex	CC	7	36	99	19.44	36.36	2.1823	0.0394
	GO:0015630	microtubule cytoskeleton	CC	7	36	94	19.44	38.30	2.1823	0.0394
	GO:0051494	negative regulation of cytoskeleton organization	BP	6	26	71	23.08	36.62	2.5002	0.0256
	GO:0008092	cytoskeletal protein binding	MF	26	194	574	13.40	33.80	2.1531	0.0404

A.13 Gene Ontology Enrichment Analysis

Major term	GO–ID	Name	Type	CAAPs	gs analyzed	gs in GO	% Changed	% Present	Z–Score	P–value
Vesicles and secretion	GO:0031974	membrane-enclosed lumen	CC	30	226	503	13.27	44.93	2.2624	0.0321
	GO:0015695	organic cation transport	BP	5	15	30	33.33	50.00	3.2839	0.0083
	GO:0022804	active transmembrane transporter activity	MF	17	98	342	17.35	28.65	2.8904	0.0072
	GO:0002028	regulation of sodium ion transport	BP	4	14	30	28.57	46.67	2.5503	0.0318
	GO:0006611	protein export from nucleus	BP	3	9	19	33.33	47.37	2.5423	0.0409
	GO:0006820	anion transport	BP	10	54	147	18.52	36.73	2.4388	0.0270
	GO:0046942	carboxylic acid transport	BP	11	66	160	16.67	41.25	2.1715	0.0471
	GO:0000045	autophagic vacuole assembly	BP	3	7	17	42.86	41.18	3.1207	0.0195
	GO:0006914	autophagy	BP	6	27	56	22.22	48.21	2.3929	0.0304
	GO:0050707	regulation of cytokine secretion	BP	6	24	61	25.00	39.34	2.7308	0.0174
	GO:0009101	glycoprotein biosynthetic process	BP	3	7	24	42.86	29.17	3.1207	0.0195
	GO:0006029	proteoglycan metabolic process	BP	3	9	29	33.33	31.03	2.5423	0.0409
	GO:0030427	site of polarized growth	CC	6	26	73	23.08	35.62	2.5002	0.0256
	GO:0043231	intracellular membrane-bounded organelle	CC	303	2931	8637	10.34	33.94	3.5115	0.0005
	GO:0022406	membrane docking	BP	4	15	31	26.67	48.39	2.3825	0.0404
	GO:0006814	sodium ion transport	BP	7	37	130	18.92	28.46	2.1008	0.0450
	GO:0055065	metal ion homeostasis	BP	16	107	286	14.95	37.41	2.1512	0.0398
	GO:0006508	proteolysis	BP	23	169	539	13.61	31.35	2.1005	0.0409
	GO:0005913	cell-cell adherens junction	CC	4	11	40	36.36	27.50	3.1619	0.0130
Protein binding	GO:0045296	cadherin binding	MF	3	8	21	37.50	38.10	2.8079	0.0292
	GO:0035258	steroid hormone receptor binding	MF	7	23	60	30.43	38.33	3.5836	0.0032
	GO:0030145	manganese ion binding	MF	3	9	39	33.33	23.08	2.5423	0.0409
	GO:0043394	proteoglycan binding	MF	3	3	14	100	21.43	5.4937	0.0007
	GO:0043393	regulation of protein binding	BP	8	33	78	24.24	42.31	3.0526	0.0077
	GO:0043242	negative regulation of protein complex disassembly	BP	5	18	45	27.78	40.00	2.7752	0.0188
	GO:0005518	collagen binding	MF	4	14	47	28.57	29.79	2.5503	0.0318
	GO:0032459	regulation of protein oligomerization	BP	3	9	16	33.33	56.25	2.5423	0.0409
	GO:0001871	pattern binding	MF	10	54	181	18.52	29.83	2.4388	0.0270
	GO:0003706	ligand-regulated transcription factor activity	MF	231	2269	6485	10.18	34.99	2.4361	0.0159
	GO:0031334	positive regulation of protein complex assembly	BP	6	29	78	20.69	37.18	2.1919	0.0419
Immune response	GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB cascade	BP	11	53	140	20.75	37.86	2.9861	0.0068
	GO:0002757	immune response-activating signal transduction	BP	17	82	194	20.73	42.27	3.7165	0.0013
	GO:0032496	response to lipopolysaccharide	BP	14	77	182	18.18	42.31	2.8144	0.0139
	GO:0042101	T cell receptor complex	CC	3	8	12	37.50	66.67	2.8079	0.0292
	GO:0007179	transforming growth factor beta receptor signaling pathway	BP	5	19	65	26.32	29.23	2.6290	0.0237
	GO:0003725	double-stranded RNA binding	MF	4	14	37	28.57	37.84	2.5503	0.0318
	GO:0009620	response to fungus	BP	3	9	28	33.33	32.14	2.5423	0.0409
	GO:0002250	adaptive immune response	BP	6	26	55	23.08	47.27	2.5002	0.0256
	GO:0000165	MAPKKK cascade	BP	10	54	142	18.52	38.03	2.4388	0.0270
	GO:0008063	Toll signaling pathway	BP	6	27	77	22.22	35.06	2.3929	0.0304
	GO:0051597	response to methylmercury	BP	3	4	10	75.00	40.00	4.6004	0.0027
	GO:0071407	cellular response to organic cyclic compound	BP	4	14	45	28.57	31.11	2.5503	0.0318
*1	GO:0005099	Ras GTPase activator activity	MF	5	22	111	22.73	19.82	2.2418	0.0429
	GO:0001708	cell fate specification	BP	4	11	44	36.36	25.00	3.1619	0.0130
	GO:0060135	maternal process involved in female pregnancy	BP	3	7	16	42.86	43.75	3.1207	0.0195
	GO:0003007	heart morphogenesis	BP	3	7	36	42.86	19.44	3.1207	0.0195
	GO:0001889	liver development	BP	5	20	70	25.00	28.57	2.4920	0.0293
*2	GO:0043631	RNA polyadenylation	BP	3	4	20	75.00	20.00	4.6004	0.0027
	GO:0005849	mRNA cleavage factor complex	CC	3	7	13	42.86	53.85	3.1207	0.0195
	GO:0000178	exosome (RNase complex)	CC	3	9	20	33.33	45.00	2.5423	0.0409
*3	GO:0045017	glycerolipid biosynthetic process	BP	9	44	105	20.45	41.90	2.6489	0.0152
	GO:0035384	thioester biosynthetic process	BP	3	9	20	33.33	45.00	2.5423	0.0409
	GO:0006633	fatty acid biosynthetic process	BP	9	49	105	18.37	46.67	2.2850	0.0388
Diverse	GO:0009068	aspartate family amino acid catabolic process	BP	3	7	16	42.86	43.75	3.1207	0.0195
	GO:0018107	peptidyl-threonine phosphorylation	BP	3	7	28	42.86	25.00	3.1207	0.0195
	GO:0009894	regulation of catabolic process	BP	25	165	440	15.15	37.50	2.7755	0.0086
	GO:0046148	pigment biosynthetic process	BP	6	24	44	25.00	54.55	2.7308	0.0174
	GO:0044036	cell wall macromolecule metabolic process	BP	3	9	16	33.33	56.25	2.5423	0.0409
	GO:0046504	glycerol ether biosynthetic process	BP	5	20	42	25.00	47.62	2.4920	0.0293
	GO:0033014	tetrapyrrole biosynthetic process	BP	4	15	26	26.67	57.69	2.3825	0.0404
	GO:0005524	ATP binding	MF	53	445	1488	11.91	29.91	2.1954	0.0313
	GO:0003824	catalytic activity	MF	222	1967	5363	11.29	36.68	4.2887	0.0000
	GO:0019012	virion	CC	311	2987	8888	10.41	33.61	3.7866	0.0002
	GO:0044444	cytoplasmic part	CC	268	2526	6373	10.61	39.64	3.6825	0.0003
	GO:0044297	cell body	CC	311	3001	8927	10.36	33.62	3.6702	0.0002

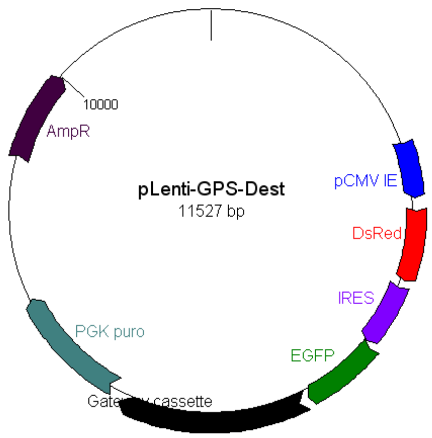
A.14 Vector Map

Vector map of the bicistronic fluorescent reporter pLenti6-GPS-*Dest*.

pLenti-GPS-Dest

Source/ Vendor:	Invitrogen derivate
Plasmid type:	lentiviral
Promoter:	CMV
Plasmid Size:	11527
Protein Tags:	N-terminal EGFP
Bacterial resistance:	Ampicillin
Notes:	DsRED and EGFP-fusion protein are expressed via an internal ribosome entry site

vector map



pCMV IE (2227~2755)
DsRed (2845~3522)
IRES (3564~4148)
EGFP (4152~4881)
Gateway cassette (4888~6597)
PGK-Puro (6679~7807)

Bibliography

[1] C.H. Johnson, T. Mori, and Y. Xu. A Cyanobacterial Circadian Clockwork. *Current biology : CB*, 18(17):R816–R825, September 2008.

