Genome-wide circadian regulation: A unique system for computational biology

Lingying Sun a, Junjie Ma a, Christoph W. Turck b, Pin Xu c,*, Guang-Zhong Wang a,*

aCAS Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China
bMax Planck Institute of Psychiatry, Department of Translational Research in Psychiatry, Munich 80804, Germany
cDepartment of Neuroscience, University of Texas Southwestern Medical Center, Dallas, TX 75390-9111, USA

A B S T R A C T

Circadian rhythms are 24-hour oscillations affecting an organism at multiple levels from gene expression all the way to tissues and organs. They have been observed in organisms across the kingdom of life, spanning from cyanobacteria to humans. In mammals, the master circadian pacemaker is located in the hypothalamic suprachiasmatic nuclei (SCN) in the brain where it synchronizes the peripheral oscillators that exist in other tissues. This system regulates the circadian activity of a large part of the transcriptome and recent findings indicate that almost every cell in the body has this clock at the molecular level. In this review, we briefly summarize the different factors that can influence the circadian transcriptome, including light, temperature, and food intake. We then summarize recently identified general principles governing genome-scale circadian regulation, as well as future lines of research. Genome-scale circadian activity represents a fascinating study model for computational biology. For this purpose, systems biology methods are promising exploratory tools to decode the global regulatory principles of circadian regulation.

© 2020 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

1. Introduction .............................................................. 1915
2. The central circadian clock in mammals ......................... 1916
3. Input factors .......................................................... 1916
   3.1. Light ............................................................ 1916
   3.2. Temperature .................................................... 1917
   3.3. Food intake ...................................................... 1917
4. The circadian transcriptome of peripheral clock ................ 1917
5. Large-scale analyses of rhythmically expressed transcripts .. 1918
6. The design principles of the circadian network output ........ 1920
7. Circadian disruption and disease ................................... 1920
8. Summary and outlook ............................................... 1921
Author statement ...................................................... 1921

Abbreviations: ABSR, Autoregressive Bayesian spectral regression; AMPK, AMP-activated protein kinase; AR, Arrhythmic feeding; ARSER, Harmonic regression based on autoregressive spectral estimation; BMAL1, The aryl hydrocarbon receptor nuclear translocator-like (ARNTL); CCD, Cortical collecting duct; CR, Calorie-restricted diet; CRY, Cryptochrome; DCT/CNT, Distal convoluted tubule and connecting tubule; DD, Dark; dark; eJTK_CYCLE, Empirical JTK_CYCLE; HF, High fat diet; JTK_CYCLE, Jonckheere-Terpstra-Kendall (JTK) cycle; KD, Ketogenic diet; LB, Ad libitum; LD, Light:dark; Liver-RE, Liver clock reconstituted BMAL1-deficient mice; LS, Lomb-Scargle; NAD, Nicotinamide adenine dinucleotides; ND, Normal diet; NR, Night-restricted feeding; PAS, PER-ARNT-SIM; PER, Period; RAIN, Rhythmicity Analysis Incorporating Nonparametric methods; RF, Restricted feeding; SCN, Suprachiasmatic nucleus; SREBP, The sterol regulatory element binding protein; TITL, Transcriptional-translational feedback loop; WT, Wild type.

* Corresponding authors.
E-mail addresses: pin.xu@UTSouthwestern.edu (P. Xu), guangzhong.wang@picb.ac.cn (G.-Z. Wang).
1. Introduction

Rhythmic behavior is one of the earliest physiological phenomena that humans are aware of. The famous Chinese pre-Qin poem “the song of ji rang”, “beginning work at sunrise and resting at sunset ” describes the circadian activity of a farmer. The scientific observation of rhythmic behavior dates back to the 17th century, with the French scientist Jean-Jacques d’Ortous de Maïran’s observation of leaf movements[1]. de Maïran found that in constant darkness, the leaves of the mimosa plant retained the pattern of opening and closing that was observed with a light–dark cycle[1]. The molecular basis of circadian behaviors was not discovered until the 20th century. In 1971, Ron Konopka and Seymour Benzer published a milestone paper that reported abnormal circadian rhythms in the behavior of fruit fly mutants, with one mutant losing its circadian rhythm and another two mutants markedly changing their circadian period[2]. They discovered that both the eclosion rhythm and the locomotor activity of the mutants were different from those of the wild type flies, and all mutations were found at a single genetic locus[2]. These pioneering studies led to more research on identifying the genetic foundation of circadian rhythms. In 1984, Hall, Rosbash, and Young successfully cloned the DNA sequence of the period gene (per) [3–5]. Young’s group was able to restore the circadian rhythmicity of fruit fly mutants by introducing the per transcript sequence into the genome of an arrhythmic fly[4]. With this finding the first circadian rhythm gene had been discovered.

Substantial progress has been made in the study of circadian clocks in the last 20 years of the 20th century. The negative feedback loop of per was proposed as the basic mechanism underlying the circadian regulatory network [6–8]. In addition, another important gene, timeless (tim), was discovered in Drosophila [9–12]. In mammals, the first circadian gene Clock was located and sequenced by Takahashi’s group [13,14]. In Neurospora crassa, the frequency (frq) gene that is homologous to the per gene in Drosophila was sequenced [15]. Several plant circadian clock-related genes were also found, including that mutant strains in period were found in green algae Chlamydomonas reinhardtii [16], and the cab gene, CCA1 and LHY in Arabidopsis was reported [17–19]. The discovery of circadian genes in cyanobacteria [20–22] further added to our understanding of the circadian mechanisms, as this oscillation does not require transcription. This finding provided further evidence that the core circadian proteins KaiA, KaiB, and KaiC produce oscillation at the molecular level in vitro [23]. One note is the fact that the sequences of many core clock genes lack obvious similarity, supporting the multiple origin hypothesis of the clock[24]. The number of published articles related to biological rhythms has been steadily increasing over the past years (Fig. 1), as research continues to expand the boundaries of knowledge in this field.

