

Gene expression

TimeCycle: topology inspired method for the detection of cycling transcripts in circadian time-series data

Elan Ness-Cohn ^{1,2,3} and Rosemary Braun^{1,2,3,4,5,6,*}

¹Department of Molecular Biosciences, Northwestern University, Evanston, IL 60208, USA, ²Biostatistics Division, Department of Preventive Medicine, Northwestern University, Chicago, IL 60611, USA, ³NSF-Simons Center for Quantitative Biology, Northwestern University, Evanston, IL 60208, USA, ⁴Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, IL 60208, USA, ⁵Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA and ⁶Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL 60208, USA

*To whom correspondence should be addressed.

Associate Editor: Janet Kelso

Received on December 7, 2020; revised on May 5, 2021; editorial decision on June 22, 2021; accepted on June 25, 2021

Abstract

Motivation: The circadian rhythm drives the oscillatory expression of thousands of genes across all tissues. The recent revolution in high-throughput transcriptomics, coupled with the significant implications of the circadian clock for human health, has sparked an interest in circadian profiling studies to discover genes under circadian control.

Result: We present TimeCycle: a topology-based rhythm detection method designed to identify cycling transcripts. For a given time-series, the method reconstructs the state space using time-delay embedding, a data transformation technique from dynamical systems theory. In the embedded space, Takens' theorem proves that the dynamics of a rhythmic signal will exhibit circular patterns. The degree of circularity of the embedding is calculated as a persistence score using persistent homology, an algebraic method for discerning the topological features of data. By comparing the persistence scores to a bootstrapped null distribution, cycling genes are identified. Results in both synthetic and biological data highlight TimeCycle's ability to identify cycling genes across a range of sampling schemes, number of replicates and missing data. Comparison to competing methods highlights their relative strengths, providing guidance as to the optimal choice of cycling detection method.

Availability and implementation: A fully documented open-source R package implementing TimeCycle is available at: <https://nesscoder.github.io/TimeCycle/>.

Contact: rbraun@northwestern.edu

Supplementary information: [Supplementary data](#) are available at *Bioinformatics* online.

1 Introduction

Circadian rhythms—physiological, behavioral and metabolic oscillations with an approximate 24-h period—are controlled by an evolutionarily conserved set of core clock genes operating at the transcriptional and protein level. Entrained by Zeitgebers (external environmental stimuli such as light, temperature and food) that modulate time-of-day specific functions, the circadian clock orchestrates a multitude of cellular processes, including nearly half of genes across all tissues (Zhang *et al.*, 2014). Although various epidemiological studies have established significant links between circadian rhythms and human health (Braun *et al.*, 2018; Chang *et al.*, 2009; Kathale and Liu, 2014; Levine *et al.*, 2020; Levi and Schibler, 2007; Patke *et al.*, 2017; Puttonen *et al.*, 2010; Roenneberg *et al.*, 2007; Videnovic *et al.*, 2014; Zhang *et al.*, 2014), the underlying

biological mechanisms coupling these phenomena remain poorly understood.

Facilitated by the development of high-throughput assays, researchers can now interrogate biological mechanisms at the molecular level by analyzing transcriptomic time-series data to identify genes under circadian control. This capability presents researchers with new analytical challenges of how best to reliably extract rhythmic signals from transcriptomic time-series data (Hutchison and Dinner, 2017). First, experimental costs constrain the frequency and length of sampling, requiring conclusions to be made from sparse or short time-series measurements. Second, circadian expression profiles often do not follow precise sinusoidal waveforms, but exhibit asymmetries, sharp peaks, additive trends and noisy fluctuations (Michael *et al.*, 2008; Ness-Cohn *et al.*, 2020; Thaben and Westermark, 2014, 2016; Wu *et al.*, 2016). To address these

challenges, a variety of cycling detection methods have been proposed (Ahdesmäki *et al.*, 2007; de Lichtenberg *et al.*, 2005; Hughes *et al.*, 2010; Hutchison *et al.*, 2018; Levine *et al.*, 2002; Michael *et al.*, 2008; Perea *et al.*, 2015; Straume, 2004; Thaben and Westermark, 2014; Wichert *et al.*, 2004; Wu *et al.*, 2016; Yang and Su, 2010). However, benchmarking studies indicate that no method is universally optimal, and different algorithms may yield conflicting results on the same data (Deckard *et al.*, 2013; Laloum and Robinson-Rechavi, 2020; Mei *et al.*, 2021; Ness-Cohn *et al.*, 2020; Wu *et al.*, 2014).

Early cycling detection methods used parametric approaches employing auto-correlation, curve fitting and Fourier analysis to decompose expression patterns into harmonic components of varying amplitude and phase (Levine *et al.*, 2002; Straume, 2004; Wichert *et al.*, 2004). These methods could successfully identify genes with relatively symmetric waveform shapes, but had poor performance on asymmetric and sharply peaked waveforms (Michael *et al.*, 2008; Thaben and Westermark, 2014). To address these shortcomings, a range of non-parametric methods were developed (Hughes *et al.*, 2010; Hutchison *et al.*, 2018; Perea *et al.*, 2015; Thaben and Westermark, 2014; Wu *et al.*, 2016; Yang and Su, 2010). While non-parametric methods sacrifice statistical power relative to parametric methods (Laloum and Robinson-Rechavi, 2020), they are applicable to a broader range of cycling waveforms and generally outperform parametric methods (Hughes *et al.*, 2010; Hutchison *et al.*, 2015, 2018; Yang and Su, 2010).

Nevertheless, limitations remain in current methods (Hughes *et al.*, 2017; Hutchison and Dinner, 2017; Laloum and Robinson-Rechavi, 2020; Ness-Cohn *et al.*, 2020). A common approach is to compare the observed gene expression time-series to a user-defined set of reference waveforms (e.g. oscillations with different degrees of sawtooth asymmetry) to calculate a periodicity score (e.g. via the Kendall τ rank correlation). Such methods may fail to detect cyclic patterns that do not fall into the predetermined profiles of reference signals, effectively limiting the scope of discovery. Moreover, these methods typically assess the statistical significance relative to a null model of a randomized time-series. Since randomized time-series may jump from low to high gene expression faster than biological translation and degradation processes allow, these methods may produce unrealistic null models and misleading significance tests.

In addition to methodological hurdles, there are also considerations of various methods' abilities to handle replicates, uneven sampling, missing data and computational efficiency. In practice, the ability of methods to adequately handle these features directly affects flexibility in experimental design. For instance, a method that can accommodate uneven sampling can allow for dense sampling at times of interest, with sparser sampling at other times. Because missing data often occur as a result of sequencing errors with greater likelihood as sample size increases (Gierliński *et al.*, 2015), researchers benefit from algorithms that can adequately handle missingness. Finally, computational efficiency allows sample sizes to grow while still processing the data in a reasonable amount of time.

Results from dynamical systems theory and topological data analysis provide alternative strategies to overcome the limitations of curve-fitting and template-based analysis methods. The recent 'SW1PerS' method (Perea *et al.*, 2015) uses time-delay embedding (Takens, 1981) to transform sliding windows of the time-series into a high-dimensional point cloud, and quantifies the circularity of the transformed signal as a measure of its periodicity. However, because the rhythmicity scores do not follow a well-defined distribution and the computational complexity of the algorithm precludes permutation tests, SW1PerS does not provide a *P*-value testing whether a gene is cycling.