[2] C. Robertson McClung. Plant Circadian Rhythms. *The Plant Cell Online*, 18(4):792–803, January 2006.

[3] J.C. Dunlap and J.J. Loros. The Neurospora Circadian System. *Journal of Biological Rhythms*, 19(5):414–424, January 2004.

[4] N. Peschel and C. Helfrich-Förster. Setting the clock – by nature: Circadian rhythm in the fruitfly Drosophila melanogaster. *FEBS Letters*, 585(10):1435–1442, May 2011.

[5] J. Bass. Circadian topology of metabolism. *Nature*, 491(7424):348–356, November 2012.

[6] F. Halberg and A.N. Stephens. Susceptibility to ouabain and physiologic circadian periodicity. *Proc. Minn. Acad. Sci.*, (27):139–143, 1959.

[7] DeMairan. *Observation botanique*, volume Hist. Acad. Imprimerie de Du Pont (Paris), August 1729.

[8] J.S. Szymanski. Die Verteilung der Ruhe- und Aktivitätsperioden bei weissen Ratten und Tanzmäusen. *Pflügers Archiv für die Gesamte Physiologie des Menschen und der Tiere*, 1918.

[9] P.C. Richter. A Behavioristic Study of the Activity of the Rat. *Comparative Psychology Monographs*, 1, 2:56, 1922.

[10] M.S. Johnson. Activity and Distribution of Certain Wild Mice in Relation to Biotic Communities. *Journal of Mammalogy*, 7(4):245–277, November 1926.

[11] O. Wahl. Neue Untersuchungen über das Zeitgedächtnis der Bienen. *Zeitschrift für vergleichende Physiologie*, 16(3):529–589, January 1932.

[12] D.E. Bianchi. An Endogenous Circadian Rhythm in Neurospora crassa. *Journal of General Microbiology*, 35(3):437–445, January 1964.

[13] E. Bünning. Zur Kenntnis der erblichen Tagesperioizität bei den Primärblättern von Phaseolus multiflorus. Jahrb. Wis, 1935.

[14] R.J. Konopka and S. Benzer. Clock Mutants of Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, 68(9):2112–2116, September 1971.

[15] P. Reddy, W.A. Zehring, D.A. Wheeler, V. Pirrotta, C. Hadfield, J.C. Hall, and M. Rosbash. Molecular analysis of the period locus in Drosophila melanogaster and identification of a transcript involved in biological rhythms. *Cell*, 38(3):701–710, October 1984.

[16] M.H. Vitaterna, D.P. King, A.M. Chang, J.M. Kornhauser, P.L. Lowrey, J.D. McDonald, W.F. Dove, L.H. Pinto, F.W. Turek, and J.S. Takahashi. Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior. *Science (New York, N.Y.)*, 264(5159):719–725, April 1994.

[17] C.S. Pittendrigh. On Temperature Independence In The Clock System Controlling Emergence Time In Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 40(10):1018–1029, October 1954.

[18] M.H. Vitaterna, J.S. Takahashi, and F.W. Turek. Overview of circadian rhythms. *Alcohol research & health: the journal of the National Institute on Alcohol Abuse and Alcoholism*, 25(2):85–93, 2001.

[19] A.C. Liu, W.G. Lewis, and S.A. Kay. Mammalian circadian signaling networks and therapeutic targets. *Nature Chemical Biology*, 3(10):630–639, October 2007.

[20] S.P. Fisher, R.G. Foster, and S.N. Peirson. The Circadian Control of Sleep. In *Circadian Clocks*, Handbook of Experimental Pharmacology, pages 157–183. January 2013.

[21] A. Kalsbeek, C.X. Yi, C. Cailotto, S. E. la Fleur, E. Fliers, and R.M. Buijs. Mammalian clock output mechanisms. *Essays in biochemistry*, 49(1):137–151, June 2011.

[22] A. Kalsbeek and E. Fliers. Daily regulation of hormone profiles. In *Circadian Clocks, Handbook of experimental pharmacology,*, number 217, pages 185–226. 2013.

[23] C.B. Green, J.S. Takahashi, and J. Bass. The Meter of Metabolism. *Cell*, 134(5):728–742, September 2008.

[24] J.S. Takahashi, H.K. Hong, C.H. Ko, and E.L. McDearmon. The genetics of mammalian circadian order and disorder: implications for physiology and disease. *Nature Reviews Genetics*, 9(10):764–775, October 2008.

[25] D.A. Bechtold, J.E. Gibbs, and A.S. Loudon. Circadian dysfunction in disease. *Trends in Pharmacological Sciences*, 31(5):191–198, May 2010.

[26] D.M. Arble, K.M. Ramsey, J. Bass, and F.W. Turek. Circadian Disruption and Metabolic Disease: Findings from Animal Models. *Best practice & research. Clinical endocrinology & metabolism*, 24(5):785–800, October 2010.

[27] D.K. Welsh, J.S. Takahashi, and S.A. Kay. Suprachiasmatic Nucleus: Cell Autonomy and Network Properties. *Annual Review of Physiology*, 72(1):551–577, 2010.

[28] C. Dibner, U. Schibler, and U. Albrecht. The Mammalian Circadian Timing System: Organization and Coordination of Central and Peripheral Clocks. *Annual Review of Physiology*, 72(1):517–549, 2010.

[29] R.Y. Moore and V.B. Eichler. Loss of a circadian adrenal corticosterone rhythm following suprachiasmatic lesions in the rat. *Brain research*, 42(1):201–206, July 1972.

[30] F.K. Stephan and I. Zucker. Circadian Rhythms in Drinking Behavior and Locomotor Activity of Rats Are Eliminated by Hypothalamic Lesions. *Proceedings of the National Academy of Sciences of the United States of America*, 69(6):1583–1586, June 1972.

[31] M.R. Ralph, R.G. Foster, F.C. Davis, and M. Menaker. Transplanted suprachiasmatic nucleus determines circadian period. *Science (New York, N.Y.)*, 247(4945):975–978, February 1990.

[32] J.A. Mohawk, C.B. Green, and J.S. Takahashi. Central and Peripheral Circadian Clocks in Mammals. *Annual Review of Neuroscience*, 35(1), 2012.

[33] F.K. Stephan. The "other" circadian system: food as a Zeitgeber. *Journal of biological rhythms*, 17(4):284–292, August 2002.

[34] A. Balsalobre, F. Damiola, and U. Schibler. A serum shock induces circadian gene expression in mammalian tissue culture cells. *Cell*, 93(6):929–937, June 1998.

[35] S. Yamazaki, R. Numano, M. Abe, A. Hida, R. Takahashi, M. Ueda, G.D. Block, Y. Sakaki, M. Menaker, and H. Tei. Resetting Central and Peripheral Circadian Oscillators in Transgenic Rats. *Science*, 288(5466):682–685, April 2000.

[36] R.A. Akhtar, A.B. Reddy, E.S. Maywood, J.D. Clayton, V.M. King, A.G. Smith, T.W. Gant, M.H. Hastings, and C.P. Kyriacou. Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus. *Current biology: CB*, 12(7):540–550, April 2002.