Although the circadian regulation, which generates persistent rhythmic behavior under constant environmental conditions, can be detected in almost all living organisms, different species need to adapt to specific environmental factors for survival. Thus, the input and output of this regulation varies tremendously across...
different organisms. This review emphasizes the impact of the different factors on transcriptome-wide activities in mammalian species (for more details of other model organisms, please refer to [25,26]). In mammals, circadian rhythms are composed of three major components: input factors, the central oscillator, and outputs [27–29]. The input factors refer to the environmental factors that affect the circadian clock, such as light, food intake, and temperature. The central oscillator, also known as the central clock, generates the oscillation behavior [28–30]. Both central and peripheral clocks consist of an interplay of positive and negative regulators, which affect transcriptional-translational feedback loops. The output of this system refers to the circadian output signals, which oscillate approximately 24 h in most organisms. In mammals, the clock system has a hierarchical structure, and suprachiasmatic nuclei (SCN) influences the peripheral clocks through the regulation of the nervous system and the endocrine system [30]. Therefore, output pathways exist between SCN and peripheral tissues at the system’s level [30].

In the following, we will first elaborate upon the core molecular mechanisms that generate those rhythms in mammals (from several model organisms), followed by describing the external factors that affect biological rhythms. Finally, we will discuss the genome-wide circadian regulation and the general design principles of the output from the entire regulatory network. Biological rhythmic regulation is a good model for systems biology research because of the complexity of the input and output characteristics at the molecular level. Although the main factor affecting the rhythmicity of gene expression have been discovered recently [31], the systematic modeling of this regulatory system under different conditions is in its infancy. For instance, although many metabolic pathways are reported to be involved in rhythmically regulating gene expression in an experimental condition dependent manner, the common principles governing the regulation of these different pathways and individual condition remain unclear. For tissue specificity modeling, the ideal model should consider tissue specific regulatory mechanisms, and the key cycling pathways together with their pathological status. The present review illustrates the importance of the circadian system as a model for investigating molecular systems biology.

2. The central circadian clock in mammals

The core molecular mechanisms dictating biological rhythms have been the central topic in the circadian field. In mammals, the molecular mechanism of circadian oscillations is formed by an autoregulatory transcriptional-translational feedback loop (TTFL). In general, the positive elements of the feedback loop activate a set of negative regulators which in turn inhibit their own transcription.

In mammals, the positive elements of the circadian clock are CLOCK (and its paralog NPAS2) and BMAL1. These basic-helix-loop-helix -PER-ARNT-SIM (PAS) transcription factors form heterodimers and bind to regulatory element containing E-boxes in Per1, Per2, Per3 and Cry1, Cry2, which are the negative elements. PER and CRY proteins dimerize and interact with CLOCK:BMAL1 to repress their own transcription. The feedback loop also involves the nuclear orphan receptors such as RORα/β/γ and REV-ERBα/β, which repress the transcription of Bmal1 [32–34].

The absence or mutation of core clock genes can affect the molecular processes of the entire circadian regulatory network (Fig. 2). There are about 900 circadian-regulated genes in the wild-type mouse liver, and the rhythmicity of more than 90% of them was perturbed following the deletion of REV-ERBα/β [37]. The REV-ERBα/β knockout mice exhibit altered circadian wheel-running behavior and have lipid metabolism disorders [37]. In the islet beta cells 1,757 genes with altered expression levels were found in Bmal1 ablation mice, with 1,074 genes decreased and 683 genes increased compared to wild-type mice [40]. The islet beta cells from Bmal1 ablation mice no longer exhibit nutrient-responsive insulin secretion, and destruction of CLOCK and BMAL1 in the mouse pancreas causes hypoinsulinemic diabetes [40]. Rhythmic expression was observed in 5457 out of 37,681 transcripts (14.5%) in the liver of inducible Bmal1 knockout mice. In WT mice, 716 out of 6,818 and 267 out of 7,824 genes were rhythmically expressed in the liver and skeletal muscle, respectively, and the majority (71% in liver and 78% in muscle) of these rhythmic genes had significantly different expression levels in Clock mutant tissue [35]. Total RNA sequencing and ribosome profiling data from the liver of Bmal1 knockout mice revealed that loss of Bmal1 expression affected mRNA accumulation at both the transcriptional and post-transcriptional levels [41].

3. Input factors

Environmental stimuli can affect the circadian rhythm as zeitgebers, such as light, temperature, hypoxia, and methamphetamine, etc. We will focus on light, temperature, and food intake, which have diverse effects in different species, and even for different developmental periods. They influence a great part of the transcriptome and are the most intensively studied stimuli so far.