To address these challenges, we introduce TimeCycle: a non-parametric, template-free algorithm based on topological data analysis with a bootstrapping procedure for statistical inference of cycling genes. Results demonstrate that TimeCycle reliably discriminates cycling and non-cycling profiles in synthetic data and reproducibly detects known circadian genes in experimental data. Below we describe the method, illustrate its application to multiple datasets, and provide a comparison to several competing methods.

2 Materials and methods

2.1 Generating synthetic data

A total of 144 unique synthetic time-course datasets, each comprising 11 000 expression profiles, were generated in R using the code outlined in (Ness-Cohn *et al.*, 2020). Each dataset consisted of a different number of replicates (1, 2, 3), sampling intervals (1-h, 2-h, 4-h), sampling durations (36-h, 48-h, 72-h, 96-h) and noise levels (10%, 20%, 30%, 40% of the waveform amplitude). Within each condition, 11 base waveforms were simulated to mimic expression patterns observed in nature: periodic patterns, non-periodic patterns, and dynamics that have a cyclic component but do not meet the strict definition of periodicity. Seven of these 11 shapes were considered cyclic (sine, peak, sawtooth, linear trend, damped, amplified, contractile), and 4 were considered non-cyclic (flat, linear, sigmoid and exponential). Further details may be found in (Ness-Cohn *et al.*, 2020).

2.2 Processing synthetic data

All 144 synthetic datasets were processed by all four cycling detection methods [TimeCycle, JTK_CYCLE (Hughes *et al.*, 2010), RAIN (Thaben and Westermark, 2014), GeneCycle (Ahdesmäki *et al.*, 2007) and SW1PerS (Perea *et al.*, 2015)], using each method's recommended parameter settings as defined by the method's documentation. Since GeneCycle and SW1PerS do not have a built-in function for dealing with replicates, replicates were averaged together, following the recommended common practice in the field (Hughes *et al.*, 2017; Ness-Cohn *et al.*, 2020). TimeCycle, JTK_CYCLE and RAIN used the replicate procedures recommended in their documentation. See the Supplement and <https://github.com/nesscoder/TimeCycle-data> for a complete list of the experimental parameters and associated source code.

2.2.1 Computing ROC, AUC and percent correct classification

For each method across all 144 synthetic datasets and missing data analysis, the receiver operating characteristic (ROC) and accompanying area under the curve (AUC) were computed using the pROC R package (Robin *et al.*, 2011). The optimal threshold for computing each method's percent correct classification is defined by Youden's *J* statistic (Youden, 1950), again computed using pROC. For each waveform shape at the classification threshold defined by the Youden's *J* index, the percent correct classification was computed by dividing the number of synthetic genes called as (non-)cyclers out of the total possible (non-)cyclers.

2.2.2 Missing data analysis

To evaluate TimeCycle's ability to handle missing data, we analyzed synthetic data sampled every 2-h for 48-h with one replicate at varying noise levels with a range of missing values. For each of the 11 000 genes in each dataset, time-points were randomly removed by sweeping from 0% missingness until only 2 points remained. ROC plots and AUC scores depicted in Figure 4 were computed as described above.

2.2.3 Outlier analysis

To evaluate TimeCycle's ability to handle outliers, we analyzed 36 synthetic datasets all with a 48-h sampling length. Each dataset consisted of a different number of replicates (1, 2, 3) and sampling intervals (1-h, 2-h, 4-h). Following (Hughes *et al.*, 2010) and (Salkovic *et al.*, 2020), for all sampled time-points, outliers were injected into the time-series at a rate of 1%. Outliers were drawn from a uniform distribution between $[\mu - 4\sigma, \mu - 3\sigma]$ and $[\mu + 3\sigma, \mu + 4\sigma]$ of the diurnal mean (μ) for each time-series (Salkovic *et al.*, 2020).

2.3 Preprocessing and analysis of microarray data

Microarray data was preprocessed as described in *Methods—Preprocessing Microarray Data* of Ness-Cohn *et al.* (2020). (See <https://github.com/nesscoder/TimeTrial> for associated source code.)

To characterize the effects of sampling schemes using real data, the three datasets were processed to ensure comparability across datasets. The Hughes (2012) and Zhang (2014) datasets were sampled every 2-h for 48-h. The Hogenesch 2009 (Hughes *et al.*, 2009) dataset, comprising data sampled every 1-h for 48-h, was down-sampled into 2 datasets sampled every 2-h for 48-h. All four datasets (2 original and 2 down-sampled) were processed by TimeCycle, JTK_CYCLE, RAIN and GeneCycle, using each method's recommended parameter settings as defined by the method's documentation (SW1Per5 was omitted from the microarray analysis since the algorithm does not produce a P -value and thus could not be meaningfully compared to the other methods). Further details may be found in the Supplement.

2.3.1 Gene dynamics reproducibility assessment via rank correlation CDF

The distribution of Spearman rank correlations was computed by comparing genes identified by each method with an $FDR < 0.05$ and $LogFC < 2$ in one dataset (e.g. Hughes *et al.*, 2009; Hughes 2012; Zhang 2014) with their pairwise expression in the other datasets. To account for the 6-h phase shift in sampling start time between the Hughes dataset (start 0_{ZT}) and the Hogenesch/Zhang (start 18_{ZT}) datasets, the last three time points (e.g. 6-h) in the Hughes dataset were moved to the start of the time-series before computing the correlation (Fig. 5D). The distribution of cycling gene correlations was compared to both the all-gene correlations and a null distribution of correlations. The all-gene correlation represents the rank correlation on a per-gene basis for all 12 868 genes, whether or not they were identified as cycling. The null distribution of correlation coefficients was constructed by randomly resampling the 12 868 genes in each dataset (thereby generating correlation coefficients from different, rather than the same, genes).

3 Results

TimeCycle is a method for classifying and quantifying cyclic patterns of gene expression in transcriptomic time-series data. Application to both synthetic and experimental data demonstrate TimeCycle's ability to efficiently discriminate cycling genes from non-cycling genes across a range of sampling schemes, number of replicates and missing data.

3.1 The TimeCycle method

TimeCycle's basic framework consists of a rescaling/normalization step, reconstruction of the state space via time-delay embedding, isolation of non-linear patterns via manifold learning and dimension reduction, quantification of the circularity of the signal using persistent homology, and comparison of that measure to a bootstrapped null distribution to assess statistical significance (Fig. 1B). We describe the main steps of the method here; a complete description pertaining to the methods preprocessing, null distribution and period estimation can be found in the supplement and TimeCycle's documentation available at <https://nesscoder.github.io/TimeCycle/>.

3.1.1 Reconstructing dynamical cycles via Takens' theorem

TimeCycle exploits Takens' theorem, a result from dynamical systems theory that proves that the time-delay embedding of a single variable observed over time will reconstruct (up to diffeomorphism) the state space of a multivariate dynamical system (Takens, 1981). A d -dimensional time-delay embedding of a time-series X is defined as a representation of that time-series in a d -dimensional space where each point is given by the coordinates $(x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-d\tau})$. As a consequence of Takens' theorem, a dynamical system with a cycle will exhibit circular patterns in the time-delay space, whereas non-periodic signals will form a mass (Fig. 1A). Using this logic, TimeCycle reconstructs the state space for each gene using time-delay embedding and quantifies the circularity of the embedding as a measure of the evidence of cycling dynamics.

