Analysis of circadian pattern reveals tissue-specific alternative transcription in leptin signaling pathway
Andrey A Ptitsyn*1 and Jeffrey M Gimble2

Address: 1Experimental Obesity Laboratory, Louisiana State University Pennington Biomedical Research Center, 6400 Perkins Rd., Baton Rouge, LA 70808, USA and 2Stem Cell Laboratory, Louisiana State University System, Pennington Biomedical Research Center, 6400 Perkins Rd., Baton Rouge, LA 70808, USA

Email: Andrey A Ptitsyn* - Andrey.Ptitsyn@colostate.edu; Jeffrey M Gimble - GimbleJM@pbrc.edu

* Corresponding author

Abstract
Background: It has been previously reported that most mammalian genes display a circadian oscillation in their baseline expression. Consequently, the phase and amplitude of each component of a signal transduction cascade has downstream consequences.

Results: Here, we report our analysis of alternative transcripts in the leptin signaling pathway which is responsible for the systemic regulation of macronutrient storage and energy balance. We focused on the circadian expression pattern of a critical component of the leptin signaling system, suppressor of cytokine signaling 3 (SOCS3). On an Affymetrix GeneChip 430A2 microarray, this gene is represented by three probe sets targeting different regions within the 3' end of the last exon. We demonstrate that in murine brown adipose tissue two downstream 3' probe sets experience circadian baseline oscillation in counter-phase to the upstream probe set. Such differences in expression patterns are a telltale sign of alternative splicing within the last exon of SOCS3. In contrast, all three probe sets oscillated in a common phase in murine liver and white adipose tissue. This suggests that the regulation of SOCS3 expression in brown fat is tissue specific. Another component of the signaling pathway, Janus kinase (JAK), is directly regulated by SOCS and has alternative transcript probe sets oscillating in counter-phase in a white adipose tissue specific manner.

Conclusion: We hypothesize that differential oscillation of alternative transcripts may provide a mechanism to maintain steady levels of expression in spite of circadian baseline variation.

Background
Circadian or approximately daily rhythms occur throughout nature. Entrained by the daily light/dark cycle, feeding behavior, sleeping pattern and/or other environmental and physiological cues, gene expression patterns also demonstrate daily oscillations. Until recently, it has been widely accepted that up to 10–15 per cent of all genes are expressed following a circadian rhythm [1], generated by...
a negative feedback molecular oscillator, aka basic circadian clock. In mammals, the central circadian clock is synchronized to the daily light/dark cycle and is located in the Suprachiasmatic Nucleus (SCN) of the hypothalamus. However, it has been demonstrated that circadian clocks are active in peripheral tissues [2-5]. The peripheral clocks are synchronized through sympathetic outputs and the controlled secretion of circulating glucocorticoids, melatonin, and other mediators. Our recent publications suggest that the scale and role of circadian oscillation in gene expression has been underestimated [6,7]. We have presented evidence of baseline gene expression oscillations that exceed the accepted 10–15% and involve nearly all expressed genes. One of the most important implications of this discovery is related to the way we model, visualize, and understand biological pathways. Since virtually every component of each biological pathway is oscillating, the timing of regulation, interaction or signal transduction is critically important as well as the phase of oscillation at which a particular event takes place.

**Results**

To illustrate the concept of oscillation within a biological pathway we have chosen the leptin pathway as an example. Leptin is an important regulator of the energy balance in the organism and contributes to the sensation of satiety after a meal. Leptin-deficient mice are prone to overeating, excessive weight gain, and the development of obesity. A similar phenotype is observed in the case of leptin-deficient humans. Leptin is highly expressed in adipose tissue and secreted into the bloodstream. The leptin receptor (LEPR) is actively expressed in the liver. Upon ligand-receptor activation, a signal is transmitted through the Janus Kinase (JAK) and STAT3 cascade, which are shared in common with related signal transduction pathways. SOCS3, whose transcription is activated by STAT3, represses both JAK and LEPR and serves to attenuate the signal.

This leptin signal transduction pathway is well studied and can be found in a number of private and public databases. We have extracted the description along with the cartoon illustrating the leptin pathway from the Kyto Encyclopedia of Genes and Genomes (KEGG)[8]. We have scaled, smoothed, and assigned to most likely phase class the data sets with microarray expression profiles from the previously published studies (see Methods). By overlapping the gene expression profile for a particular tissue with the general leptin signaling cartoon, we produced a diagram illustrating the relative abundance of transcripts for each component of the pathway at a given time of the day (Figure 1).

As shown on Figure 1 in brown adipose tissue both leptin and leptin receptor are expressed in a circadian pattern with two peaks over the two-day observation period. The acrophase of leptin receptor is slightly ahead of the leptin acrophase, indicating increase in receptor production in anticipation of the signal. All other components of the leptin signaling pathway shown on Figure 1 have expression levels synchronized with the expression of the leptin receptor, i.e. they oscillate in the same phase.