![Fig. 2. The transcriptome-wide effects of core clock gene mutations (Clock, Bmal1 and Rev-erb α/β) on the circadian clock of mouse liver. Data source: Clock mutant [35]; Clock (-/-) [36]; Rev-erb α/β(-/-) [37]; Bmal1(-/-) [38]; Bmal1(-/-) [39].](image-url)
3.1. Light

For most organisms, the light–dark cycle is the most apparent environmental factor that influences circadian behavior[42]. Completely blind individuals may have free-running circadian rhythms, although they can sense diurnal changes of environmental factors other than light, suggesting that light is one of the most important circadian stimuli[43]. In mammals, light signals are detected by retina, transmitted through retinohypothalamic tract to the SCN and change expression of clock genes (by cAMP response element-binding protein), causing daily phase changes. Light signals activate mitogen- and stress-activated protein kinase (MAPK) pathway and induce the expression of genes containing CAMP response element binding protein (CREB)[44]. In addition, the light signal regulates the phosphorylation of the translation initiation factor eIF4E by affecting MAPK/MNK pathway, and further regulates the translation of PER1 and PER2[45]. Early microarray analysis identified hundreds of cycling transcripts in the SCN under constant darkness, most of which are SCN-specific[46]. Specifically restoring Clock function in the brain of Clock mutants rescues the rhythmic expression of large numbers of cycling genes[47]. Interestingly, a much larger gene set (4,569 genes) was recently identified showing rhythmic expression in the SCN under light–dark condition, and an unexpected group of more than 700 genes was observed that peaked twice per day, indicating the complexity of their transcriptome[48]. In general, more genes show rhythmic expression under light–dark (LD) condition compared to dark–dark (DD) condition in peripheral tissues. For instance, 2960 and 2302 circadian transcripts were observed in epidermis in LD and DD conditions of wide type mouse, respectively[49], which indicates that different pathways are involved in regulating rhythmic gene expression under these two light conditions.

3.2. Temperature

Similar to light, temperature also changes rhythmically in the natural environment. Temperature has a fundamental effect on physiology, and variations in temperature are key to maintain the stability of circadian rhythms[50]. An important property of circadian clock is temperature compensation[51], i.e. the period of a free-running cycle remains relatively stable under different temperature conditions. The temperature compensation mechanism enables organisms to maintain the stability of the circadian clock under temperature fluctuations, thereby ensuring the consistency of physiological and behavioral rhythms. However, the mechanism of temperature compensation is not fully understood, and it may be related to the heat-shock pathway[52,53]. In mammals, temperature entrainment acts as a universal synchronizing factor, which resets circadian oscillations[54]. Subsequently, this signal acts as the input for the entrainment of other organs in the body.

3.3. Food intake

In addition to environmental cues, factors such as feeding, fasting, and social contact also affect the circadian clock. In the peripheral tissues, the circadian clock is an intrinsic time-keeping mechanism to cope with the changing external environment and maintain internal homeostasis. Food, including feeding and fasting, is an important synchronizer of the rhythmicity of gene expression, liver mass, and even cell size[36,55–58]. Restricted feeding (RF) can alter the circadian rhythms of fish, birds, and many mammals[59]. Time restricted feeding reduced body weight and cholesterol levels, and at the same time improved insulin sensitivity in rodents[60]. Under DD, about 15% of transcripts in the mouse liver are rhythmically expressed, which is driven by both food intake and the circadian regulatory network[56,61]. Time restricted feeding restored the oscillation of many transcripts in oscillator-deficient mouse livers[56]. Feeding- or fasting-induced transcripts accumulated as the amount of feeding or fasting time increased. After feeding, Per1 mRNA levels were downregulated, while Per2 expression levels were upregulated[56]. In ad lib-fed wild-type mice, 2997 transcripts had circadian expression patterns, while only 368 transcripts among them expressed circadian patterns under fasting conditions[56]. This indicates that temporal pattern of food intake in combination with circadian clock drive the expression of rhythmic genes in the liver[56]. The feeding activity of most animals has its own rhythmicity. This rhythm is important for maintaining the synchronization of metabolism and behavior. Under normal circumstances, multiple organs or tissues participate in the digestion, absorption, and metabolism of food after feeding. A suite of proteins are involved in this process, including NAD-dependent enzymes (e.g., sirtuins and poly[(ADP-ribose) polymerases], and protein kinases (e.g., AMP-activated protein kinase) [48], which in turn affect the expression of rhythmic genes. Reducing the rhythmic expression of BMAL1 and REV-ERBα in the liver and muscle led to an inhibition of their target genes[62]. In both liver and muscle, around 70 (about 30%) differentially expressed BMAL1-target genes had higher expression levels under fasting conditions, and these genes were induced in a BMAL1-dependent manner[62].

A time-restricted feeding pattern produces a sharp feeding-fasting cycle that consolidates rhythmic expression of genes and activation of various metabolic pathways[63] (Fig. 3). Compared with cell-autonomous hepatic clock, signals provoked by rhythmic food intake were more potent at driving circadian gene expression in mouse liver[64]. In this study, mice were fed with an automatic feeding system according to three patterns: arrhythmic (AR) feeding, night-restricted (NR) feeding, and ad libitum (LB) feeding. More than 1000 genes (1454 in total) of the cycling transcriptome in the liver of ad libitum fed mice lost rhythmicity in AR fed mice, but core clock genes remained unchanged. Collectively, these observations suggest that arrhythmic feeding does not affect the rhythmicity of core clock genes (Bmal1, Clock, Cry1, and Per1), and the rhythmic feeding has a strong effect on the rhythmicity of gene expression in the liver[64].