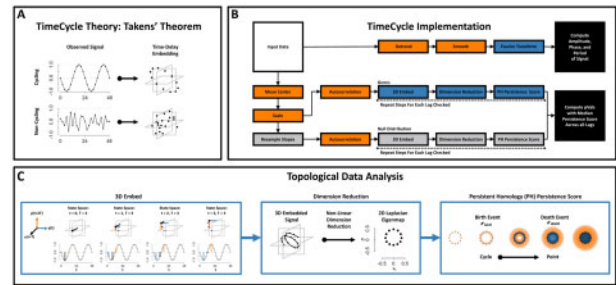


Fig. 1. TimeCycle Theory and Implementation. (A) Takens' theorem proves that the dynamics of a periodic signal will exhibit circular patterns in the embedded space. A periodic function (Top) embedded in the state space forms a circle. A non-periodic function (Bottom) embedded in the state space forms a mass. (B) The TimeCycle algorithm assesses statistical significance with respect to a null distribution of persistence scores generated from random time-series with the same marginal distribution of finite differences as the original time-series; this ensures that the random time-series are constrained by the same transcription and degradation rates present in the original data. Period, phase and amplitude are estimated separately using a fast Fourier transform procedure. (C) Left: For a given time-series, the method reconstructs the state space using time-delay embedding, a data transformation technique for dynamic systems. Middle: Dampened and trending signals form spirals and helices in the embedded space. Non-linear dimension reduction using Laplacian Eigenmaps yields embedded points preserving the local circular geometry of the 3-D helix in a 2-D space. Right: Persistent Homology (PH) Persistence Score. Points in the embedding are connected if they are at most a distance 2ϵ apart, as the radius ϵ increases. A topological cycle appears at some ϵ_{birth} and disappears at ϵ_{death} , and a persistence score is calculated as $Persistence = \epsilon_{death} - \epsilon_{birth}$. Periodic signals have high persistence, while non-periodic signals have low persistence.

3.1.2 Parameter choice and detrending via dimension reduction

Time-delay embedding requires two parameters, the embedding dimensionality d and the delay lag τ . A perfectly sinusoidal signal will form a circle in $d=2$ dimensions, and hence a two-dimensional embedding would be sufficient in the ideal case. However, biological signals are often not strictly periodic, but exhibit drifts in the oscillation. A common approach is to detrend the time-series by fitting it to a line and analyzing the residuals (Yang and Su, 2010). While this removes linear trends, the detrending procedure can introduce false positives (Ness-Cohn *et al.*, 2020), as non-rhythmic signals—e.g. sigmoidal and exponential—may appear rhythmic after detrending (Supplementary Fig. S1). Instead, we observe that any non-periodic component of the signal may be represented by higher dimensions in the embedded space. Hence, we embed the time-series in three dimensions (Fig. 1C: Left) and use non-linear dimension reduction to recover the periodic component. An illustrative example is given in Figure 1C (Middle). A drifting oscillation will form a helix in the 3-D embedded space, with the linear trend contributing to the elongation of the helix along one coordinate and the circular component preserved in the other two. TimeCycle uses Laplacian Eigenmaps (Belkin and Niyogi, 2003), a non-linear dimension reduction (NLDR) technique, to project the 3-D embedded data back into a 2-D space, preserving the circular geometry.

This approach has two advantages over linear-fit detrending. First, it is in principle applicable to any type of drift, and is not necessarily confined to removing linear trends. Second, it will only yield a circular pattern in the 2-D space if the periodic component is strong relative to the drift. That is, if the 'drift coordinate' (e.g. the axis of the helix in Fig. 1C) more faithfully preserves the local geometry of the data than the circular 'cycling coordinates', it will be chosen by the NLDR, and the resulting 2-D representation will not have a detectable circular topology. Together, these features overcome the drawbacks of linear-fit detrending.

The optimal choice for the other parameter, the lag τ , is less obvious. While the underlying mathematical theory holds for all $\tau > 0$ (Takens, 1981), in practice the presence of noise and short time-series lengths lead to differences in the embedding as a function of τ (Garland *et al.*, 2016). Because the underlying theory is insensitive to τ , it yields no guidance on how to derive an 'optimal' lag or set

criteria for its selection. A common approach, due to Fraser and Swinney, is to choose τ to maximize the independence between embedded points (Fraser and Swinney, 1986); more recently, Meiss and colleagues proposed a method to choose τ that minimizes the curvature (Deshmukh *et al.*, 2020). In practice, however, the best choice is often application-dependent. Because our goal here is to detect rhythmicity (rather than fully reconstruct the state space), we sweep through values of τ , computing the degree of circularity for each. If a signal is truly rhythmic, on average the state space manifold across lags should be approximately circular.

3.1.3 Quantification of cycling via persistent homology

We quantify the circularity of the embedded data using *persistent homology* (PH) (Edelsbrunner *et al.*, 2002), an algebraic means to measure topological structures (i.e. components, holes, voids and higher dimensional analogs) in a dataset. From each point of the embedded signal, a d -sphere of radius ϵ is incrementally grown (Fig. 1C: Right). When a pair of points have intersecting spheres (i.e. are at most 2ϵ apart), a line is drawn between those points. When a complete cycle is formed amongst a set of points at some ϵ_b , this is defined as a ‘birth event’: the initial appearance of a topological hole, a closed cycle lying entirely inside the spheres surrounding the embedded data. As ϵ is increased, the topological hole eventually closes at a radius ϵ_d defining the ‘death event’. The persistence score of a topological feature is defined as the difference between the death and birth radii $\epsilon_d - \epsilon_b$. Several such features, with associated persistence scores, may be present in the data, the largest of which corresponds to the most unambiguously periodic dynamics. A rhythmic signal will produce a more circular embedding, resulting in a larger maximum persistence score, than a non-rhythmic signal.

3.1.4 Significance testing

A common approach for assessing statistical significance in cycling detection algorithms is to compare the statistics of an observed signal to a null model comprising random time series (Hughes *et al.*, 2010; Hutchison *et al.*, 2018; Thaben and Westermark, 2014). Such null models may be unrealistic, however, since gene expression in the biological context is constrained by transcription and degradation rates, and an assumption of random, independent time-series may contain changes in gene expression that are biologically unattainable. Instead, TimeCycle’s null model is obtained by permuting the finite differences in gene expression between sampled time-points. The distribution of gene expression changes under the null is thus identical to that of the observed data, ensuring that the resulting random time-series reflect biological transcription and degradation constraints (Fig. 1B). A persistence score is computed for each resampled time-series to generate a reference distribution of persistence scores under the null (no cycling), conditioned on the observed rates of change. The observed persistence score is then compared to this reference distribution to obtain a P -value.

A significant consideration for this method is the computational complexity of the persistence score calculation [$\mathcal{O}(N \log N)$ in the number of time points for $2-d$ data], which will need to be performed for each gene and each resampled time-series. Indeed, a drawback of the SW1PerS algorithm is its inability to compute a P -value due to the computational cost of the method. To overcome this challenge, we devised a scheme to allow for hypothesis testing with improved computational efficiency: prior to the embedding and PH analysis, each gene is mean centered and scaled to unit variance, allowing a common set of permuted time-series to be used for all genes to test the statistical significance of cycling.