However, SOCS3 is the exception to this pattern. In Figure 1, SOCS3 expression is present in two separate and contradictory plots oscillating in opposite phases; the acrophase (time of the highest point) of one profile corresponds to the bathyphase (time of the lowest point) of the other at an equivalent time point. In Figure 1, we have plotted 2 of the 3 probesets representing SOCS3 on the Affymetrix Mouse Expression Genechip. The observation that alternative probesets interrogating expression levels of the same gene could demonstrate such a contrast was novel and called for further investigation.

We extracted the target sequences for the individual probe sets represented on the GeneChip from the supporting materials provided by Affymetrix through the Developer Network. Then, we used these target sequences as queries in a BLAST search against the non-redundant set of nucleotide sequences and mouse ESTs. The target sequences were also aligned to the genomic sequence corresponding to the murine SOCS3 gene. The results of this mapping are shown in Figure 2.

All three Affymetrix probesets from SOCS3 were derived from the alternative consensi that result from alignment of murine ESTs to the genomic sequence (Additional File 1). All three target sequences from which the probesets are derived represent overlapping stretches of the 3' UTR of the SOCS3 gene. Comparing the relative positions of the target sequences and alignment of the ESTs from which they were derived, we came to the conclusion that the Affymetrix probesets represented alternative transcripts of different length, consistent with alternative polyadenylation. The presence of two potential polyadenylation signals within the 3' UTR of SOCS3 gene targeted by the probeset supports this hypothesis. The 3' polyadenylation signal has a canonical AATAAA sequence starting at position 11675. 5' to the canonical signal at position 11492 there is an alternative sequence AATAAT, also described in the literature as a potential PolyA signal [9]. The primary Affymetrix probeset 1416576_at targets the most 5' part of the 3' UTR, common for both alternative transcripts. The most downstream probeset 1455899_x_at targets the extreme 3' end specific to the longer version of SOCS3 transcript. The third probe set representing 1456212_x_at targets the area up- and downstream of the weaker alternative polyA signal. The target sequence for this probeset
is contained within the longer transcript completely, but only partially present in the shorter alternative version.

If our conclusion about Affymetrix probesets representing alternatively polyadenylated transcripts of different length is true, the middle of three probesets containing probes for both transcripts should demonstrate a mixed pattern. Indeed, as presented on Figure 2, expression profile of the intermediate probeset 1456212_x_at looks more erratic than the other two profiles. This profile also shows a remarkable similarity to the arithmetic sum of the expression profiles for the flanking 1416576_at and 1455899_x_at probesets.

The described phenomenon of alternative transcripts with contrasting oscillation patterns is not unique to SOCS3. We have analyzed the oscillation pattern in the components of the same pathway in different tissues. In liver we found no evidence of probesets with different phase of oscillation. Some of the probesets corresponding to the same transcript demonstrate different degree of noise, but allow no conclusive statement regarding the phase of oscillation. In white adipose tissue SOCS3 probesets also show no evidence of alternative oscillating pattern. However, the other component of the same pathway, JAK, has alternative probesets oscillating in counter-phase, just like SOCS3 in the brown adipose tissue.

Discussion

The notion that all genes being expressed in an oscillating manner has a few important implications. First, a steady line gene expression, i.e. a constant level of transcript abundance over time, is an abstraction for which we see no evidence in experimental data. If some transcripts are to be maintained at a certain level of abundance, this level can be constant only in relation to some other transcripts, which also means that these genes have to oscillate in the same phase. Even so called "housekeeping" genes, abundant and essential for the survival of a cell, in fact do experience daily oscillation of the baseline transcription. Affymetrix chips can register this oscillation because relative intensity of expression for each probeset is normalized to the mean expression level of the control probes. Control probesets represent all possible phases and their number

**Figure 1**

Circadian oscillation in leptin signaling pathway in mouse brown fat. This is a fragment of the KEGG pathway with superimposed plots of circadian expression profiles for the essential components of the pathway. Every gene involved in leptin signaling demonstrates explicit circadian pattern in baseline oscillation. All except one oscillate in the same phase. Leptin itself is actively expressed in brown adipose tissue and its intensity oscillates with a phase shift, slightly behind leptin receptor (LEPR) and other components of leptin signaling. Remarkably, suppressor of cytokine signal (SOCS3) has two alternative transcripts oscillating in counter-phase.
is large enough to have all possible phases represented approximately equally. All oscillations are averaged out and disappear when control sets are pooled in silico to normalize the expression of the rest of the chip. This resulting steady line is artificial, but solid enough to make all other phases of oscillation detectable. The second important implication is that the negative-feedback molecular clock is unlikely to be the main generator of all oscillations. It is hard to see a reason why genes from the oxidative phosphorylation pathway should be regulated by a circadian rhythm. But the opposite is more likely: multiple publications point that oxidative phosphorylation pathway is a natural oscillator and time-separation of biochemical reactions is essential for the aerobic metabolism [10-13].

