*Systems biology*

# WAVECLOCK: wavelet analysis of circadian oscillation

Tom S. Price[1],*, Julie E. Baggs[2], Anne M. Curtis[3], Garret A. FitzGerald[2] and John B. Hogenesch[2]

[1]Institute of Psychiatry, Kings College London, [2]Institute for Translational Medicine and Therapeutics, University of Pennsylvania and [3]Bio Directorate, Science Foundation Ireland (SFI)

## ABSTRACT

**Summary:** Oscillations in mRNA and protein of circadian clock components can be continuously monitored *in vitro* using synchronized cell lines. These rhythms can be highly variable due to culture conditions and are non-stationary due to baseline trends, damping and drift in period length. We present a technique for characterizing the modal frequencies of oscillation using continuous wavelet decomposition to non-parametrically model changes in amplitude and period while removing baseline effects and noise.

**Availability:** The method has been implemented as the package *waveclock* for the free statistical software program *R* and is available for download from http://cran.r-project.org/

**Contact:** thomas.price@iop.kcl.ac.uk

**Supplementary information:** Supplementary figures are available at *Bioinformatics* online.

## 1 INTRODUCTION

Animals and plants adapt to the daily light:dark cycle through the action of endogenous biological clocks. Circadian rhythms in mammals are coordinated by a central pacemaker located in the suprachiasmatic nucleus (SCN) but autonomous circadian clocks are also present in peripheral tissues (Stratmann and Schibler, 2006). After serum shock, cellular models of clock function—such as transformed cell lines, mouse embryonic fibroblasts and dissociated SCN neurons—recapitulate oscillator function and allow continuous monitoring of the concentration of clock components through the use of reporter genes either as transcriptional (promoter-driving luciferase or GFP) or translational fusions (e.g. Per2::Luciferase) (Hastings, 2005; Yoo *et al.*, 2004). Although oscillations in cell cultures can persist for up to 20 days (Yoo *et al.*, 2004), the amplitude of the overall signal becomes damped due to desynchronization as the period of individual cells drifts over time (Nagoshi *et al.*, 2004; Rougemont and Naef, 2007). Another challenge for stable measurement is the initial spike following serum shock. Finally, these cellular models are increasingly being challenged with small molecules, as well as environmental and genetic perturbations, all of which can impact oscillator function. Many of these perturbagens have decay properties that change over the course of an experiment (e.g. pharmacodynamics of small molecules, $\rho$H). The analytic problem, therefore, is how to characterize a progressively damped signal of non-constant period in the presence of random noise and experimental challenge.

## 2 METHODS

Continuous wavelet transformation is a multiscale smoothing technique that projects the unidimensional time-series data into two-dimensional time-frequency space. In this transformed space, circadian oscillations typically appear as a ridge of slowly changing frequency with decreasing amplitude over time. We used nonparametric methods to identify these ridges and perform a reverse transformation to reconstruct the modal frequencies in the original signal.

We chose to use the complex-valued Morlet wavelet which preserves phase information and mainly smoothes over the time dimension allowing finer resolution in the frequency domain (Torrence and Compo, 1998). The correspondingly coarse resolution in the time dimension matters little in terms of feature detection since we expect slow changes in frequency and amplitude. A drawback of this approach is that it perpetuates 'edge effects' by which spurious features can be introduced by smoothing beyond the boundaries of the time dimension.

To locate ridges in the time-frequency space, we employed the 'crazy climbers' algorithm (Carmona *et al.*, 1998). This algorithm uses a simulated annealing method to identify features that are sharply defined in the frequency dimension but smooth in the time dimension. These features are then linked together to construct contiguous ridges that do not overlap. Finally, the modal signals are reconstructed using a smoothing spline procedure that projects each ridge back into a time series. In this way, the modal frequencies are characterized without making parametric assumptions about how their period and amplitude change over time.

The wavelet method requires no prior detrending or smoothing to remove baseline effects and noise. The ridge-finding and mode-reconstruction procedures will remove most random variability, which tends not to be localized to any particular region of the time-frequency space, and if applied within an appropriate bandwidth will ignore smooth baseline trends which appear as low-frequency features. The maximum frequency signal that can be detected is twice the sampling frequency. For the sake of computational efficiency, our implementation of the wavelet transform requires that the data series have regularly spaced time points. Consequently, missing data in the series must be imputed by some means.

We packaged the wavelet analysis in a function *waveclock* that performs these analytic steps in sequence, plots the results and returns several useful summaries of the output. The default settings are optimized for cell luminescence time-series cycling at circadian frequencies. The outputs include the instantaneous period, amplitude and phase of each mode, the mean and median period, and the variance of the reconstructed waves. By default, these outputs exclude measurements within the 'cone of influence' within $\sqrt{2}$ wavelengths from the beginning and end of the data where edge

---

*To whom correspondence should be addressed.

effects may apply (Torrence and Compo, 1998). This means that the time series must be at least $2\sqrt{2}$ times the wavelength of the lowest modal frequency in the signal. One approach to minimizing edge effects is to reflect the data at the boundaries of the time series prior to conducting the analysis: this is implemented in *waveclock* as a default procedure.