3.1.5 Parameter estimation

Often, researchers desire estimates of the period, phase and amplitude of the genes detected as cycling. Because the time-delay embedding and PH computation does not provide these estimates, a separate computation in TimeCycle is used to generate these results. To estimate the period, amplitude and phase of the oscillations (Fig. 1B), signals are linearly detrended and smoothed via a moving average before fitting the signal to the first three harmonics of the fast

Fourier transform (FFT). The period, phase and amplitude are computed from the FFT fit. (Note that when performing the less computationally costly parameter estimation, the rescaling procedure used in the cycle detection step is omitted to preserve amplitude measures.)

3.1.6 Replicate time-points

A practical consideration when designing transcriptomic time-series experiment is the trade-off between replicates, sampling resolution and sampling length in relation to experimental cost (Hughes *et al.*, 2017; Ness-Cohn *et al.*, 2020). While cycling detection methods do not require replicates (Hughes *et al.*, 2010; Hutchison *et al.*, 2015, 2018; Perea *et al.*, 2015; Thaben and Westermark, 2014; Wu *et al.*, 2014; Yang and Su, 2010); technical replicates are necessary for performing additional analysis beyond the scope of cycling detection—e.g. differential expression and differential time-course profiles (Anders and Huber, 2010; Love *et al.*, 2014). As such, researchers benefit from algorithms that can incorporate replicates to improve detection of cycling genes (Hutchison *et al.*, 2015). Adhering to best practice guidelines outlined in previous studies (Hughes *et al.*, 2017; Ness-Cohn *et al.*, 2020), TimeCycle averages replicate time-points. The averaged signal is then processed following all steps outlined above.

3.2 Application and evaluation using simulated gene expression data

To comprehensively evaluate TimeCycle’s capabilities across different patterns of temporal gene expression, we applied TimeCycle and four other methods—JTK_CYCLE (Hughes *et al.*, 2010), RAIN (Thaben and Westermark, 2014), GeneCycle (Ahdesmäki *et al.*, 2007) and SW1PerS (Perea *et al.*, 2015)—to synthetic data described in our previous benchmarking work (Ness-Cohn *et al.*, 2020). JTK_CYCLE and RAIN were selected for comparison as the primary methods used in the field of circadian rhythm detection. GeneCycle was selected as an alternative method highlighted in previous benchmarking studies to be robust to outliers (Ahdesmäki *et al.*, 2007; Deckard *et al.*, 2013; Laloum and Robinson-Rechavi, 2020). SW1PerS was selected as the only other method utilizing topological data analysis. While other methods such as ARSER (Yang and Su, 2010) and booteJTK (Hutchison *et al.*, 2018) were considered based on previous benchmarking studies (Deckard *et al.*, 2013; Laloum and Robinson-Rechavi, 2020; Ness-Cohn *et al.*, 2020; Wu *et al.*, 2014), they were ultimately omitted from analysis as they do not natively handle missing data. Synthetic datasets varied in number of replicates (1, 2, 3), sampling intervals (1-h, 2-h, 4-h), sampling length (36-h, 48-h, 72-h, 96-h) and noise levels (10%, 20%, 30%, 40% of signal amplitude), across 11 base waveform shapes. Seven of these 11 shapes were considered cyclic (sine, peak, sawtooth, oscillations about a linear trend, damped, amplified, contractile) and 4 were considered non-cyclic (flat, linear, sigmoid and exponential). For each waveform in each condition, 1000 ‘genes’ were simulated with varying amplitudes, phases and shape parameters (e.g. the envelope for damped/amplified waves), yielding in total 11 000 simulated genes for each of the 144 sampling and noise conditions. [Further details of the synthetic datasets can be found in the Methods section and Supplementary Materials of Ness-Cohn *et al.* (2020).]

3.2.1 Accuracy and sampling considerations

A summary of TimeCycle results in the synthetic data is found in Figure 2. Receiver-operating characteristic (ROC) curves were computed for each of the 144 synthetic datasets, and the area under the curve (AUC) was used to compare classification accuracy across methods. An AUC of 1 represents perfect classification, while an AUC of 0.5 represents a classification no better than pure chance. AUC values are shown as a function of the number of samples collected per time-series. TimeCycle exhibits AUCs ≥ 0.8 across most sampling schemes, with the best performance observed for data sampled every 2-h for 48-h—the scheme primarily used in previous

benchmarking studies (Deckard *et al.*, 2013; Hutchison *et al.*, 2018; Ness-Cohn *et al.*, 2020; Perea *et al.*, 2015; Thaben and Westermarck, 2014; Wu *et al.*, 2014).

Examining the type I and type II error rates relative to the FDR adjusted P -values across the varying noise levels, we find that TimeCycle and JTK_Cycle are more conservative in comparison to RAIN and GeneCycle (Fig. 2B). This can also be seen in Supplementary Figure S2, which illustrates that the null distribution of P -values is biased toward 1 for both TimeCycle and JTK_Cycle (more strongly biased for JTK_Cycle), whereas small P -values are overrepresented under the null for RAIN and GeneCycle (leading to a greater chance of false positives) for non-cycling genes. Across all methods, false negative error rates increase with noise. (SW1PerS was omitted from the type I and type II error analysis as it does not produce P -values, but rather a periodicity score.)

Outside the 48-h sampling schemes, TimeCycle exhibits good performance for longer time-series sampled every 1 or 2-h, but reduced performance for sparser sampling (every 4-h compared to 1-h and 2-h), even for longer time-series. This is attributable to sparsely sampled manifolds in the reconstructed state space, resulting in an insufficient number of data-points for cycle formation. This suggests that, for a fixed number of samples, denser sampling for a shorter duration may be favorable to longer, sparser sampling.

It is also instructive to examine how the classification accuracy for different waveforms changes with the sampling strategy. We find that for time-series lasting 36-h, strong symmetric cyclers are robustly detected, while asymmetric cyclers (such as the sawtooth) do not have sufficient points to close the cycle in the embedded space (that is, observations along the ‘sharp’ rise or fall of an asymmetric waveform may be missed when fewer than two complete periods are assayed, leading to a ‘C’ shape in the embedded space). At 72-h and 96-h, linear trending and damped oscillations are more likely to be classified as non-cycling due to the fact that as the time-series sampling is extended, the trends and dampening become more pronounced, dominating the underlying oscillation. Because these signals are not in fact strictly periodic (returning to precisely the same state at regular intervals), this can be a desirable or undesirable behavior depending on whether the user wishes to classify linear trending and damped oscillations as cycling. Details for an implementation with alternative parameters to detect oscillations with linear trends can be found in TimeCycle’s documentation.

3.2.2 Comparison to other methods

To compare TimeCycle’s ability to distinguish cycling from non-cycling waveforms to that of other methods, we examined the pairwise comparison of raw P -values generated by each method. To simulate results for an ideal dataset, comparisons were made using the recommended sampling scheme—every 2-h for 48-h with one replicate—under low (10%) noise conditions (Fig. 3). (Similar plots for varying noise levels can be found in Supplementary Figs S3–S5).

The lower triangle of Figure 3 depicts the pairwise scatterplot of the $-\log_{10}P$ for each synthetic gene across methods, with the larger points designating the average $-\log_{10}P$ for a given waveform shape. (Note that, as previously discussed, SW1PerS does not compute a P -value; instead, the periodicity score is shown.) The results are generally well-correlated, with non-cycling waveforms clustering at low significance and sinusoidal waveforms ranked high; differences between methods are most noticeable for trending, sawtooth and contractile waveforms. Given the algorithmic similarities of RAIN/JTK_CYCLE (via the Jonckheere-Terpstra test) and TimeCycle/SW1PerS (via topological analysis), it is unsurprising that these methods are more highly correlated within method pairings than across methodological approaches.