Indeed in murine tissue samples we find genes related to oxidative phosphorylation expressed in the most clear noise-free circadian rhythm. It is likely that in higher eukaryotes the respiratory cycle is synchronized with the metabolic cycle and thus expression patterns of all other genes are modulated by the oscillation of the main energy production cycle.

Microarrays are designed to provide a snapshot of relative abundance of thousands of transcripts at a time. The abundance of transcripts depends on many factors, of which one of the most important is the turnover rate of a particular mRNA. The ratio between the rate of transcription, maturation, localization, and the rate of degradation
of mRNA determines the number of active copies available for translation. The structure of 3' UTR determines localization and stability of the transcript. The great majority of genes are represented by a mixed population of transcripts resulting from the alternative processing during mRNA maturation, while alternative polyadenylation is arguably the most common cause for such variation. It has been previously reported that oscillating genes can have alternative transcripts of different length [14,15]. Different length of 3' UTR results in a different turnover rate of PER3 protein, one of the key elements of the circadian molecular clock in mammals. Changes in PER3 turnover may result in accumulation of available PER3 mRNA and increased rates of protein synthesis, which leads to alteration of sleeping patterns, feeding behavior, and other periodic activities. Since alternative transcripts of SOCS3 are slightly different lengths and 3' UTR structure are very common it is logical to suggest that alternative transcripts may have different observed oscillating properties, resulting from the different turnover rates of alternative transcripts.

The observed pattern of a longer transcript oscillating in an opposite phase to the shorter version may have a special function. We hypothesize that this phenomenon may represent the cellular mechanism generating a "steady line" non-oscillating production of an important protein. Both SOCS3 and JAK are essential elements of a number of signal transduction pathways that may not be synchronized. Signal transduction events may take place at different times and a steady number of mature JAK and SOCS3 proteins have to be maintained in anticipation of such signals. A recent publication suggests that high abundance of SOCS3 transcripts is responsible for leptin resistance in mouse brown adipose tissue, but not in white adipose tissue or liver. The other component of the same signal transduction cascade, JAK has alternative transcripts oscillating in counter-phase in white adipose tissue, but not in liver or brown adipose tissue. We propose a hypothesis that alternative transcripts in counter-phase represent the evolutionary developed compensation for oscillating transcription rate. This function provides a steady non-oscillating abundance of essential proteins in eukaryotic cell.

Methods

Murine circadian gene expression data

We have completed independent circadian studies in AKR/J mice acclimated to a 12 hr light: 12 hr dark cycle, harvesting sets of 3–5 mice at 4 hr intervals in duplicates over a 24 hr period [4]. Total RNA samples from inguinal (IWAT) white adipose tissue, brown adipose tissue (BAT), and liver have been assayed by Affymetrix microarrays. A few genes have been selected for validation with RT-PCR for the expression profile of representative circadian rhythm genes in all 3 tissues. The transcriptomic data set contained over 22,000 gene expression profiles for each of 3 different tissues. In the current study, we have used only the murine liver data. Since each time point was sampled twice, the following Fourier transform for each profile can be re-arranged into a short time series that represents two complete circadian cycles. Smoothing and median subtraction procedure has been applied to all data sets.

Data pre-processing

Profiles have been smoothened by a 3rd degree polynomial procedure and median-subtracted. For smoothing we use seven-point Savitzky-Golay algorithm [17]. To take advantage of all points in the time series a single-pass smoothing has been applied in a circular manner, with the last points contributing to smoothing the starting points. For better compatibility, the same smoothing and median subtraction procedure has been applied to all data sets.
Phase classification
We have assigned phase to each expression time series by computing cross-correlation
\[
R(f) = \frac{\sum_{i=0}^{N-1} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=0}^{N-1} (x_i - \bar{x})(y_i - \bar{y})}
\]
where \(x\) is a gene expression time series of \(N\) points and \(y\) is an artificially generated profile of ideal cosine function
\[
y_i = \cos\left(\frac{2\pi}{p} * i\right), \text{ where } p \text{ is the number of time points in a complete circadian cycle, for example 6 in Zvoníc et al. [4] data set.}
\]
To account for all phases the artificial cosine curve profile has been regenerated with a phase shift by 1 time point. The highest correlation among all possible phase shifts was assigned as the most probable phase. The significance of periodicity was not assessed at this point, it was done separately by three independent procedures described above. All expression profiles were sorted (classified) first by the assigned phase then by ascending p-value estimated by one of the described algorithms.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
AAP has developed the algorithms, the software and performed data analysis and algorithm testing and benchmarking. AAP and JMG wrote the paper.