4. The circadian transcriptome of peripheral clock

The circadian regulatory network exists in almost all organs and cells throughout the body and the autonomous circadian rhythmicity can be maintained independent of SCN[29,47,68–72]. Peripheral clocks were first reported in liver, kidney, heart, pancreas and then in other major organs [68–70]. It is important to note that usually 10–15% of transcribed genes are circadian regulated in one tissue, many of which are specific to the phenotype of that tissue. Although circadian regulatory network exists in every organ, in liver, rhythmic food intake (RF) has a higher impact on the rhythmic gene expression, which indicates the independence of the cellular autonomous liver clock[64]. Furthermore, in mice with liver-specific deletion of Bmal1, the rhythmic expression of hepatic glucose-related genes is disturbed[73], which may be caused by Bmal1-related chromatin interactions[74]. There is a close connection between circadian rhythm and bone metabolism especially for the metabolic functions of osteoblasts, osteoclasts and chondrocytes. The regulation of circadian clock is important for bone and cartilage homeostasis[75]. Musculoskeletal shows decreased muscle force, high bone mass, arthritis and tendon calcification in mice deficient in various core clock genes[76]. For tendon, time-series microarray studies have shown that
745 genes (4.6% of the expressed genes) show rhythmic oscillations[77]. For peripheral clocks, the microenvironment of cell or tissue stiffness impacts on circadian control[78]. Yang et al., identified 594 cycling genes in the breast of mice, and found that the older stroma becomes stiffer, which leads to a reduction in the amplitude of circadian clocks[79]. The disruption of tissue clocks damages tissue homeostasis, resulting in an increased risk of diseases such as metabolic disorder, age-related diseases and cancer [78].

The mechanism of communication between circadian clock genes in the mammalian peripheral tissue and the liver was studied by Koronowski et al.[80] using a tissue-specific clock mouse model in which the liver clock was reconstituted in BMAL1-deficient animals (liver-RE). They found that although the liver oscillated in the absence of other clocks in vivo, the circadian expression was reduced to 10% of normal rhythmic transcripts, i.e., 218 oscillating transcripts[80]. Many rhythmic behaviors were reduced or lost when liver-RE mice were placed under constant dark condition. The same phenomenon also occurred in the epidermis of RE mice (epidermis-RE)[49], suggesting that the circadian clock in the liver operates independently to other clocks, yet remains dependent on the light–dark cycle, as light synchronizes circadian clocks in the liver without the presence of other BMAL1-dependent clocks. These results are consistent with previous reports that the circadian rhythmicity between the SCN and the liver can be decoupled by feeding[69], and interestingly, a phase shift of 10 h in 2 days can be achieved in the liver under restricted feeding conditions. The circadian rhythms of the liver, the SCN, and other peripheral tissues can be similarly aligned by feeding[69]. These findings are important for exploring the relationship between circadian rhythms in peripheral tissues. High-throughput transcriptomic and metabolomic analyses of mice with lung cancer revealed that a unique set of transcripts and metabolites were cycling exclusively in livers of tumor-bearing mice[81]. Further data shows that lung cancer had no effect on the core clock, but instead that the liver metabolism is reprogrammed by altered inflammatory responses through the STAT3-Socs3 pathway[81]. This process led to disruption of AKT, AMPK, and SREBP signaling, and resulted in changes in insulin, glucose, and lipid metabolism[81]. Collectively, these findings illustrate the complex regulatory connectivity between the liver and the lung, which have potential clinical importance during treatments of hepatic and lung diseases.

### 5. Large-scale analyses of rhythmically expressed transcripts

The regulatory targets of core circadian clock genes (Clock, Bmal1, Per1, Per2, Cry1 and Cry2) are many as each of these transcription factors can bind to tens of thousands of loci[32,61]. With the development of microarray and high-throughput sequencing technology, increasing numbers of rhythmically expressed genes have been detected, often across multiple biological processes.

The design of large-scale circadian experiments is not trivial, as experimental challenges exist in both the animal model generating process and circadian behavior design process. The precision and accuracy of the measurement and the degree of rhythmicity should be determined beforehand and are dependent on the specific experimental aims (for a more detailed discussion, refer to[82]). To ensure sufficient statistical power, the typical recommendations for the experimental design are: 1. Sampling at a 4-hour interval or even more frequently; 2. With at least 2 or 3 biological replicates per each time point or 2 cycles.

More sensitive algorithms are now applied to the identification of rhythmically expressed genes, many of which are high-quality methods to calculate related oscillation parameters, including period, phase, and amplitude[82]. These algorithms include COSOPT[83], JTK_CYCLE[84], ARSER[85], HAYSTACK[86], Lomb-Scargle[87], RAIN[88], CircWave[89], eJTK_CYCLE[90], ABSR[91], fisher’s G test, cosine or sinusoidal wave-fitting algorithm, Fourier analysis, and BIO_CYCLE[92]. In Table 1 we show the genome-wide changes in circadian expression factors using JTK_CYCLE takes several days to complete[84]. Conversely, analysis of the same dataset using COSOPT takes only several hours[84]. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

**Fig. 3.** The transcriptome-wide effects of differential feeding patterns of the circadian clock in mice. The small cubes represent the type of mouse food. Different colors of cubes are used to distinguish foods with different nutrients. Brown cubes: normal food; Yellow cubes: high fat diet; Pink cubes: ketogenic diet. Data source: AR, arrhythmic feeding[64]; CR, calorie-restricted diet[65]; RF, restricted feeding[56]; KD, ketogenic diet[66]; HF, high fat diet[67]; FAST[62]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Table 1
Genome-wide impact of circadian regulatory networks and the corresponding computational methods used in typical mouse studies.