The upper triangle of Figure 3 measures the fraction of correctly classified signals for each waveform shape, i.e. the conditional probability of correctly classifying the gene as cycling or non-cycling given the waveform. Here, the decision threshold for classifying a given gene as cycling was chosen via Youden’s J statistic (Youden, 1950) computed from AUC over all genes. For the sampling scheme of 48-h every 2-h, TimeCycle’s accuracy was comparable to other methods for most waveform shapes with three exceptions: First,

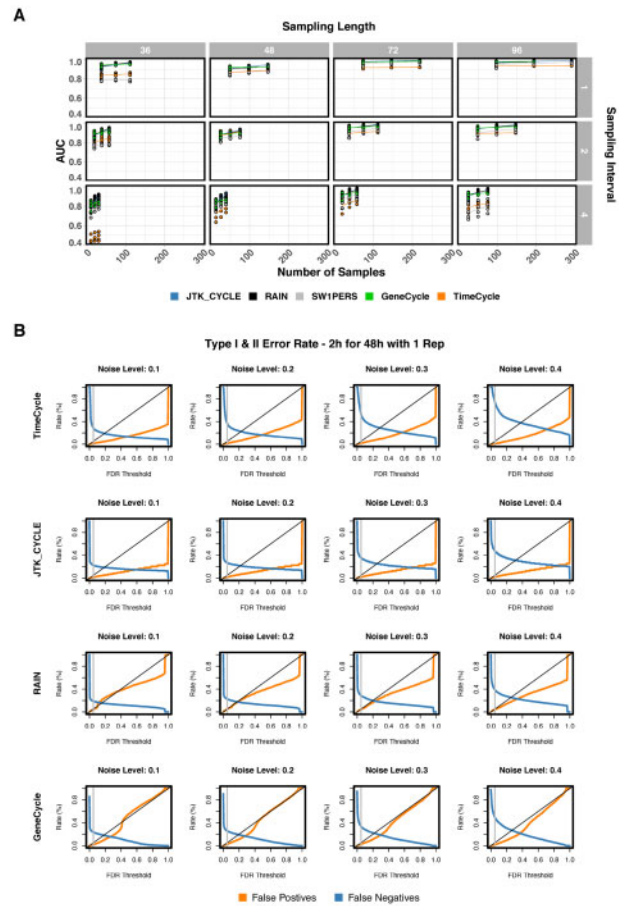


Fig. 2. Synthetic data overview. (A) A total of 144 unique synthetic time-course datasets were generated in R with known ground truth, as described in Ness-Cohn *et al.* (2020). Each dataset consisted of a different number of replicates (1,2,3), sampling intervals (1-h, 2-h, 4-h), sampling lengths (36-h, 48-h, 72-h, 96-h) and noise levels (10%, 20%, 30%, 40%) as a percentage of the waveform amplitude. AUC scores for each method were computed across all 144 datasets. Lines represent the best linear fit across replicates and noise levels within a specified sampling scheme. Replicate time-series were averaged together for the GeneCycle and SW1PerS algorithm, since neither algorithm has a built-in method for handling replicates. (B) Type I and Type II error rates at varying FDR thresholds across methods for synthetic data sampled every 2-h for 48-h with 1 replicate. FDR = 0.05 marked by gray vertical line. (SW1PerS does not compute a P -value, but rather a periodicity score and was thus omitted from analysis.)

TimeCycle shows improved performance in detecting peaked and contractile waveforms across methods when samples are taken every 2-h across 48-h. Second, TimeCycle also shows improved detection of peaked waveforms in comparison to JTK_CYCLE, RAIN and GeneCycle, with a slight decrease in comparison to SW1PerS. Third, TimeCycle shows decreased performance in detecting linear trends across all methods. To mitigate this shortcoming, alternative parameters for the improved detection of genes with linear trends can be found in TimeCycle’s documentation. While TimeCycle outperforms other methods in certain circumstances, our result suggest that no method is uniformly best at picking up all types of circadian dynamics under all sampling schemes.

3.2.3 Missing data and outliers

Since missing data often occur as a result of sequencing errors with greater probability as the sample size increases (Gierliński *et al.*, 2015), researchers benefit from algorithms that can handle missingness. TimeCycle imputes the missing points via the linear interpolation algorithm as described by Moritz *et al* in their imputeTS R package (Moritz and Bartz-Beielstein, 2017). However, by introducing imputed data, one also introduces new assumptions—

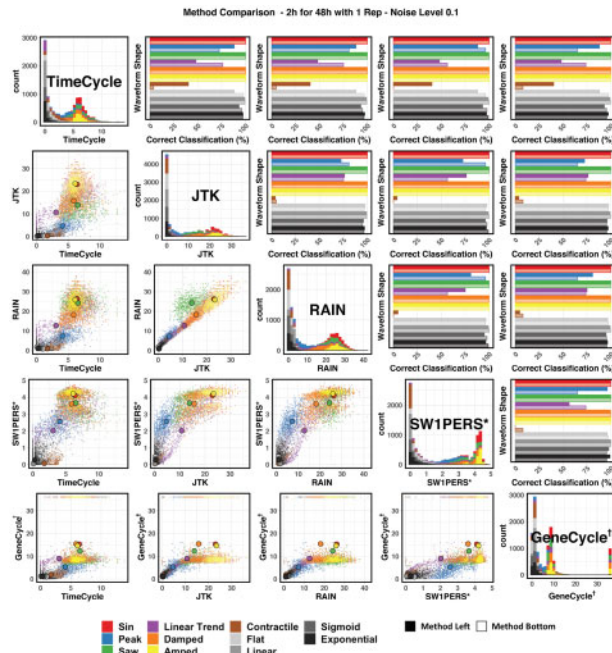


Fig. 3. 48 h Synthetic Data Method Comparison—Noise Level 0.1. Comparison of TimeCycle, JTK_CYCLE, RAIN, SWIPERS* and GeneCycle† results for synthetic data of various waveform shapes. LOWER TRIANGLE: scatterplot of $-\log_{10}P$ for each synthetic gene. The larger colored points represent the average $-\log_{10}P$ designated by waveform shape. DIAGONAL: Histogram of $-\log_{10}P$ for each synthetic gene by method. UPPER TRIANGLE: Classification accuracy comparison by waveform category. Darker bars correspond to the method listed in the row; lighter bars correspond to the method listed in the column. Each value corresponds to the fraction of correctly classified genes (true positive for cycling/non-cycling as appropriate) for a given waveform type, using a classification threshold obtained from Youden's J index for the ROC across all genes. * SWIPERS* does not compute a P -value, but rather a periodicity score. † GeneCycle results of $P=0$ were set to machine precision (2.2×10^{-16}) for visualization purposes.

specifically, smoothness in the trajectory and hence small gene expression changes before/after the missing observation. These properties must be shared by the null model to ensure that the significance testing is not biased by the imputation. Hence, we perform the previously-described resampling procedure post-imputation, effectively taking the imputation assumptions into account when defining the null distribution.