[37] K.F. Storch, O. Lipan, I. Leykin, N. Viswanathan, F.C. Davis, W.H. Wong, and C.J. Weitz. Extensive and divergent circadian gene expression in liver and heart. *Nature*, 417(6884):78–83, April 2002.

[38] S. Panda, M.P. Antoch, B.H. Miller, A.I. Su, A.B. Schook, M. Straume, P.G. Schultz, S.A. Kay, J.S. Takahashi, and J.B. Hogenesch. Coordinated Transcription of Key Pathways in the Mouse by the Circadian Clock. *Cell*, 109(3):307–320, May 2002.

[39] B.H. Miller, E.L. McDearmon, S. Panda, K.R. Hayes, J. Zhang, J.L. Andrews, M.P. Antoch, J.R. Walker, K.A. Esser, J.B. Hogenesch, and J.S. Takahashi. Circadian and CLOCK-controlled regulation of the mouse transcriptome and cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(9):3342–3347, February 2007.

[40] M.E. Hughes, L. DiTacchio, K.R. Hayes, C. Vollmers, S. Pulivarthy, Julie E. Baggs, S. Panda, and J.B. Hogenesch. Harmonics of Circadian Gene Transcription in Mammals. *PLoS Genetics*, 5(4), April 2009.

[41] A.B. Reddy, N.A. Karp, E.S. Maywood, E.A. Sage, M. Deery, J.S. O'Neill, G.K.Y. Wong, Jo Chesham, Mark Odell, K.S. Lilley, C.P. Kyriacou, and M.H. Hastings. Circadian Orchestration of the Hepatic Proteome. *Current Biology*, 16(11):1107–1115, June 2006.

[42] T. Tsuji, T. Hirota, N. Takemori, N. Komori, H. Yoshitane, M. Fukuda, H. Matsumoto, and Y. Fukada. Circadian proteomics of the mouse retina. *PROTEOMICS*, 7(19):3500 –3508, October 2007.

[43] M.J. Deery, E.S. Maywood, J.E. Chesham, M. Sládek, N.A. Karp, Edward W. Green, P.D. Charles, A.B. Reddy, C.P. Kyriacou, K.S. Lilley, and M.H. Hastings. Proteomic Analysis Reveals the Role of Synaptic Vesicle Cycling in Sustaining the Suprachiasmatic Circadian Clock. *Current Biology*, 19(23):2031–2036, December 2009.

[44] T.A. Martino, N. Tata, G.A. Bjarnason, M. Straume, and M.J. Sole. Diurnal protein expression in blood revealed by high throughput mass spectrometry proteomics and implications for translational medicine and body time of day. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 293(3):R1430–R1437, September 2007.

[45] M. Moller, T. Sparre, N. Bache, P. Roepstorff, and H. Vorum. Proteomic analysis of day–night variations in protein levels in the rat pineal gland. *PROTEOMICS*, 7(12):2009–2018, June 2007.

[46] S.-H. Yoo, S. Yamazaki, P.L. Lowrey, K. Shimomura, C.H. Ko, E.D. Buhr, S.M. Siepka, H.Ky. Hong, W. J. Oh, O. J. Y.oo, M. Menaker, and J. S. Takahashi. PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proceedings of the National Academy of Sciences of the United States of America*, 101(15):5339–5346, April 2004.

[47] Benoît Kornmann, Olivier Schaad, Hermann Bujard, Joseph S Takahashi, and Ueli Schibler. System-Driven and Oscillator-Dependent Circadian Transcription in Mice with a Conditionally Active Liver Clock. *PLoS Biol*, 5(2):e34, January 2007.

[48] K.A. Lamia, K.F.. Storch, and C.J. Weitz. Physiological significance of a peripheral tissue circadian clock. *Proceedings of the National Academy of Sciences of the United States of America*, 105(39):15172–15177, September 2008.

[49] J.C. Dunlap. Molecular Bases for Circadian Clocks. *Cell*, 96(2):271–290, January 1999.

[50] Steven M. Reppert and David R. Weaver. Coordination of circadian timing in mammals. *Nature*, 418(6901):935–941, August 2002.

[51] M. Gallego and D.M. Virshup. Post-translational modifications regulate the ticking of the circadian clock. *Nature Reviews Molecular Cell Biology*, 8(2):139–148, February 2007.

[52] L.P. Shearman, S. Sriram, D.R. Weaver, E.S. Maywood, I. Chaves, B. Zheng, K. Kume, C.C. Lee, G.T. J. van der Horst, M.H. Hastings, and S.M. Reppert. Interacting Molecular Loops in the Mammalian Circadian Clock. *Science*, 288(5468):1013–1019, December 2000.

[53] N. Preitner, F. Damiola, L. Lopez-Molina, J. Zakany, D. Duboule, U. Albrecht, and U. Schibler. The orphan nuclear receptor REV-ERBalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell*, 110(2):251–260, July 2002.

[54] S. Reischl and A. Kramer. Kinases and phosphatases in the mammalian circadian clock. *FEBS Letters*, 585(10):1393–1399, May 2011.

[55] T. Hirota, J.W. Lee, W.G. Lewis, E.E. Zhang, G. Breton, X. Liu, M. Garcia, E.C. Peters, J.-P. Etchegaray, D. Traver, P.G. Schultz, and S.A. Kay. High-throughput chemical screen identifies a novel potent modulator of cellular circadian rhythms and reveals CKIα as a clock regulatory kinase. *PLoS biology*, 8(12):e1000559, 2010.

[56] B. Maier, S. Wendt, J.T. Vanselow, T. Wallach, S. Reischl, S. Oehmke, A. Schlosser, and A. Kramer. A large-scale functional RNAi screen reveals a role for CK2 in the mammalian circadian clock. *Genes & Development*, 23(6):708–718, March 2009.

[57] E.E. Zhang, A.C. Liu, T. Hirota, L.J. Miraglia, G. Welch, P.Y. Pongsawakul, X. Liu, A. Atwood, J.W.3rd Huss, J. Janes, A.I. Su, J.B. Hogenesch, and S.A. Kay. A genome-wide RNAi screen for modifiers of the circadian clock in human cells. *Cell*, 139(1):199–210, October 2009.

[58] T. Hirota, W.G Lewis, A.C Liu, J.W. Lee, P.G. Schultz, and S.A. Kay. A chemical biology approach reveals period shortening of the mammalian circadian clock by specific inhibition of GSK-3beta. *Proceedings of the National Academy of Sciences of the United States of America*, 105(52):20746–20751, December 2008.

[59] A. Hirano, K. Yumimoto, R. Tsunematsu, M. Matsumoto, M. Oyama, H. Kozuka-Hata, T. Nakagawa, D. Lanjakornsiripan, K.I. Nakayama, and Y. Fukada. FBXL21 Regulates Oscillation of the Circadian Clock through Ubiquitination and Stabilization of Cryptochromes. *Cell*, 152(5):1106–1118, February 2013.

[60] S.M. Siepka, S.-H. Yoo, J. Park, W. Song, V. Kumar, Y. Hu, C. Lee, and J.S. Takahashi. Circadian Mutant Overtime Reveals F-box Protein FBXL3 Regulation of Cryptochrome and Period Gene Expression. *Cell*, 129(5):1011–1023, June 2007.

[61] S.I.H. Godinho, E.S. Maywood, L. Shaw, V. Tucci, A.R. Barnard, L. Busino, M. Pagano, R. Kendall, M.M. Quwailid, M.R. Romero, J. O'neill, J.E. Chesham, D. Brooker, Z. Lalanne, M.H. Hastings, and P.M. Nolan. The after-hours mutant reveals a role for Fbxl3 in determining mammalian circadian period. *Science (New York, N.Y.)*, 316(5826):897–900, May 2007.

[62] L. Busino, F. Bassermann, A. Maiolica, C. Lee, P.M. Nolan, S.I.H. Godinho, G.F. Draetta, and M. Pagano. SCFFbxl3 controls the oscillation of the circadian clock by directing the degradation of cryptochrome proteins. *Science (New York, N.Y.)*, 316(5826):900–904, May 2007.

[63] B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach. Global quantification of mammalian gene expression control. *Nature*, 473(7347):337–342, May 2011.

[64] G.M. Cooper. Protein Degradation. In *The Cell: A Molecular Approach.* 2nd edition, 2000.