| Species | Number of rhythmic transcripts/genes | Organ | Condition | Genotype | Computational methods | Ref. |
|---------|--------------------------------------|-------|-----------|----------|-----------------------|------|
| 1 Mus musculus | 2,960 | Epidermis | 12:12 LD | WT | JTK_CYCLE | [49] |
|           | 1,107 | 1,018 | 2,302 | 476 | 836 | Reconstituted (RE) Bmal1 KO WT RE Bmal1 KO |
| 2 Mus musculus | 1,061 | Liver | AR | WT | F24, Meta Cycle, RAIN, Harmonic Regression |
|           | 2,718 | 1,454 | 2,146 (15.3%) | 830 (4%) | WT | JTK_CYCLE | [62] |
| 3 Mus musculus | 3,153 | Liver | LB Fast | WT | JTK_CYCLE | [62] |
|           | 476 | 1,107 | 1,018 | 2,302 | 476 | Bmal1 KO |
| 4 Mus musculus | 3,144 | Epidermal stem cells | Adult | WT | JTK_CYCLE | [67] |
|           | 2,109 | 1,363 | 2,376 | 1,979 | 2,221 | JTK_CYCLE | [49] |
|           | 2,221 | 2,221 | 2,507 | 3,210 | 3,661 | JTK_CYCLE | [49] |
| 5 Mus musculus | 4,201 | Liver | Young & ND Old & ND | WT | JTK_CYCLE | [65] |
|           | 3,239 | 6,404 | 5,232 | 2,773 | 1,246 | JTK_CYCLE | [65] |
| 6 Mus musculus | 1,520 | Liver | Control chow | WT | JTK_CYCLE | [66] |
|           | 3,140 | 1,511 | 1,300 | 1,300 | 1,502 | JTK_CYCLE | [66] |
| 7 Mus musculus | 5,457 (14.5%) | Liver | 12:12 DD | WT | JTK_CYCLE | [39] |
|           | 1 | 2,869 (11%) | 2,730 (10%) | WT | JTK_CYCLE | [95] |
| 8 Mus musculus | 4,569 (24%) | Islet | 12:12 LD | WT | JTK_CYCLE | [40] |
|           | 3,905 (27%) | 2,460 | 1,330 | 1,502 | 1,220 | JTK_CYCLE | [40] |
| 9 Mus musculus | 1,197 | Liver | 12:12 LD | WT | JTK_CYCLE | [96] |
|           | 1,197 | 1,285 | 1,794 | 2,217 | 1,067 | JTK_CYCLE | [96] |
| 10 Mus musculus | 1,067 | Lung | 12:12 LD | WT | JTK_CYCLE | [96] |
|           | 1,067 | 1,067 | 1,067 | 1,067 | 1,067 | JTK_CYCLE | [96] |
| 11 Mus musculus | 1,067 | Lung | 12:12 LD | WT | JTK_CYCLE | [96] |
|           | 1,067 | 1,067 | 1,067 | 1,067 | 1,067 | JTK_CYCLE | [96] |
| 12 Mus musculus | 1,067 | Lung | 12:12 LD | WT | JTK_CYCLE | [96] |
|           | 1,067 | 1,067 | 1,067 | 1,067 | 1,067 | JTK_CYCLE | [96] |
| 13 Mus musculus | 1,067 | Lung | 12:12 LD | WT | JTK_CYCLE | [96] |
|           | 1,067 | 1,067 | 1,067 | 1,067 | 1,067 | JTK_CYCLE | [96] |
| 14 Mus musculus | 1,067 | Lung | 12:12 LD | WT | JTK_CYCLE | [96] |
|           | 1,067 | 1,067 | 1,067 | 1,067 | 1,067 | JTK_CYCLE | [96] |
| 15 Mus musculus | 1,067 | Lung | 12:12 LD | WT | JTK_CYCLE | [96] |
|           | 1,067 | 1,067 | 1,067 | 1,067 | 1,067 | JTK_CYCLE | [96] |
| 16 Mus musculus | 1,067 | Lung | 12:12 LD | WT | JTK_CYCLE | [96] |
|           | 1,067 | 1,067 | 1,067 | 1,067 | 1,067 | JTK_CYCLE | [96] |
| 17 Mus musculus | 1,067 | Lung | 12:12 LD | WT | JTK_CYCLE | [96] |

(continued on next page)
dealing with low-resolution data, ARSER will have a lower false negative rate, while JTK_CYCLE will report phase less accurately, together with a lower number of cycling genes. However, when high-resolution time data are processed, JTK_CYCLE performs better than ARSER, with lower false positive rates[119].