To investigate how TimeCycle performs when data are missing, we simulated missingness in the synthetic datasets (Fig. 4). For each of the 11 000 genes, we randomly removed points sweeping from 0% missingness until only 2 points remained. The AUC was computed at each missingness value (Fig. 4A). At 50% missingness in the synthetic data, TimeCycle shows robust classification with a slow decrease in AUC for each noise level. We observe that the difference in AUC between noise levels is greater than the degradation in AUC as a function of missingness (Fig. 4A and B). The differences are primarily attributable to decreases in sensitivity, with high levels of noise and missingness compromising the ability to detect cycling genes. We find little decrease in specificity, confirming that the 'smoothing' induced by the imputation procedure does not generate false positives. Missing data analyses for JTK_CYCLE, RAIN and GeneCycle can be found in Supplementary Figures S6–S8.

As an additional check of method robustness, we performed an outlier analysis across all 36 variants of the 48-h sampling lengths. Outliers were injected into the time-series at a rate of 1% of all sampled time-points and were drawn from a uniform distribution between $[\mu - 4\sigma, \mu - 3\sigma]$ and $[\mu + 3\sigma, \mu + 4\sigma]$ of the diurnal mean (μ) for each time-series, following previous studies (Hughes *et al.*, 2010; Salkovic *et al.*, 2020). TimeCycle exhibited a very slight and non-significant decrease in average AUC \pm SD across all noise levels (0.86 ± 0.05 versus 0.87 ± 0.05 , with and without outliers,

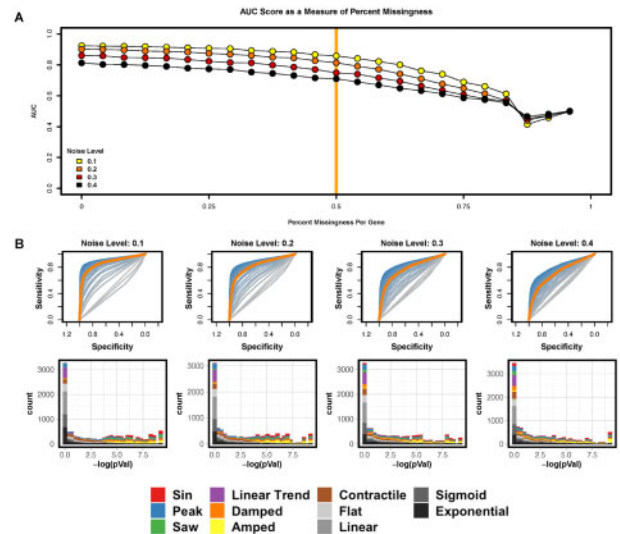


Fig. 4. Missing data results. (A) TimeCycle AUC scores across varying levels of percent missingness per gene when sampled every 2-h for 48 h with one replicate. Results are shown for each level of noise. 50% missingness is highlighted by the vertical orange line. (B) TOP: ROC curves for each percent missingness depicted in the AUC score plot in panel A above. ROC curves scaled from Blue (0% missingness) to gray (96% missingness). 50% missingness is again highlighted by the orange curve. BOTTOM: Histogram of $-\log(p)$ at the different noise levels corresponding to the orange 50% missingness ROC plot.

respectively); other methods also remained unchanged (Supplementary Fig. S9). As expected from the unchanged AUC, the type I and type II error rates were also unaltered. This suggests that the performance of TimeCycle, as well as other methods, is robust to outliers.

3.3 Application to biological data: reproducibility analysis

While synthetic data has the advantage of known ground truth, it also has the drawback of not necessarily being representative of the real biological datasets. On the other hand, measuring a method's accuracy using real data is limited, as the ground truth is not generally known. Instead, one may test the reproducibility of the results, under the assumption that a true biological signal should be consistently detected across multiple studies of the same condition.

To this end, we took a 'cross-study concordance' approach in which we tested the ability of each method to consistently characterize a set of 12 868 genes measured in three independent mouse liver time-series expression sets (Hughes *et al.*, 2009, 2012; Zhang *et al.*, 2014). We expect that a method that accurately detects cycling genes should do so reproducibly in all three datasets, whereas a method that is no better than chance will yield divergent results in the three datasets. Moreover, the genes detected as cycling in common across the dataset should exhibit similar dynamics across all three datasets. By contrast, if a method is overly permissive, leading to a high false-positive rate, genes may be classified as 'cycling' across datasets without actually having any commonality in their dynamics. We evaluate both of these by (i) examining the overlap in genes classified as cycling in the various datasets [following the Methods of Ness-Cohn *et al.* (2020)] and (ii) computing the rank correlation ρ of the expression profiles of the genes detected as cycling at FDR < 0.05 and LogFC > 2 . Together, our analysis quantifies whether the cycling detection is reproducible.

The datasets used in this analysis were comparable, but independent time-series studies of gene expression in mouse liver, referred to hereafter as Hogenesch (Hughes *et al.*, 2009), Hughes (Hughes *et al.*, 2012) and Zhang (Zhang *et al.*, 2014). These datasets are summarized in Table 1. The Hogenesch dataset (Hughes *et al.*, 2009) was downsampled to two datasets sampled every 2-h