Additional material

Additional file 1
Graphical overview of the BLAST alignment of mouse genome fragment containing SOCS3 sequence and homologous mouse ESTs.
Click here for file [http://www.biomedcentral.com/content/supplementary/1471-2105-8-S7-S15-S1.doc]

Acknowledgements
We would like to acknowledge Steven Conrad from the Departments of Bioinformatics and Computational Biology, Medicine, and Emergency Medicine, Louisiana State University Shreveport Health Sciences Center for his help with the Affymetrix microarray processing.

This work was funded in part by the Pennington Biomedical Research Foundation (JMG).

This article has been published as part of BMC Bioinformatics Volume 8 Supplement 7, 2007: Proceedings of the Fourth Annual MCBIOS Conference. Computational Frontiers in Biomedicine. The full contents of the supplement are available online at [http://www.biomedcentral.com/1471-2105/8/S7/S15]

References
1. Hogenesch JB, Panda S, Kay S, Takahashi JS: Circadian transcriptional output in the SCN and liver of the mouse. Novartis Found Symp 2003, 253:171-180. discussion 152–177, 102–179, 180–173 passim
2. Panda S, Antoch MP, Miller BH, Su AI, Schook AB, Straume M, Schultz PG, Kay SA, Takahashi JS, Hogenesch JB: Coordinated transcription of key pathways in the mouse by the circadian clock. Cell 2002, 109(3):307-320.
3. Ueda HR, Chen W, Adachi A, Wakamatsu H, Hayashi S, Takasugi T, Nagano M, Nakahama K, Suzuki Y, Sugano S, et al.: A transcription factor response element for gene expression during circadian night. Nature 2002, 418(6877):534-539.
4. Zvoníc S, Ptitsyn AA, Conrad SA, Scott LK, Floyd ZE, Kilroy G, Wu X, Goh BC, Mynatt RL, Gibb JM: Characterization of peripheral circadian clocks in adipose tissues. Diabetes 2006, 55(4):962-970.
5. Storch KF, Lipan O, Leykin I, Viswanathan N, Davis FC, Wong WH, Weitz CJ: Extensive and divergent circadian gene expression in liver and heart. Nature 2002, 417(6884):78-83.
6. Ptitsyn AA, Zvoníc S, Conrad SA, Scott LK, Mynatt RL, Gibb JM: Circadian clocks are resounding in peripheral tissues. PLoS Comput Biol 2005, 2(3):e16.
7. Ptitsyn AA, Zvoníc S, Gibb JM: Permutation test for periodicity in short time series data. BMC Bioinformatics 2006, 7(Suppl 2):S10.
8. Kyoto Encyclopaedia of Genes and Genomes. .
9. Yang H, Melera PW: A genetic polymorphism within the third poly(A) signal of the DHFR gene alters the polyadenylation pattern of DHFR transcripts in CHL cells. Nucleic Acids Res 1994, 22(13):2694-2702.
10. Selkow E: On the Mechanism of Single-Frequency Self-Oscillations in Glycolysis. A Simple Kinetic Model. European Journal of Biochemistry 1968, 4(1):79-86.
11. Selkow E: Stabilization of Energy Charge, Generation of Oscillation and Multiple Steady States in Energy Metabolism as a Result of Purely Stoichiometric Regulation. European Journal of Biochemistry 1975, 59(1):151-157.
12. Tu BP, Kudlicki A, Rowicka M, McKnight SL: Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes. Science 2005, 310(5751):1152-1158.
13. Klevecz RR, Bolen J, Forrest G, Murray DB: A genomewide oscillation in transcription gates DNA replication and cell cycle. Proc Natl Acad Sci USA 2004, 101(5):1200-1205.
14. Archer SN, Robillard DL, Skene DJ, Smits M, Williams A, Arendt J, von Schantz M: A length polymorphism in the circadian clock gene Per3 is linked to delayed sleep phase syndrome and extreme diurnal preference. Sleep 2003, 26(4):413-415.
15. Viola AU, Archer SN, James LM, Groeger JA, Lo JC, Skene DJ, von Schantz M, Dijk DJ: PER3 Polymorphism Predicts Sleep Structure and Waking Performance. Curr Biol 2007.
16. Enriori PJ, Evans AE, Sinnayah P, Jobst EE, Tonelli-Lemos L, Bilas SK, Glavas MM, Grayson BE, Perello M, Nillni EA, et al.: Diet-Induced Obesity Causes Severe but Reversible Leptin Resistance in Arcuate Melanocortin Neurons. Cell Metab 2007, 5(3):181-194.
17. Savitzky A, Golay M: Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry 1964, 36:1627-1639.