Many transcriptomic datasets have been published with the development of high-throughput sequencing technology, and additional cycling genes have been identified in each tissue using these datasets[46,102,104]. An early study using microarray identified 650 cycling transcripts in the liver and SCN with only a small number overlapping with each other[46]. Mouse liver is one of the most used tissues for circadian research and 1371 intron and 2037 exon cycling transcripts were reported recently[61]. Oster H et al.[89] performed whole-genome microarray hybridization to identify circadian genes in the mouse adrenal gland, and found that about 5% of the genes are under the circadian control. To study the role of circadian rhythms on the physiology and behavior of mice, Zhang R et al.[96] combined RNA-seq and DNA array technology to obtain quantitative time-dependent transcription data of 12 mouse tissues, including both mRNAs and non-coding RNAs. They found that 43% of protein-coding genes display circadian oscillations in at least one mouse tissue, and that the expression of many rhythmic genes peaked before dawn and dusk. A well-controlled analysis using human skeletal muscle revealed that 5748 pre-mRNA/mRNA transcripts are rhythmically expressed[106]. Most of the rhythmic genes identified in multiple organs were found in mammals with nocturnal habits, and there are no time series expression datasets from different tissues or brain regions of humans or species that are closely related to humans. Based on this, Mure et al.[114] performed genome sequencing on 64 tissue samples from baboons, including 22 brain regions. They found that most protein-coding genes in baboons have rhythmic expression[114].

**6. The design principles of the circadian network output**

The increase in the number of known rhythmic genes in different species has contributed to our overall understanding of the characteristics of rhythmic genes, which supports deeper understanding of the design principles of a circadian clock system. So far, four important design principles have been revealed: (a) Several studies have indicated that the distribution of cycling genes is tissue-specific[35,46,114], with only a small overlap between tissues. The recently released baboon circadian transcriptome shows that less than 1% (<10) of rhythmic genes were shared among tissues[114]. Comparison of cycling genes in the mouse SCN and liver revealed that a total of 28 genes were shared between these two tissues[46]. (b) There is a strong positive correlation between expression level and the amplitude of rhythmically expressed genes across different tissues in mice (E-A correlation). Rhythmic gene expression can explain most of the variation observed in the amplitude of circadian genes, suggesting that it is a major factor in the regulation of rhythmic gene behavior[31]. (c) The energetic cost of expressing a cycling gene is significantly higher than that of a non-cycling gene in mice, fruit flies, and yeast[120], as the former has significantly higher expression levels than the latter. (d) In simulation experiments, switching between the sequence of cycling and noncycling genes led to an increased overall cost of expression for the whole transcriptome[120]. Moreover, the overall energy cost of expressing the whole transcriptome is increased if all the genes are shuffled in the genome. In summary, the higher the expression level of the cycling genes, the more energy is consumed during the transcription process. As a consequence transcriptional systems tend to downregulate these highly expressed genes when their function is not needed, and rhythmically regulating highly expressed genes can effectively reduce the energy cost of transcription[31,120]. This strategy is part of the overall design requirement for reducing the energetic cost of expressing the whole transcriptome. In addition, the circadian regulatory system tends to regulate paralogous copies that consume more energy through evolution[120].

**7. Circadian disruption and disease**

Disrupted circadian rhythms can result in many diseases and affect downstream gene expression. The factors that cause circadian rhythm disorders can be roughly divided into three types: gene-level factors, such as mutations in clock genes, physiological factors, such as obesity and aging, and environmental factors, such as irregular work schedules, jet lag, and shift works.

Mutations or deletions of core clock genes directly impact the circadian system, leading to circadian disorders. Smal1 gene knockout or conditional knockout mice lose behavioral rhythms, such as
The effects of four different sleep-wake cycles on the circadian transcriptome of human blood. Shift work or mistimed sleep or sleep loss can lead to transcriptome-wide circadian changes. The clock (left) represents the sleep or wake condition and right boxes represents changes in each condition. Data source: Night shift [105]; Sleep deprivation [109]; Mistimed sleep [138]; Normal sleep (melatonin) [138].

Fig. 4. The effects of four different sleep-wake cycles on the circadian transcriptome of human blood. Shift work or mistimed sleep or sleep loss can lead to transcriptome-wide circadian changes. The clock (left) represents the sleep or wake condition and right boxes represents changes in each condition. Data source: Night shift [105]; Sleep deprivation [109]; Mistimed sleep [138]; Normal sleep (melatonin) [138].
improving drug efficacy or reducing side effects [139–141]. A considerable number of rhythm genes are targets for drug action [96]. Therefore, studies of circadian regulation are highly significant for drug development and disease treatment.

In summary, the circadian clock affects the expression of many different genes that are involved in various processes. Rhythm genes are usually expressed at high levels, and are often involved in the metabolism of glucose, carbon, lipids, etc. The transcription system rhythmically regulates the highly expressed genes according to the needs of each biological process [120]. If these high-expression genes are downregulated when they are not in need, the overall energy cost of transcription and translation can be reduced. This strategy allows the energy distribution within the genetic system to be optimized, and homeostasis to be maintained. This review of biological rhythms promotes the understanding of the molecular mechanisms of circadian regulation at the system's level and highlights the importance of studying the circadian clock in systems biology as a model system.

Author statement

GZW conceptualized the manuscript. LS drafted the manuscript. JM curated the data. GZW and PX supervised the manuscript; LS, GZW, PX and CWT reviewed and edited the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 81827901, 31600960 and No. 31871333) and the National Key R&D Program of China (2016YFC0901700 and 2016YFC1303100). We appreciate Luoying Zhang and Hung-Chun Chang for helpful feedback on the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2020.07.002.