## 3  RESULTS

We demonstrated this analytic technique using experimental data generated to investigate the robustness of the circadian clock to genetic perturbations. Small interfering RNAs (siRNAs) were used to knockdown clock components in immortalized human osteosarcoma cells that are amenable to RNA interference (RNAi). Cells stably expressing luciferase under the control of the Bmal1 promoter were synchronized with dexamethasone, and subsequently had robust oscillations in bioluminescence with a period length of ∼24 h for more than 6 days in culture (Baggs *et al.*, *Elucidating network structures of the circadian clock underlying robustness*, Network features of the mammalian circadian clock). The luminescence was measured every 10 min for 120 h. The dataset and analytic functions are archived in the *R* library *waveclock*.

We demonstrated using a titration series that Bmal1 knockdown diminishes the amplitude of the oscillation in a dose-dependent manner (Supplementary Fig. 1A). Oscillations were both strong and persistent in the negative siRNA condition, whereas in the intermediate siBmal1 0.2X condition the oscillations were reduced in initial amplitude and quickly damped. No rhythmicity was detected in the siBmal1 1X condition. The instantaneous periods and amplitudes output from the *waveclock* function can be displayed graphically (Supplementary Fig. 1B). Also, the wavelet scalogram (time-frequency plot) can be interpreted directly (Supplementary Fig. 1C). Damping in the siBmal1 0.2X condition is evident from the fading intensity of color over time at circadian periods: this denotes decreasing amplitude. The modal frequency at close to 24 h period is denoted by a solid line and slopes upward showing its increasing period over time.

Although *waveclock* does not report error metrics for the estimates of period, amplitude or phase, the quantitative outputs can be used for hypothesis testing in a replicated experimental design. For example, using the variance of the reconstructed wave as a measure of the power at circadian frequencies, we showed not only that knocking down single genes such as Bmal1 and Cry1 decreased the amplitude of circadian oscillation compared with the negative siRNA condition, but also that their combinatorial knockdown further suppressed circadian oscillation relative to the single-gene conditions (Baggs *et al.*, ibid.).

The measurement of period is granular: its precision is governed by the number of 'voices' (frequency bins per octave) used in the wavelet transform. The default setting of 96 voices gives fine resolution in the frequency domain and is suitable for testing hypotheses about subtle changes in period length. For example, we were able to show that siCry1 knockdown decreases period length and siCry2 knockdown increases period length in a dose-dependent manner (Baggs *et al.*, ibid.). Lowering the number of voices gives a quicker and more robust analysis that can track modes even during relatively fast shifts in period length, at the expense of decreased precision. In an analysis of murine blood pressure using 12 voices per octave, we detected ultradian rhythms of ∼2 h

period superimposed on the diurnal rhythm (Curtis *et al.*, 2007; Supplementary Fig. 2). In arrhythmic Bmal1 KO mice, the ultradian rhythm is preserved but the diurnal rhythm is entirely absent. These physiological measurements are noisier than the bioluminescence data and the ultradian oscillations do not have a constant period, making them much harder to characterize.

As a further example of how the wavelet technique can identify superimposed oscillations at different frequencies, we detected both 24 h and 12 h rhythms in a transcript of the murine Dsc2 gene in mouse liver taken from a 48 h series (http://bioinf.itmat.upenn.edu/circa/; Supplementary Fig. 3).

## 4  DISCUSSION

Current commercial algorithms to measure circadian parameters from cellular luminescence data fit simple parametric models that assume either exponentially damped (e.g. Lumicycle; Actimetrics, Evanston IL) or constant amplitude (e.g. FFT-NLLS; Izumo *et al.*, 2006) and fixed period length (Supplementary Fig. 1B). Consequently the results are sensitive to the time window being analyzed when the period drifts. The *waveclock* method presents an attractive alternative by estimating period length and amplitude as smooth functions over time: these quantities can then be summarized as desired. The graphical display is useful as it can easily convey extreme changes due to technical issues (e.g. an opened incubator), or the presence of signals at other frequencies. Finally, we have shown proof of concept that this technique can be used to analyze physiological measurements, such as blood pressure and gene expression data. In the interest of public dissemination of research, we have made this application open source and available through the *R* network.

## REFERENCES

Carmona,R. *et al.* (1998) *Practical Time-Frequency Analysis: Gabor and Wavelet Transforms with an Implementation in S.* Academic Press, San Diego, CA.

Curtis,A.M. *et al.* (2007) Circadian variation of blood pressure and the vascular response to asynchronous stress. *Proc. Natl Acad. Sci. USA*, **104**, 3450–3455.

Hastings,M.H. (2005) Circadian biology: fibroblast clocks keep ticking. *Curr. Biol.*, **15**, R16–R18.

Izumo,M. *et al.* (2006) Quantitative analyses of circadian gene expression in mammalian cell cultures. *PLoS Comp. Biol.*, **2**, e136.

Nagoshi,E. *et al.* (2004) Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells. *Cell*, **119**, 693–705.

Rougemont,J. and Naef,F. (2007) Dynamical signatures of cellular fluctuations and oscillator stability in peripheral circadian clocks. *Mol. Syst. Biol.*, **3**, 93.

Stratmann,M. and Schibler,U. (2006) Properties, entrainment, and physiological functions of mammalian peripheral oscillators. *J. Biol. Rhythms*, **21**, 494–506.

Torrence,C. and Compo,G.P. (1998) A practical guide to wavelet analysis. *Bull. Amer. Meteorol. Soc.*, **79**, 61–78.

Yoo,S.H. *et al.* (2004) Period2: luciferase real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proc. Natl Acad. Sci. USA*, **101**, 5339–5346.