[65] X. Zheng and A. Sehgal. Speed control: cogs and gears that drive the circadian clock. *Trends in neurosciences*, 2012.

[66] M. R. Ralph and M. Menaker. A mutation of the circadian system in golden hamsters. *Science*, 241(4870):1225–1227, February 1988.

[67] P.L. Lowrey, K. Shimomura, M.P. Antoch, S. Yamazaki, P.D. Zemenides, M.R. Ralph, M. Menaker, and J.S. Takahashi. Positional syntenic cloning and functional characterization of the mammalian circadian mutation tau. *Science (New York, N.Y.)*, 288(5465):483–492, April 2000.

[68] E.J. Eide, M.F. Woolf, H. Kang, P. Woolf, W. Hurst, F. Camacho, E.L. Vielhaber, A. Giovanni, and D.M. Virshup. Control of mammalian circadian rhythm by CKIepsilon-regulated proteasome-mediated PER2 degradation. *Molecular and cellular biology*, 25(7):2795–2807, April 2005.

[69] C.A. Lamb, H.C. Dooley, and S.A. Tooze. Endocytosis and autophagy: Shared machinery for degradation. *BioEssays*, 35(1):34–45, 2013.

[70] M.J. Clague and S. Urbé. Ubiquitin: Same Molecule, Different Degradation Pathways. *Cell*, 143(5):682–685, November 2010.

[71] N. Mizushima, Y. Ohsumi, and T. Yoshimori. Autophagosome Formation in Mammalian Cells. *Cell Structure and Function*, 27(6):421–429, 2002.

[72] M.H. Glickman and A. Ciechanover. The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiological Reviews*, 82(2):373–428, January 2002.

[73] K. Vanselow, J.T. Vanselow, P.O. Westermark, S. Reischl, B. Maier, T. Korte, A. Hermann, H. Herzel, A. Schlosser, and A. Kramer. Differential effects of PER2 phosphorylation: molecular basis for the human familial advanced sleep phase syndrome (FASPS). *Genes & Development*, 20(19):2660–2672, October 2006.

[74] C.T. Tang, S. Li, C. Long, J. Cha, G. Huang, L. Li, S. Chen, and Y. Liu. Setting the pace of the Neurospora circadian clock by multiple independent FRQ phosphorylation events. *Proceedings of the National Academy of Sciences*, page pnas.0904898106, June 2009.

[75] Y. Tsuchiya, M. Akashi, M. Matsuda, K. Goto, Y. Miyata, K. Node, and E. Nishida. Involvement of the protein kinase CK2 in the regulation of mammalian circadian rhythms. *Science signaling*, 2(73):ra26, 2009.

[76] M. Gallego, H. Kang, and D.M. Virshup. Protein phosphatase 1 regulates the stability of the circadian protein PER2. *The Biochemical journal*, 399(1):169–175, October 2006.

[77] I. Schmutz, S. Wendt, A. Schnell, A. Kramer, I.M. Mansuy, and U. Albrecht. Protein Phosphatase 1 (PP1) Is a Post-Translational Regulator of the Mammalian Circadian Clock. *PLoS ONE*, 6(6):e21325, June 2011.

[78] D.M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef. Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science*, 332(6028):472 –474, April 2011.

[79] Y. Fujimoto, K. Yagita, and H. Okamura. Does mPER2 protein oscillate without its coding mRNA cycling?: post-transcriptional regulation by cell clock. *Genes to Cells*, 11(5):525–530, 2006.

[80] C. Lee, J.-P. Etchegaray, F.R.A. Cagampang, A.S.I. Loudon, and S.M. Reppert. Posttranslational Mechanisms Regulate the Mammalian Circadian Clock. *Cell*, 107(7):855–867, December 2001.

[81] D. Gatfield and U. Schibler. Proteasomes Keep the Circadian Clock Ticking. *Science*, 316(5828):1135–1136, May 2007.

[82] N. Kurabayashi, T. Hirota, M. Sakai, K. Sanada, and Y. Fukada. DYRK1A and glycogen synthase kinase 3beta, a dual-kinase mechanism directing proteasomal degradation of CRY2 for circadian timekeeping. *Molecular and Cellular Biology*, 30(7):1757–1768, April 2010.

[83] K.A. Lamia, U.M. Sachdeva, L. DiTacchio, E.C. Williams, J.G. Alvarez, D.F. Egan, D.S. Vasquez, H. Juguilon, S. Panda, R.J. Shaw, C.B. Thompson, and R.M. Evans. AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation. *Science (New York, N.Y.)*, 326(5951):437–440, October 2009.

[84] S.-H. Yoo, J.A. Mohawk, S.M. Siepka, Y. Shan, S.K. Huh, H.-K. Hong, I. Kornblum, V. Kumar, N. Koike, M. Xu, J. Nussbaum, X. Liu, Z. Chen, Z.J. Chen, C.B. Green, and J.S. Takahashi. Competing E3 Ubiquitin Ligases Govern Circadian Periodicity by Degradation of CRY in Nucleus and Cytoplasm. *Cell*, 152(5):1091–1105, February 2013.

[85] T. Tamaru, Y. Isojima, G.T. J. van der Horst, K. Takei, K. Nagai, and K. Takamatsu. Nucleocytoplasmic shuttling and phosphorylation of BMAL1 are regulated by circadian clock in cultured fibroblasts. *Genes to Cells*, 8(12):973–983, 2003.

[86] H. Yoshitane, T. Takao, Y. Satomi, N.-H. Du, T. Okano, and Y. Fukada. Roles of CLOCK phosphorylation in suppression of E-box-dependent transcription. *Molecular and cellular biology*, 29(13):3675–3686, July 2009.

[87] E.J. Eide, E.L. Vielhaber, W.A. Hinz, and D.M. Virshup. The circadian regulatory proteins BMAL1 and cryptochromes are substrates of casein Kinase Iepsilon(CKIepsilon). *Journal of Biological Chemistry*, March 2002.

[88] T. Tamaru, J. Hirayama, Y. Isojima, K. Nagai, S. Norioka, K. Takamatsu, and P. Sassone-Corsi. CK2α phosphorylates BMAL1 to regulate the mammalian clock. *Nature Structural & Molecular Biology*, 16(4):446–448, 2009.

[89] K. Sanada, T. Okano, and Y. Fukada. Mitogen-activated Protein Kinase Phosphorylates and Negatively Regulates Basic Helix-Loop-Helix-PAS Transcription Factor BMAL1. *Journal of Biological Chemistry*, 277(1):267–271, April 2002.

[90] S. Sahar, L. Zocchi, C. Kinoshita, E. Borrelli, and P. Sassone-Corsi. Regulation of BMAL1 Protein Stability and Circadian Function by GSK3β-Mediated Phosphorylation. *PLoS ONE*, 5(1):e8561, January 2010.

[91] M.L. Spengler, K.K. Kuropatwinski, M. Schumer, and M. Antoch. A serine cluster mediates BMAL1-dependent CLOCK phosphorylation and degradation. *Cell Cycle*, 8(24):4138–4146, December 2009.

[92] H.S. Shim, H. Kim, J. Lee, G.H. Son, S. Cho, T.H. Oh, S.H. Kang, D.-S. Seen, K.H. Lee, and K. Kim. Rapid activation of CLOCK by Ca2+-dependent protein kinase C mediates resetting of the mammalian circadian clock. *EMBO reports*, 8(4):366–371, April 2007.

[93] J. Lee, Y. Lee, M.J. Lee, E. Park, S.H. Kang, C.H. Chung, K.H. Lee, and K. Kim. Dual Modification of BMAL1 by SUMO2/3 and Ubiquitin Promotes Circadian Activation of the CLOCK/BMAL1 Complex. *Molecular and Cellular Biology*, 28(19):6056–6065, October 2008.

[94] M.-D. Li, H.-B. Ruan, M.E. Hughes, J.-S. Lee, J.P. Singh, S.P. Jones, M.N. Nitabach, and X. Yang. O-GlcNAc Signaling Entrains the Circadian Clock by Inhibiting BMAL1/CLOCK Ubiquitination. *Cell metabolism*, 17(2):303–310, February 2013.

[95] Y.-T. Ma, H. Luo, W.-J. Guan, H. Zhang, C. Chen, Z. Wang, and J.D. Li. O-GlcNAcylation of BMAL1 regulates circadian rhythms in NIH3T3 fibroblasts. *Biochemical and biophysical research communications*, 431(3):382–387, February 2013.