References

[1] Roenneberg T, Merrow M. Circadian clocks – the fall and rise of physiology. Nat Rev Mol Cell Biol 2005;6(12):965–71.
[2] Konopka R, Benzer S. Clock mutants of Drosophila melanogaster. Proc Natl Acad Sci U S A 1971;68(9):2112–6.
[3] Bargiello TA, Young MW. Molecular genetics of a biological clock in Drosophila. Proc Natl Acad Sci U S A 1984;81(7):2142–6.
[4] Bargiello TA, Jackson FR, Young MW. Restoration of circadian behavioural rhythms by gene transfer in Drosophila. Nature 1984;312(5996):752–4.
[5] REDDY P et al. Molecular analysis of the period locus in Drosophila melanogaster and identification of a transcript involved in biological rhythms. Cell 1984;38(3):701–10.
[6] Reddy P et al. Molecular analysis of the period clock gene in Drosophila. Cell 1984;38(3):701–10.
[7] HARDIN PE, HALL JC, ROSBASH M. Feedback of the Drosophila period gene activity on its own expression. Cell 1998;93(7):1207–17.
[8] HARDIN PE, HALL JC, ROSBASH M. The implications of multiple circadian clock origins. PLoS Biol 2005;3(7):e262.
[9] LI S, ZHANG L. Circadian Control of Global Transcription. Biomed Res Int 2015;2015:87809.
[10] DORSETY J, KAY SA. Circadian control of global gene expression patterns. Annu Rev Genet 2010;44:419–44.
[11] TAKASHI JS, ZATZ M. Regulation of circadian rhythm. Science 1982;217(4565):1104–11.
[12] COLEWILL CS. Linking neural activity and molecular oscillations in the SCN. Nat Rev Neurosci 2011;12(10):533–69.
[13] MOHAWK JA, GREEN CB, TAKASHI JS. Central and peripheral circadian clocks in mammals. Annu Rev Neurosci 2012;35:445–62.
[14] DIBNER C, SCHILBER U, ALBRECHT U. The mammalian circadian timing system: organization and coordination of central and peripheral clocks. Annu Rev Physiol 2010;72:517–49.
[15] CHENG Y, CHI Y, ZHANG L, WANG G-Z. A single factor dominates the behavior of circadian clock genes in mouse organs. BMC Genomics 2015;20(1):879.
[16] TAKASHI JS. Transcriptional architecture of the mammalian circadian clock. Nat Rev Genet 2017;18(3):164–79.
[17] SHEARMAN LP et al. Interacting molecular loops in the mammalian circadian clock. Science 2000;288(5481):1013–9.
[18] PARICH CL, GREEN CB, TAKASHI JS. Molecular architecture of the mammalian circadian clock. Trends Cell Biol 2014;24(2):90–9.
[19] MILLER BH et al. Circadian and CLOCK-controlled regulation of the mouse transcriptome and cell proliferation. Proc Natl Acad Sci U S A 2007;104(9):3342–7.
[20] ECKEL-MAHAN KL et al. Coordination of the transcriptome and metabolome by the circadian clock. Proc Natl Acad Sci U S A 2012;109(14):5541–6.
[21] CHOI H et al. Regulation of circadian behaviour and metabolism by REV-ERB-alpha and REV-ERB-beta. Nature 2012;485(7396):123–7.
[22] KONDRAVOT RV, KONDRAVOT AA, GORBAicheva YA, VYKHovANETS OV, ANTOCH MP. Early aging and age-related pathologies in mice deficient in BMAL1, the core component of the circadian clock. Genes Dev 2006;20(14):1868–73.
[23] YANG G et al. Timing of expression of the core clock gene Bmal1 influences its effects on aging and survival. Sci Transl Med 2016;8(324):324ra316.
[24] PERELIS M et al. Pancreatic beta cell enhancers regulate rhythmic transcription of genes controlling insulin secretion. Science 2015;350(6261):aac4250.
[25] ATGER F et al. Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. Proc Natl Acad Sci U S A 2015;112(15):e5679–e5688.
[26] LeGATES TA, Fernandez DC, Hattar S. Light as a central modulator of circadian rhythms. Annu Rev Physiol 2010;72:517–49.
[27] Kondo T et al. Circadian rhythms by gene transfer in Drosophila. Nature 1984;312(5996):752–4.
[28] Nakajima M et al. Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science 2005;308(5720):414–5.
[29] ROSBASH M. The implications of multiple circadian clock origins. PLoS Biol 2005;3(7):e262.
[30] Wang ZY, Tobin EM. Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. Cell 1998;93(7):1207–17.
[31] Kondo T et al. Circadian clock rhythms in prokaryotes: luciferase as a reporter of circadian gene expression in cyanobacteria. Proc Natl Acad Sci U S A 1993;90(12):5672–6.
[32] Kondo T et al. Circadian clock mutants of cyanobacteria. Science 1994;266(5188):1233.
[33] Ishizu M et al. Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria. Science 1998;281(5382):1519–23.
[34] Nakajima M et al. Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science 2005;308(5720):414–5.
[35] Rosbash M. The implications of multiple circadian clock origins. PLoS Biol 2005;3(7):e262.
Kornmann B, Schaad O, Bujard H, Takahashi JS, Schibler U. System-driven and Yeung CY et al. Gremlin-2 is a BMP antagonist that is regulated by the Dudek M, Meng QJ. Running on time: the role of circadian clocks in the Gonçalves CF, Meng QJ. Timing metabolism in cartilage and bone: links Lamia KA, Storch KF, Weitz CJ. Physiological significance of a peripheral tissue Rensing L, Monnerjahn C. Heat Shock Proteins and Circadian Rhythms. Hughes ME et al. Guidelines for genome-scale analysis of biological rhythms. J Biol Rhythms 2017;32(5):380–93.