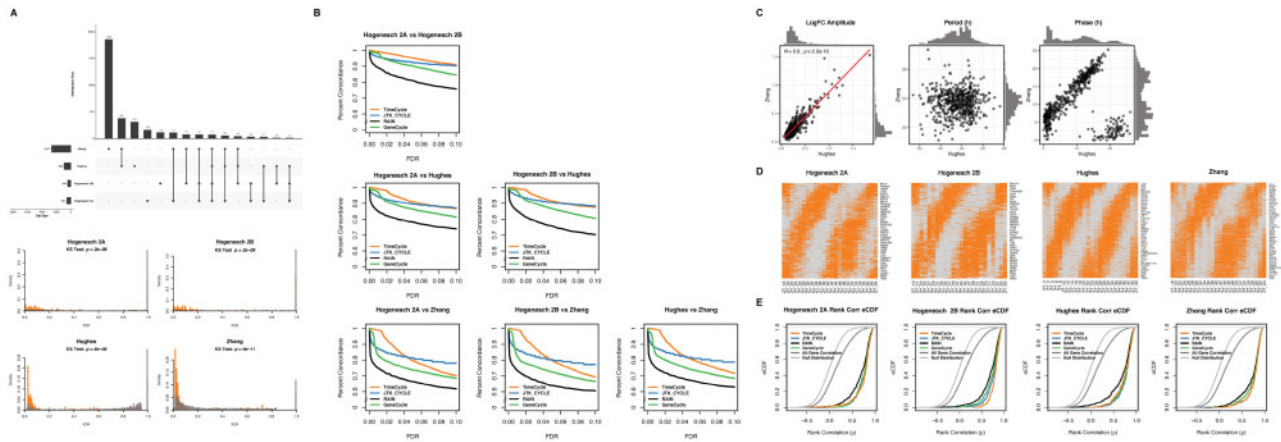


Fig. 5. TimeCycle Biological Data Results. (A) TOP: Upset plot showing overlap between significant cycling genes sets detected by TimeCycle at an FDR < 0.05 across 3 distinct mouse liver time-series datasets as described in (Ness-Cohn *et al.*, 2020). The Zhang and Hughes datasets were sampled every 2-h for 48-h. The Hogenesch study, sampled every 1-h for 48-h, was downsampled into two datasets sampled every 2-h for 48-h (Hogenesch 2A and Hogenesch 2B). BOTTOM: Distribution of FDR-adjusted P -values of the known circadian genes (orange) versus all genes (gray). Circadian genes were extracted from the CGDB database (Li *et al.*, 2017). Only circadian genes that were experimentally validated in mouse liver tissue through low-throughput methods were included in the analysis. (B) Percentage of genes concordantly called cycling or non-cycling across studies at varying FDR thresholds in each pair of studies; comparisons to other methods are also shown. (C) LogFC amplitude, period and phase scatterplot comparison of genes identified as cycling in both the Zhang and Hughes datasets with an FDR < 0.05. (D) Heatmap of the cycling genes detected in each dataset with an FDR < 0.05 ordered by phase. Orange and gray represents gene expression above and below diurnal mean, respectively. (E) Rank correlation CDFs of genes identified by each method with an FDR < 0.05 and LogFC < 2 in one dataset compared to their pairwise expression in the other datasets. The null distribution represents the rank correlation of a random sampling of all 12 868 genes with replacement in the pairwise comparisons of all datasets. All gene correlation represents the rank correlation on a per gene basis for all 12 868 genes in the pairwise comparisons of all datasets.

for 48-h (Hogenesch 2A and Hogenesch 2B) to be consistent with the sampling schemes of the Hughes (Hughes *et al.*, 2012) and Zhang (Zhang *et al.*, 2014) studies. Following application of TimeCycle, genes with FDR < 0.05 were classified as cycling; the overlap in the cycling genes for the various datasets is given (Fig. 5A: Top). Comparing the distribution of the FDR-corrected P -values for the known circadian genes relative to all genes shows a statistically significant shift across all four datasets, with circadian genes more likely to have significant P -values (Fig. 5A: Bottom), confirming the intuition that known circadian genes should show enriched cycling. Furthermore, the known circadian genes *Per1*, *Cry1*, *Npas2* and *Clock* were detected across multiple studies.

To investigate whether the reproducibility was robust to the (somewhat arbitrary) choice of significance threshold, we calculated the percent of genes concordantly classified as cycling or non-cycling for each pair of studies as a function of the FDR significance threshold (Fig. 5B). The concordance is generally higher for TimeCycle relative to other methods at conservative significance thresholds ≤ 0.05 . This implies that genes identified as cycling by TimeCycle are likely to be reproducible.

The Hughes and Zhang studies differed in the sampling start phase by 6-h, presenting an opportunity to examine whether genes detected as cycling in both had similar estimated periods and amplitudes, but differing phases, as would be expected if the cycling detection is accurate (Ness-Cohn *et al.*, 2021). We find that genes detected as cycling in both the Zhang and Hughes datasets with an FDR < 0.05 also had highly reproducible amplitudes across the various datasets (Pearson correlation $r = 0.9$, Fig. 5C). We also find that genes identified as cycling clustered around a period of 24 h, indicating that these are indeed circadian oscillators. Finally, plotting the computed phase of genes detected reflects the 6-h phase shift (modulo 24 h) as expected (Fig. 5C). Heatmaps of cycling genes detected across datasets are shown (Fig. 5D).

We then examined more broadly whether genes that are detected as cycling across different studies exhibit similar dynamics, and compared TimeCycle's results to other methods. A method that has a high rate of false-positives may indeed identify many genes as cycling in common across studies, but by chance rather than detecting a meaningful signal. Hence, we assess whether the genes that are identified as cycling in multiple studies of the same tissue also have reproducible dynamics. For each study and each gene, we computed

the rank correlation of its time-series in the reference study with the same gene's time-series in the other studies. We expect that genes under true circadian control should exhibit more correlated dynamics amongst the studies than the set of all genes. We thus computed empirical CDFs of the correlations for genes identified as cycling (FDR < 0.05) by each method (Since SW1PerS does not produce a P -value to classify genes as cycling/non-cycling, it was omitted from this comparison.) (Fig. 5E). In this figure, the null distribution represents the rank correlation of a random sampling of all 12 868 genes with replacement in the pairwise comparisons of all datasets, and the all-gene correlation represents the rank correlations for all 12 868 genes in the pairwise comparisons of all datasets. TimeCycle, JTK_CYCLE and GeneCycle show a statistically significant shift in the distribution of rank correlations in comparison to RAIN, the all-gene correlation and the null distribution in all datasets (all $P \leq 2.2 \times 10^{-16}$, Kolmogorov-Smirnov test). We conclude from this that TimeCycle reliably detects cycling dynamics across multiple studies of the same tissue.

4 Discussion

We have presented TimeCycle, a new method that leverages results from dynamical systems theory and topology to detect patterns of cyclic expression in time-series experiments. TimeCycle reconstructs the state space for the dynamical system governing each gene using time-delay embedding, and quantifies how cyclic the embedding is using persistence homology. Statistical significance is assessed by comparing the persistence scores to those that obtain from a resampled null model. TimeCycle accurately detects rhythmic transcripts in both synthetic and real biological data, and is robust to missingness, noise and non-cyclic components in the dynamics.

A few methodological innovations distinguish TimeCycle from other cycling-detection algorithms. In contrast to methods that compare gene expression profiles to templates of expected cycling patterns, TimeCycle reconstructs the underlying dynamical system directly from the observed data. This enables TimeCycle to articulate more complex dynamics than can be easily considered using template-based approaches. Additionally, the method to construct the null distribution is both computationally efficient and biologically representative. SW1PerS (a prior method that also used

Table 1. Experimental design biological data

	Hogenesch (Hughes 2009)	Hughes (Hughes 2012)	Zhang (Zhang 2014)
Genotype	Wild-type C57BL/6J	Wild-type C57BL/6J	Wild-type C57BL/6J
Age	6 weeks	Not Specified	6 weeks
Entertainment condition	12h : 12hLD → DD	12h : 12hLD → DD	12h : 12hLD → DD
Entrainment length	1 Week	1 Week	1 Week
Microarray platform	Affy Mouse 430 2.0	Affy MoEx 1.0 ST	Affy MoGe 1.0 ST
Sampling scheme	Every 1-h for 48-h	Every 2-h for 48-h	Every 2-h for 48-h
ZT start time	ZT 18	ZT 0	ZT 18
Samples per time point	3–5 Male	2 Male and 2 Female	3 Male
GEO accession number	GSE11923	GSE30411	GSE54650

Note: Three mouse liver time-series expression sets (Hughes 2009; Hughes 2012; Zhang 2014) were analyzed from the Gene Expression Omnibus database (GEO). In each experiment, Wild-type C57BL/6J mice were entrained to a 12-h light, 12-h dark environment for a week before being released into constant darkness, a standard practice in the field for assessing circadian function. The Hogenesch study sampled mice every 1-h for 48-h, while the Hughes and Zhang studies sampled every 2-h for 48-h. Other variations in the datasets include microarray platform, ZT time of sampling and number of mice sampled per time point. Adapted from Ness-Cohn *et al.* (2020).

persistence homology rather than a template) did not implement hypothesis testing due to the computational cost, while template-based methods that implement hypothesis testing do so by resampling the time-series in a manner that can generate biologically implausible null models. TimeCycle is thus an improvement in both regards.