[96] I. Kwon, J. Lee, S.H. Chang, N.C. Jung, B.J. Lee, G.H. Son, K. Kim, and K.H. Lee. BMAL1 Shuttling Controls Transactivation and Degradation of the CLOCK/BMAL1 Heterodimer. *Molecular and Cellular Biology*, 26(19):7318–7330, January 2006.

[97] M. Muratani and W.P. Tansey. How the ubiquitin–proteasome system controls transcription. *Nature Reviews Molecular Cell Biology*, 4(3):192–201, March 2003.

[98] N. Koike, S.-H. Yoo, H.-C. Huang, V. Kumar, C. Lee, T.-K. Kim, and J.S. Takahashi. Transcriptional Architecture and Chromatin Landscape of the Core Circadian Clock in Mammals. *Science*, 338(6105):349–354, October 2012.

[99] S. Kojima, El.L. Sher-Chen, and C.B. Green. Circadian control of mRNA polyadenylation dynamics regulates rhythmic protein expression. *Genes & Development*, 26(24):2724–2736, December 2012.

[100] D. Ma, S. Panda, and J.D. Lin. Temporal orchestration of circadian autophagy rhythm by C/EBPbeta. *The EMBO Journal*, 30(22):4642–4651, November 2011.

[101] G.E. Duffield, J.D. Best, B.H. Meurers, A. Bittner, J.J. Loros, and J.C. Dunlap. Circadian Programs of Transcriptional Activation, Signaling, and Protein Turnover Revealed by Microarray Analysis of Mammalian Cells. *Current Biology*, 12(7):551–557, April 2002.

[102] M. Horiguchi, S. Koyanagi, A.M. Hamdan, K. Kakimoto, N. Matsunaga, C. Yamashita, and S. Ohdo. Rhythmic Control of the ARF-MDM2 Pathway by ATF4 Underlies Circadian Accumulation of p53 in Malignant Cells. *Cancer Research*, 73(8):2639–2649, April 2013.

[103] U. Stelzl, U. Worm, M. Lalowski, C. Haenig, F.H. Brembeck, H. Goehler, M. Stroedicke, M. Zenkner, A. Schoenherr, S. Koeppen, J. Timm, S. Mintzlaff, C. Abraham, N. Bock, S. Kietzmann, A. Goedde, E. Toksöz, A. Droege, S. Krobitsch, B. Korn, W. Birchmeier, H. Lehrach, and E.E. Wanker. A Human Protein-Protein Interaction Network: A Resource for Annotating the Proteome. *Cell*, 122(6):957–968, September 2005.

[104] E.V. McCarthy, J.E. Baggs, J.M. Geskes, J.B. Hogenesch, and C.B. Green. Generation of a Novel Allelic Series of Cryptochrome Mutants via Mutagenesis Reveals Residues Involved in Protein-Protein Interaction and CRY2-Specific Repression. *Molecular and Cellular Biology*, 29(20):5465–5476, October 2009.

[105] S. Reischl, K. Vanselow, P.O. Westermark, N. Thierfelder, B. Maier, H. Herzel, and A. Kramer. β-TrCP1-Mediated Degradation of PERIOD2 Is Essential for Circadian Dynamics. *Journal of Biological Rhythms*, 22(5):375–386, January 2007.

[106] H.-C.S. Yen, Q. Xu, D.M. Chou, Z. Zhao, and S.J. Elledge. Global Protein Stability Profiling in Mammalian Cells. *Science*, 322(5903):918–923, July 2008.

[107] M.J. Emanuele, A.E.H. Elia, Q. Xu, C.R. Thoma, L. Izhar, Y. Leng, A. Guo, Y.-N. Chen, J. Rush, P.W.-C. Hsu, H.-C.S. Yen, and S.J. Elledge. Global Identification of Modular Cullin-RING Ligase Substrates. *Cell*, 147(2):459–474, October 2011.

[108] A.C. Zambon, S. Gaj, I. Ho, K. Hanspers, K. Vranizan, C.T. Evelo, B.R. Conklin, A.R. Pico, and N. Salomonis. GO-Elite: a flexible solution for pathway and ontology over-representation. *Bioinformatics*, 28(16):2209–2210, August 2012.

[109] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor symposia on quantitative biology*, 51 Pt 1:263–273, 1986.

[110] G.K. Smyth, R. Gentleman, V.J. Carey, W. Huber, R.A. Irizarry, and S. Dudoit. limma: Linear Models for Microarray Data. Statistics for Biology and Health, pages 397–420. Springer New York, January 2005.

[111] V.V. Verkhusha, I.M. Kuznetsova, O.V. Stepanenko, A.G. Zaraisky, M.M. Shavlovsky, K.K. Turoverov, and V.N. Uversky. High stability of Discosoma DsRed as compared to Aequorea EGFP. *Biochemistry*, 42(26):7879–7884, July 2003.

[112] Y.A. Bochkov and A.C. Palmenberg. Translational efficiency of EMCV IRES in bicistronic vectors is dependent upon IRES sequence and gene location. *Biotechniques*, 41(3):283–284, 286, 288 passim, September 2006.

[113] A.G. Evdokimov, M.E. Pokross, N.S. Egorov, A.G. Zaraisky, I.V. Yampolsky, E.M. Merzlyak, A.N. Shkoporov, I. Sander, K.A. Lukyanov, and D.M. Chudakov. Structural basis for the fast maturation of Arthropoda green fluorescent protein. *EMBO Reports*, 7(10):1006–1012, October 2006.

[114] B.J. Bevis and B.S. Glick. Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed). *Nature Biotechnology*, 20(1):83–87, January 2002.

[115] S.R. Kain. Green fluorescent protein (GFP): applications in cell-based assays for drug discovery. *Drug Discovery Today*, 4(7):304–312, July 1999.

[116] X. Li, X. Zhao, Y. Fang, X. Jiang, T. Duong, C. Fan, C.-C. Huang, and S.R. Kain. Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter. *Journal of Biological Chemistry*, 273(52):34970–34975, December 1998.

[117] T. Wallach, K. Schellenberg, B. Maier, R.K.R. Kalathur, P. Porras, E.E. Wanker, M.E. Futschik, and A. Kramer. Dynamic Circadian Protein–Protein Interaction Networks Predict Temporal Organization of Cellular Functions. *PLoS Genet*, 9(3):e1003398, March 2013.

[118] P. Lamesch, N. Li, S. Milstein, C. Fan, T. Hao, G. Szabo, Z. Hu, K. Venkatesan, G. Bethel, P. Martin, J. Rogers, S. Lawlor, S. McLaren, A. Dricot, H. Borick, M.E. Cusick, J. Vandenhaute, I. Dunham, D.E. Hill, and M. Vidal. hORFeome v3.1: A resource of human open reading frames representing over 10,000 human genes. *Genomics*, 89(3):307–315, March 2007.

[119] J.-F. Rual, T. Hirozane-Kishikawa, T. Hao, N. Bertin, S. Li, A. Dricot, N. Li, J. Rosenberg, P. Lamesch, P.-O. Vidalain, T.R. Clingingsmith, J.L. Hartley, D. Esposito, D. Cheo, T. Moore, B. Simmons, R. Sequerra, S. Bosak, L. Doucette-Stamm, C. Le Peuch, J. Vandenhaute, M.E. Cusick, J.S. Albala, D.E. Hill, and M. Vidal. Human ORFeome Version 1.1: A Platform for Reverse Proteomics. *Genome Research*, 14(10b):2128–2135, October 2004.