[47] Hughes ME et al. Brain-specific rescue of Clock reveals system-driven transcriptional rhythms in peripheral tissue. PLoS Genet 2012;8(7):e1002835.

[48] Pembroke RG, Babbs A, Davies KE, Oliver PL. Temporal transcriptions suggest that twin-peaking genes reset the clock. Elife 2015;4/5.

[49] PS, et al. BMAL1-Driven Tissue Clocks Respond Independently to Light to Maintain Homeostasis. Cell 2019;177(6):1436–47 e1412.

[50] Boothroyd CE, Wijnken H, Noef F, Saez L, Young MW. Integration of light and temperature in the regulation of circadian gene expression in Drosophila. PLoS Genet 2007;3(9):e165.

[51] Zimmerman WF, Pintendrigh CS, Pavlidis T. Temperature compensation of the circadian oscillation in Drosophila pseudobscura and its entrainment by temperature cycles. J Insect Physiol 1968;14(5):669–84.

[52] Young WW, Young MK. Temperature compensation and temperature sensation in the circadian clock. Proc Natl Acad Sci U S A 2015;112(46):E6284–6292.

[53] Rensing L, Menonjerh C. Heat Shock Proteins and Circadian Rhythms. Chronobiol Int 1996;13(4):239–50.

[54] Buhr ED, Yoo SH, Takahashi JS. Temperature as a universal resetting cue for mammalian circadian oscillators. Science 2010;330(6002):379–85.

[55] Sintuf F, et al. Diurnal Oscillations in Lung Mass and Cell Size Accompany Ribosome Assembly Cycles. Cell 2017;168(4):651–63 e614.

[56] Volmers C et al. Time of feeding and the intrinsic circadian clock drive temperature in the regulation of circadian gene expression in Drosophila. J Biol Rhythms 2017;32(10):547–56.

[57] Liu Y, Hong CL, Lim S, Song S. Finding clock genes in genes: a Bayesian approach to estimate periodicity. Biomed Res Int 2016;2016:3071475.

[58] Asher G, Sassone-Corsi P. Time for food: the intimate interplay between nutrition, metabolism, and the circadian clock. Mol Cell 2015;61(1):84–92.

[59] Koeke N et al. Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. Science 2012;338(6105):349–54.

[60] Aeschbach D, Buxton OB, Guyenet PG. Clocks and Circadian Rhythms in Metabolic Tissue. Cell Rep 2018;25(12):3299–314 e3296.

[61] Kinochi K et al. Fasting Imparts a Switch to Alternative Daily Pathways in Liver and Muscle. Cell Rep 2018;22(5):1225–35.

[62] Zheng L, Lahner NF, Bence HI, Hughes ME, Hogenes JB. A circadian gene expression atlas in mammals: implications for biology and medicine. Proc Natl Acad Sci U S A 2014;111(45):18219–24.

[63] Rensing L, Monnerjahn C. Heat Shock Proteins and Circadian Rhythms. Chronobiol Int 1996;13(4):239–50.

[64] Mange F et al. Diurnal regulation of RNA polymerase III transcription is under temperature control. Genes Dev 2019;33(23):2731–36.

[65] spaghetti, P. What time is it? J Biol Rhythms 2017;32(10):547–56.

[66] Schick S et al. Identifying Novel Transcriptional Regulators with Circadian Functions Controlling Insulin Sensitivity and Glucose Metabolism. Elife 2016;5.

[67] Haspel JA et al. Circadian rhythm reprogramming during lung inflammation. Nat Commun 2014;5:4753.

[68] Tanaka K et al. Mouse circadian clock genes are controlled by the intrinsic muscle clock. Mol Metab 2014;3(1):29–41.

[69] Feng B et al. Circadian enhancers coordinate multiple phases of rhythmic gene transcription in vivo. Cell 2014;159(3):1140–52.

[70] Le Martelot G et al. Genome-wide RNA polymerase II profiles and RNA accumulation reveal kinetics of transcription and associated epigenetic changes during circadian cycles. PLoS Biol 2012;10(1):e1001442.

[71] Neri L et al. Pancreatic β-cell enhancers regulate rhythmic transcription of genes controlling insulin secretion. Science 2015;350(6261):aac4250.

[72] Meten JS, Rodriguez J, Abruzzi KC, Rosbash M. Nascent-Seq reveals novel genes associated with mouse circadian transcriptional regulation. Elife 2012;1:e00011.

[73] Weitz CJ. Temporal transcriptions suggest that twin-peaking genes reset the clock. Elife 2015;4/5.

[74] Sintuf F, et al. Diurnal Oscillations in Lung Mass and Cell Size Accompany Ribosome Assembly Cycles. Cell 2017;168(4):651–63 e614.

[75] Kinochi K et al. Fasting Impacts a Switch to Alternative Daily Pathways in Liver and Muscle. Cell Rep 2018;25(12):3299–314 e3296.

[76] Aeschbach D, Buxton OB, Guyenet PG. Clocks and Circadian Rhythms in Metabolic Tissue. Cell Rep 2018;22(5):1225–35.

[77] Zheng L, Lahner NF, Bence HI, Hughes ME, Hogenes JB. A circadian gene expression atlas in mammals: implications for biology and medicine. Proc Natl Acad Sci U S A 2014;111(45):18219–24.

[78] Rensing L, Monnerjahn C. Heat Shock Proteins and Circadian Rhythms. Chronobiol Int 1996;13(4):239–50.