[120] D.S. Gerhard, L. Wagner, E.A. Feingold, C.M. Shenmen, L.H. Grouse, G. Schuler, S.L. Klein, S. Old, R. Rasooly, P. Good, M. Guyer, A.M. Peck, J.G. Derge, D. Lipman, F.S. Collins, W. Jang, S. Sherry, M. Feolo, L. Misquitta, E. Lee, K. Rotmistrovsky, S.F. Greenhut, C.F. Schaefer, K. Buetow, T.I. Bonner, D. Haussler, J. Kent, M. Kiekhaus, T. Furey, M. Brent, C. Prange, K. Schreiber, N. Shapiro, N.K. Bhat, R.F. Hopkins, F. Hsie, T. Driscoll, M.B. Soares, T.L. Casavant, T.E. Scheetz, M.J. Brownstein, T.B. Usdin, S. Toshiyuki, P. Carninci, Y. Piao, D.B. Dudekula, M.S.H. Ko, K. Kawakami, Y. Suzuki, S. Sugano, C.E. Gruber, M.R. Smith, B. Simmons, T. Moore, R. Waterman, S.L. Johnson, Y. Ruan, C.L. Wei, S. Mathavan, P.H. Gunaratne, J. Wu, A.M. Garcia, S.W. Hulyk, E. Fuh, Y. Yuan, A. Sneed, C. Kowis, A. Hodgson, D.M. Muzny, J. McPherson, R.A. Gibbs, J. Fahey, E. Helton, M. Ketteman, A. Madan, S. Rodrigues, A. Sanchez, M. Whiting, A. Madari, A.C. Young, K.D. Wetherby, S.J. Granite, P.N. Kwong, C.P. Brinkley, R.L. Pearson, G.G. Bouffard, R.W. Blakesly, E.D. Green, M.C. Dickson, A.C. Rodriguez, J. Grimwood, J. Schmutz, R.M. Myers, Y.S.N. Butterfield, M. Griffith, O.L. Griffith, M.I. Krzywinski, N. Liao, R. Morin, R. Morrin, D. Palmquist, A.S. Petrescu, U. Skalska, D.E. Smailus, J.M. Stott, A. Schnerch, J.E. Schein, S.J.M. Jones, R.A. Holt, A. Baross, M.A. Marra, S. Clifton, K.A. Makowski, S. Bosak, and J. Malek. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome research*, 14(10B):2121–2127, October 2004.

[121] G. Temple, D.S. Gerhard, R. Rasooly, E.A. Feingold, P.J. Good, C. Robinson, A. Mandich, J.G. Derge, J. Lewis, D. Shoaf, F.S. Collins, W. Jang, L. Wagner, C.M. Shenmen, L. Misquitta, C.F. Schaefer, K.H. Buetow, T.I. Bonner, L. Yankie, M. Ward, L. Phan, Al. Astashyn, G. Brown, C. Farrell, J. Hart, M. Landrum, B.L. Maidak, M. Murphy, T. Murphy, B. Rajput, L. Riddick, D. Webb, J. Weber, W. Wu, K.D. Pruitt, D. Maglott, A. Siepel, B. Brejova, M. Diekhans, R. Harte, R. Baertsch, J. Kent, D. Haussler, M. Brent, L. Langton, C.L.G. Comstock, M. Stevens, C. Wei, M.J. van Baren, K. Salehi-Ashtiani, R.R. Murray, L. Ghamsari, E. Mello, C. Lin, C. Pennacchio, K. Schreiber, N. Shapiro, A. Marsh, E. Pardes, T. Moore, A. Lebeau, M. Muratet, B. Simmons, D. Kloske, S. Sieja, J. Hudson, P. Sethupathy, M. Brownstein, N. Bhat, J. Lazar, H. Jacob, C.E. Gruber, M.R. Smith, J. McPherson, A.M. Garcia, P.H. Gunaratne, J. Wu, D. Muzny, R.A. Gibbs, A.C. Young, G.G. Bouffard, R.W. Blakesley, J. Mullikin, E.D. Green, M.C. Dickson, A.C. Rodriguez, J. Grimwood, J. Schmutz, R.M. Myers, M. Hirst, T. Zeng, K. Tse, M. Moksa, M. Deng, K. Ma, D. Mah, J. Pang, G. Taylor, E. Chuah, A. Deng, K. Fichter, A. Go, S. Lee, J. Wang, M. Griffith, R. Morin, R.A. Moore, M. Mayo, S. Munro, S. Wagner, S.J.M. Jones, R.A. Holt, M.A. Marra, S. Lu, S. Yang, J. Hartigan, M. Graf, R. Wagner, S. Letovksy, J.C. Pulido, K. Robison, D. Esposito, J. Hartley, V.E. Wall, R.F. Hopkins, O. Ohara, and S. Wiemann. The completion of the Mammalian Gene Collection (MGC). *Genome Research*, 19(12):2324–2333, January 2009.

[122] Y. Lee, R. Chen, H. Lee, and C. Lee. Stoichiometric Relationship among Clock Proteins Determines Robustness of Circadian Rhythms. *Journal of Biological Chemistry*, 286(9):7033–7042, April 2011.

[123] K. Kume, M.J. Zylka, S. Sriram, L.P. Shearman, D.R. Weaver, X. Jin, E.S. Maywood, M.H. Hastings, and S.M. Reppert. mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell*, 98(2):193–205, July 1999.

[124] A. Czarna, A. Berndt, H.R. Singh, A. Grudziecki, A.G. Ladurner, G. Timinszky, A. Kramer, and E. Wolf. Structures of Drosophila Cryptochrome and Mouse Cryptochrome1 Provide Insight into Circadian Function. *Cell*, 153(6):1394–1405, June 2013.

[125] Kazuhiro Yagita, Filippo Tamanini, Maya Yasuda, Jan H J Hoeijmakers, Gijsbertus T J van der Horst, and Hitoshi Okamura. Nucleocytoplasmic shuttling and mCRY-dependent inhibition of ubiquitylation of the mPER2 clock protein. *The EMBO Journal*, 21(6):1301–1314, March 2002.

[126] F. Spörl, K. Schellenberg, T. Blatt, H. Wenck, K.P. Wittern, A. Schrader, and A. Kramer. A Circadian Clock in HaCaT Keratinocytes. *Journal of Investigative Dermatology*, 131(2):338–348, 2011.

[127] C. Jouffe, G. Cretenet, L. Symul, E. Martin, F. Atger, F. Naef, and F. Gachon. The Circadian Clock Coordinates Ribosome Biogenesis. *PLoS Biol*, 11(1):e1001455, January 2013.

[128] F.X.R. Sutandy, J. Qian, C.-S. Chen, and H. Zhu. Overview of protein microarrays. *Current protocols in protein science / editorial board, John E. Coligan*, Chapter 27:Unit 27.1, April 2013.

[129] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, and B. Kuster. Quantitative mass spectrometry in proteomics: a critical review. *Analytical and Bioanalytical Chemistry*, 389(4):1017–1031, 2007.

[130] M.S. Robles and M. Mann. Proteomic approaches in circadian biology. In *Handbook of experimental pharmacology*, volume 217, pages 389–407. 2013.

[131] X. Jiang, P. Coffino, and X. Li. Development of a method for screening short-lived proteins using green fluorescent protein. *Genome Biology*, 5(10):R81, 2004.

[132] T. Schneider-Poetsch, J. Ju, D.E. Eyler, Y. Dang, S. Bhat, W.C. Merrick, R. Green, B. Shen, and J.O. Liu. Inhibition of Eukaryotic Translation Elongation by Cycloheximide and Lactimidomycin. *Nature chemical biology*, 6(3):209–217, March 2010.

[133] M.K. Doherty, D.E. Hammond, M.J. Clague, S.J. Gaskell, and R.J. Beynon. Turnover of the Human Proteome: Determination of Protein Intracellular Stability by Dynamic SILAC. *Journal of Proteome Research*, 8(1):104–112, January 2009.

[134] A. Varshavsky. The N-end rule pathway and regulation by proteolysis. *Protein Science : A Publication of the Protein Society*, 20(8):1298–1345, August 2011.

[135] C.-S. Hwang, A. Shemorry, and A. Varshavsky. N-Terminal Acetylation of Cellular Proteins Creates Specific Degradation Signals. *Science*, 327(5968):973–977, February 2010.

[136] J.E. Baggs, T.S. Price, L. DiTacchio, S. Panda, G.A. FitzGerald, and J.B. Hogenesch. Network Features of the Mammalian Circadian Clock. *PLoS Biology*, 7(3), March 2009.

[137] A. Relógio, P.O. Westermark, T. Wallach, K. Schellenberg, A. Kramer, and H. Herzel. Tuning the Mammalian Circadian Clock: Robust Synergy of Two Loops. *PLoS Computational Biology*, 7(12), December 2011.

[138] M.K. Bunger, L.D. Wilsbacher, S.M. Moran, C. Clendenin, L.A. Radcliffe, J.B. Hogenesch, M.C. Simon, J.S. Takahashi, and C.A. Bradfield. Mop3 Is an Essential Component of the Master Circadian Pacemaker in Mammals. *Cell*, 103(7):1009–1017, December 2000.

[139] I. Yamanaka, S. Koinuma, Y. Shigeyoshi, Y. Uchiyama, and K. Yagita. Presence of robust circadian clock oscillation under constitutive over-expression of mCry1 in rat-1 fibroblasts. *FEBS letters*, 581(21):4098–4102, August 2007.

[140] Y. Fan, A. Hida, D.A. Anderson, M. Izumo, and C.H. Johnson. Cycling of CRYPTOCHROME proteins is not necessary for circadian-clock function in mammalian fibroblasts. *Current biology: CB*, 17(13):1091–1100, July 2007.

[141] G. Soler, J.M. Bautista, J.A. Madrid, and G.M. Salido. Circadian rhythms in enzymatic activity of rat liver arginase and glucose 6-phosphate dehydrogenase. *Chronobiologia*, 15(3):205–212, September 1988.

[142] M. Aramendy, S. Seibert, P. Treppmann, K. Richter, G. Ahnert-Hilger, and U. Albrecht. Synaptophysin is involved in resetting of the mammalian circadian clock. *Journal of Circadian Rhythms*, 11(1):11, October 2013.

[143] S. Punia, K.K. Rumery, E.A. Yu, C.M. Lambert, A.L. Notkins, and D.R. Weaver. Disruption of gene expression rhythms in mice lacking secretory vesicle proteins IA-2 and IA-2ß. *American Journal of Physiology - Endocrinology and Metabolism*, 303(6):E762–E776, September 2012.

[144] F. Stricher, C. Macri, M. Ruff, and S. Muller. HSPA8/HSC70 chaperone protein: Structure, function, and chemical targeting. *Autophagy*, 9(12), October 2013.

[145] L. Fu, H. Pelicano, J. Liu, P. Huang, and C.C. Lee. The Circadian Gene Period2 Plays an Important Role in Tumor Suppression and DNA Damage Response In Vivo. *Cell*, 111(1):41–50, October 2002.

[146] S. Gery, N. Komatsu, L. Baldjyan, A. Yu, D. Koo, and H.P. Koeffler. The circadian gene per1 plays an important role in cell growth and DNA damage control in human cancer cells. *Molecular cell*, 22(3):375–382, May 2006.

[147] M.V. Plikus, C. Vollmers, D. de la Cruz, A. Chaix, R. Ramos, S. Panda, and C.-M. Chuong. Local circadian clock gates cell cycle progression of transient amplifying cells during regenerative hair cycling. *Proceedings of the National Academy of Sciences of the United States of America*, 110(23):E2106–2115, June 2013.

[148] F. Lévi, E. Filipski, I. Iurisci, X.M. Li, and P. Innominato. Cross-talks between circadian timing system and cell division cycle determine cancer biology and therapeutics. *Cold Spring Harbor symposia on quantitative biology*, 72:465–475, 2007.

[149] T. Hunt and P. Sassone-Corsi. Riding tandem: circadian clocks and the cell cycle. *Cell*, 129(3):461–464, May 2007.

[150] E. Nagoshi, C. Saini, C. Bauer, T. Laroche, F. Naef, and U. Schibler. Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells. *Cell*, 119(5):693–705, November 2004.

[151] M. Yeom, J.S. Pendergast, Y. Ohmiya, and S. Yamazaki. Circadian-independent cell mitosis in immortalized fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*, 107(21):9665–9670, May 2010.

[152] T. Matsuo, S. Yamaguchi, S. Mitsui, A. Emi, F. Shimoda, and H. Okamura. Control mechanism of the circadian clock for timing of cell division in vivo. *Science (New York, N.Y.)*, 302(5643):255–259, October 2003.

[153] E. Kowalska, J.A. Ripperger, D.C. Hoegger, P. Bruegger, T. Buch, T. Birchler, A. Mueller, U. Albrecht, C. Contaldo, and S.A. Brown. Feature Article: NONO couples the circadian clock to the cell cycle. *Proceedings of the National Academy of Sciences of the United States of America*, December 2012.

[154] A. Gréchez-Cassiau, B. Rayet, F. Guillaumond, M. Teboul, and F. Delaunay. The circadian clock component BMAL1 is a critical regulator of p21WAF1/CIP1 expression and hepatocyte proliferation. *The Journal of biological chemistry*, 283(8):4535–4542, February 2008.

[155] K. Ünsal Kaçmaz, T.E. Mullen, W.K. Kaufmann, and A. Sancar. Coupling of Human Circadian and Cell Cycles by the Timeless Protein. *Molecular and Cellular Biology*, 25(8):3109–3116, April 2005.

[156] V. Janssens and J. Goris. Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochemical Journal*, 353(Pt 3):417–439, February 2001.

[157] M. Le Breton, P. Cormier, R. Bellé, O. Mulner-Lorillon, and J. Morales. Translational control during mitosis. *Biochimie*, 87(9–10):805–811, September 2005.

[158] C.-H. Johnson. Circadian clocks and cell division. *Cell Cycle*, 9(19):3864–3873, October 2010.

List of Publications

T. Wallach, **K. Schellenberg**, B. Maier, R.K.R. Kalathur, P. Porras, E.E. Wanker, M.E. Futschik, and A. Kramer. Dynamic Circadian Protein-Protein Interaction Networks Predict Temporal Organization of Cellular Functions. *PLoS Genet*, 9(3):e1003398, 2013.

F. Spörl, S. Korge, K. Jürchott, M. Wunderskirchner, **K. Schellenberg**, S. Heins, A. Specht, C. Stoll, R. Klemz, B. Maier, H. Wenck, A. Schrader, D. Kunz, T. Blatt, and A. Kramer. Krüppel-like factor 9 is a circadian transcription factor in human epidermis that controls proliferation of keratinocytes. *Proceedings of the National Academy of Sciences*, 109(27):10903-0908, 2012.

A. Relógio, P.O. Westermark, T. Wallach, **K. Schellenberg**, A. Kramer, and H. Herzel. Tuning the Mammalian Circadian Clock: Robust Synergy of Two Loops. *PLoS Computational Biology*, 7(12), 2011.

K. Schellenberg, F. Spörl, T. Blatt, H. Wenck, K.P. Wittern, A. Schrader, and A. Kramer. A Circadian Clock in HaCaT Keratinocytes. *Journal of Investigative Dermatology*, 131(2):338-48, 2011.

Abbreviations

βGAL .... β-galactosidase
βTrCP .... F-box protein beta-transducin repeat containing protein
Δ .... difference
μ .... micro
σ .... standard deviation
lacZ .... gene of β-galactosidase
A .... area (stands for whole cell area analyzed in flow cytometry)
ad .... *latin* to
Amp .... Ampicillin
Ampk .... adenosine monophosphate-activated protein kinase
ampli. .... amplification
Ap4m1 .... Adaptor-related protein complex 4, mu 1 subunit
aq. dest .... *latin* aqua destillata
Arfgap3 .... ADP-ribosylation factor GTPase activating protein 3
Arg1 .... Arginase 1
ATP .... adenosine triphosphate
ATPase .... adenylpyrophosphatase
BCA .... bicinchoninic acid
Bhlhe40 .... basic helix-loop-helix family, member e40, also known as Dec1
Bhlhe41 .... basic helix-loop-helix family, member e41, also known as Dec2
Bla .... Blasticidin
Bmal1 .... ~ Arntl, Aryl Hydrocarbon Receptor Nuclear Translocator-like
BP .... Biological process
bp .... base pair
BSA .... bovine serum albumin
CC .... cellular component
CCG .... clock–controlled genes
cDNA .... complementary deoxyribonucleic acid
cent .... centered
Chl .... Chloramphenicol

CIP .... Calf Intestinal Alkaline Phosphatase
CK1/2 .... casein kinase 1/2
Clock .... Circadian Locomotor Output Cycles Kaput
Cq .... cycle in quantitative real–time PCR, based on regression analysis
Cry1/∼2 .... Cryptochrome 1/∼2
Cry1mut .... Cry1 mutant with single amino acid change G336D
CT .... circadian time
Cy3/5 .... cyanine dyes 3/5
D .... Aspartic Acid
d .... days
d1/ d4Egfp .... destabilized EGFP variants with either 1– or 4–hrs protein half–lifes
Da .... Dalton
det. .... detected
dex .... dexamethasone
DMEM .... Dulbecco´s modified Eagle Medium
DMSO .... dimethyl sulfoxide
DNA .... deoxyribonucleic acid
dNTP's .... deoxynucleotide triphosphates
DsRed .... *Discosoma species* red fluorescent protein
E6S .... six–times repeat of E–box motif fused to luciferase
EDTA .... ethylenediaminetetraacetic acid
Egfp .... enhanced green fluorescent protein
EtBr .... ethidium bromide
EtOH .... ethanol
euk. .... eukaryotic
FACS .... fluorescence activated cell sorting
Fads1 .... Fatty acid desaturase 1
FASPS .... familial advanced sleep phase syndrome
FBS .... fetal bovine serum
FITC-A .... Fluorescein isothiocyanate
frac .... fraction
FSC .... Forward Scatter
FW .... forward primer
G .... Glycine
GeneID .... gene identifier

GO .... Gene Ontology
GPS .... Global Protein Stability
gs .... gene symbol
Gsk3β .... glycogen synthase kinase-3 beta
h .... human
HBS .... HEPES–buffered Saline
HEK293 .... human embryonic cells 293
HEK293T .... human embryonic kidney cells expressing the SV40 large T-antigen
hr(s) .... hour(s)
Hspa8 .... Heat shock 70 kDa protein 8
int. .... intensity measured on array
IRES .... internal ribosome entry site
k .... kilo
Kana .... Kanamycin
kb .... kilobase
lacZ .... gene of β-galactosidase
LB .... Luria Broth
LUC .... luciferase
LV .... lentiviral
M .... molar
m .... murine
Mapk .... mitogen-activated protein kinase
MF .... Molecular function
MFI .... median fluorescence intensity
min .... minute
Mio. .... million
mM .... millimolar
MOI .... multiplicity of infection
Mr1 .... Major histocompatibility complex, class I-related
mRNA .... messenger RNA
n.d. .... not detected
n.s. .... not significant
Nde1 .... NudE neurodevelopment protein 1
NEB .... New England BioLabs
neg. .... negative

Neo .... Neomycin
norm. .... normalized
Npas2 .... Neuronal PAS domain protein 2
Npm1 .... Nucleophosmin (nucleolar phosphoprotein B23, numatrin)
Nr1d1 .... Nuclear Receptor Subfamily 1, Group D, Member 1 (alias Rev–Erbα)
Nr1d2 .... Nuclear Receptor Subfamily 1, Group D, Member 2 (alias Rev–Erbβ)
Nutf2 .... Nuclear transport factor 2
ORF .... open reading frame
OX .... overexpression
pCMV .... cytomegalovirus promoter
PCR .... polymerase chain reaction
PE .... Phycoerythrin
PEG .... Polyethylene glycol
Per1/$\sim$2 .... Period 1/$\sim$2
Per2βTrCP .... Per2 mutated in βTrCP binding site
Per2mut7 .... Per2 mutated in CK1ϵ/δ phosphorylation sites
PFA .... paraformaldehyde
Pih1d1 .... PIH1 domain containing 1
Plk1 .... Polo-like kinase 1
pos. .... positive
Pp2a .... Protein phosphatase 2
Ppp1ca .... protein phosphatase 1 alpha, catalytic subunit
Ppp2r1b .... Protein phosphatase 2, regulatory subunit A, beta
pres. .... present
prok. .... prokaryotic
ps .... probe set
PSI .... protein stability index (according to [106])
puro .... Puromycin
qRT–PCR .... quantitative real–time PCR
R .... statistical software
r. .... relative
r_s .... Spearman's Rank Correlation Coefficient
Ranbp3 .... RAN binding protein 3
rcf .... relative centrifugal force
reg. .... regulatory

rel.	relative
resis.	resistance
RNA	ribonucleic acid
Rorα	RAR-related orphan receptor A
rpm	rotations per minute
RT	room temperature
RV	reverse primer
SCN	suprachiasmatic nucleus
SD	standard deviation
SDS	sodium dodecyl sulfate
SDS–PAGE	SDS–Polyacrylamid Gel Electrophoresis
sec	seconds
Sec61A2	Sec61 alpha 2 subunit (S. cerevisiae)
sel.	selected
SILAC	stable isotope labeling by amino acids in cell culture
SSC	Side Scatter
st	stable
SV40	Simian virus 40
t	time-point
T_M	melting temperature
Thoc6	THO complex 6 homolog (Drosophila)
thr	threshold
tRNA	transfer RNA
TU	transduction unit
U	units
U–2 OS	human osteosarcoma cell line
Ub	ubiquitin
UniHI	Unified Human Interactome
unst	unstable
UTR	untranslated region
UV	ultraviolet
V	volt
w/v	weight per volume

Danksagung

Mit diesen Zeilen möchte ich denen danken, die mich wissenschaftlich, geistig sowie menschlich während der Zeit meiner Doktorarbeit unterstützt haben.

An erster Stelle möchte ich Prof. Dr. Achim Kramer danken, dass er mir die Möglichkeit gegeben hat, meine Doktorarbeit auf dem hochspannenden Gebiet der Molekularen Chronobiologie anzufertigen. Vielen Dank für die sehr gute wissenschaftliche Betreuung, vor allem für die stets konstruktiven Anregungen und Gespräche.

Prof. Dr. Markus Wahl möchte ich für die freundliche Übernahme der Begutachtung und dem Interesse an dieser Arbeit herzlich danken.

Ich danke Prof. Dr. Steven Elledge und Dr. Hsueh-Chi Sherry Yen für das Generieren und Bereitstellen der U–2 OS GPS–hORFeome V5.1 Bibliothek.

Mein besonderer Dank gilt Dr. Karsten Jürchott, der mir bei den Microarray–Analysen mit Rat und Tat zur Seite stand und mich in die Tiefen von ‚R' einführte.

Für den Einblick und die Diskussion über Ihre unveröffentlichten Daten möchte ich Dr. Charo Robles sowie Sarah Lück und Dr. Pål Westermark herzlich danken.

Der gesamten Arbeitsgruppe ‚Chronobiologie' möchte ich für die sehr angenehme Arbeitsatmosphäre, die Unterstützung im Labor und die vielen hilfreichen Diskussionen und Anregungen danken. Für die gesicherte Versorgung mit Materialien und der Unterstützung in bürokratischen Angelegenheiten danke ich Astrid Grudziecki und Annette Hayungs für Ihre große Hilfe im Laboralltag. Ich danke Rupert Öllinger, Sandra Korge und Sebastian Jäschke für Ihre uneingeschränkte Hilfe im Labor und Unterstützung in meinem Projekt. Sabrina Klemz und vor allem Roman Klemz möchte ich für die zahlreichen konstruktiven Gespräche und aufmunternden Worte herzlich danken. Mein ganz besonderer Dank, für sein Interesse und seine Unterstützung in meinem Projekt, vor allem aber für seine uneingeschränkte Loyalität, gilt Thomas Wallach, der zu einem wahren Freund geworden ist. Für die zahlreichen sportlichen, musikalischen und geselligen Aktivitäten außerhalb des Labors danke ich den gesamten ‚Chronos' – Ihr habt diese Zeit intensiv und unvergesslich gemacht.

Meinen Freunden, vor allem Maria Hennig und Marion Gottschald möchte ich für Ihr Verständnis, den gelungenen Ausgleich und Ermunterungen in dieser Zeit sehr danken.

Für die uneingeschränkte Unterstützung, Liebe und Kraft auf allen Lebenswegen richtet sich mein ganz besonderer Dank an meine Lieben, meinen Bruder und seiner Familie, vor allem an meine Eltern, die immer für mich da sind und es mir ermöglicht haben meine Ziele zu verfolgen. Schließlich möchte ich insbesondere Chris Schöndube von Herzen danken, dass er mir den Raum und die Zeit gegeben hat meine Ziele zu verwirklichen, vor allem aber für seine Liebe und Unterstützung in guten wie in schweren Zeiten.

Statement of Authorship

I declare on oath that I completed this work on my own and that information which has been directly or indirectly taken from other sources has been noted as such. Neither this, nor a similar work, has been published or presented to an examination committee.

Berlin, den 28.11.2013 Katja Schellenberg