From a practical standpoint, we identified strengths and weaknesses of TimeCycle's ability to detect cycling transcripts under varying conditions. We find that TimeCycle is better able to detect sharply peaked waveforms and waveforms where the period appears variable, while JTK_Cycle, RAIN and GeneCycle are more robust with respect to linear trends. This is in keeping with prior work (Ness-Cohn *et al.*, 2020) demonstrating that no cycling detection method is consistently 'best' for all genes. Instead, it is incumbent upon the researcher to consider the patterns of gene expression that are of the greatest interest and choose a method accordingly.

The results also highlight the importance of constructing biologically representative null models. By resampling the finite differences from the gene expression time-series to construct a null distribution of persistence scores, TimeCycle tests whether an observed gene has a stronger cycling behavior than expected by chance, conditioned upon the speed at which the expression is capable of changing. This improves upon SW1PerS, which does not compute a *P*-value, and also upon methods such as RAIN that randomize the time-series itself. We note that this method for constructing the null distributions could also be adapted for other methods, and emphasize the need for method developers to consider biological constraints when devising null models.

A practical consequence of TimeCycle's methodological features is that the genes detected as cycling by TimeCycle are highly reproducible. Applied to three independent studies of mouse liver gene expression, TimeCycle consistently identified genes as cycling in multiple studies, and those genes were shown to exhibit reproducible dynamics.

Finally, we note that the experimental sampling design remains a crucial factor for the reliability of any cycling detection method. As with other methods (Hughes *et al.*, 2017; Ness-Cohn *et al.*, 2020), TimeCycle performs best when applied to time-series spanning 48-h sampled every 2-h, and is considerably less accurate for shorter and sparser sampling. These findings underscore the important role that experimental design—not only the method choice—plays in the analysis of circadian data.

Funding

Research reported in this publication was supported by the NSF-Simons Center for Quantitative Biology at Northwestern University, an NSF-Simons MathBioSys Research Center. This work was supported by a grant from the Simons Foundation/SFARI [597491-RWC] and the National Science Foundation [1764421]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science

Foundation and Simons Foundation. This research was supported in part through the computational resources and staff contributions provided for the Quest high performance computing facility at Northwestern University which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.

Conflict of Interest: none declared.

Data availability

Data and source code to reproduce the analysis and figures in this article are available at: <https://github.com/nesscoder/TimeCycle-data>.

Acknowledgements

E.N.C. and R.B. devised the TimeCycle method and designed the study; E.N.C. developed the TimeCycle package, analyzed the data and produced the tutorials/documentation; E.N.C. and R.B. wrote the article.

References

- Ahdesmäki, M. *et al.* (2007) Robust regression for periodicity detection in non-uniformly sampled time-course gene expression data. *BMC Bioinformatics*, **8**, 233.
- Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.
- Belkin, M. and Niyogi, P. (2003) Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.*, **15**, 1373–1396.
- Braun, R. *et al.* (2018) Universal method for robust detection of circadian state from gene expression. *Proc. Natl. Acad. Sci. USA*, **115**, E9247–E9256.
- Chang, A.M. *et al.* (2009) Sleep timing and circadian phase in delayed sleep phase syndrome. *J. Biol. Rhythms*, **24**, 313–321.
- de Lichtenberg, U. *et al.* (2005) Comparison of computational methods for the identification of cell cycle-regulated genes. *Bioinformatics*, **21**, 1164–1171.
- Deckard, A. *et al.* (2013) Design and analysis of large-scale biological rhythm studies: a comparison of algorithms for detecting periodic signals in biological data. *Bioinformatics*, **29**, 3174–3180.
- Deshmukh, V. *et al.* (2020) Using curvature to select the time lag for delay reconstruction. *Chaos*, **30**, 063143.
- Edelsbrunner, H. *et al.* (2002) Topological persistence and simplification. *Discrete Comput. Geometry*, **28**, 511–533.
- Fraser, A.M. and Swinney, H.L. (1986) Independent coordinates for strange attractors from mutual information. *Phys. Rev. A*, **33**, 1134–1140.
- Garland, J. *et al.* (2016) Leveraging information storage to select forecast-optimal parameters for delay-coordinate reconstructions. *Phys. Rev. E*, **93**, 022221.
- Gierliński, M. *et al.* (2015) Statistical models for RNA-seq data derived from a two-condition 48-replicate experiment. *Bioinformatics*, **31**, 3625–3630.
- Hughes, M.E. *et al.* (2009) Harmonics of circadian gene transcription in mammals. *PLoS Genet.*, **5**, e1000442.

- Hughes, M.E. *et al.* (2010) JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J. Biol. Rhythms*, **25**, 372–380.
- Hughes, M.E. *et al.* (2012) Brain-specific rescue of Clock reveals system-driven transcriptional rhythms in peripheral tissue. *PLoS Genet.*, **8**, e1002835.
- Hughes, M.E. *et al.* (2017) Guidelines for genome-scale analysis of biological rhythms. *J. Biol. Rhythms*, **32**, 380–393.
- Hutchison, A.L. and Dinner, A.R. (2017) Correcting for dependent P-values in Rhythm Detection. *bioRxiv*, 118547. 10.1101/118547
- Hutchison, A.L. *et al.* (2015) Improved statistical methods enable greater sensitivity in rhythm detection for genome-wide data. *PLoS Comput. Biol.*, **11**, e1004094.
- Hutchison, A.L. *et al.* (2018) Bootstrapping and empirical bayes methods improve rhythm detection in sparsely sampled data. *J. Biol. Rhythms*, **33**, 339–349.
- Kathale, N.D., and Liu, A.C. (2014) Prevalence of cycling genes and drug targets calls for prospective chronotherapeutics. Proceedings of the National Academy of Sciences of the United States of America, **111**, 15869–15870. 10.1073/pnas.1418570111 25368193
- Laloum, D. and Robinson-Rechavi, M. (2020) Methods detecting rhythmic gene expression are biologically relevant only for strong signal. *PLoS Comput. Biol.*, **16**, e1007666.
- Levi, F. and Schibler, U. (2007) Circadian rhythms: mechanisms and therapeutic implications. *Annu. Rev. Pharmacol. Toxicol.*, **47**, 593–628.
- Levine, D.C. *et al.* (2020) NAD⁺ controls circadian reprogramming through PER2 nuclear translocation to counter aging. *Mol. Cell*, **78**, 835–849.e7.
- Levine, J.D. *et al.* (2002) Signal analysis of behavioral and molecular cycles. *BMC Neuroscience*, **3**, 1.
- Li, S. *et al.* (2017) CGDB: a database of circadian genes in eukaryotes. *Nucleic Acids Res.*, **45**, D397–D403.
- Love, M.I. *et al.* (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
- Mei, W. *et al.* (2021) Genome-wide circadian rhythm detection methods: systematic evaluations and practical guidelines. *Brief. Bioinf.*, **22**, 1–13.
- Michael, T.P. *et al.* (2008) Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS Genet.*, **4**, e14.
- Moritz, S. and Bartz-Beielstein, T. (2017) imputeTS: time series missing value imputation in R. *R. J.*, **9**, 207–218.
- Ness-Cohn, E. *et al.* (2020) TimeTrial: an interactive application for optimizing the design and analysis of transcriptomic time-series data in circadian biology research. *J. Biol. Rhythms*, **35**, 439–451.
- Ness-Cohn, E. *et al.* (2021) Comment on “Circadian rhythms in the absence of the clock gene Bmal1”. *Science*, **372**, eabe9230.
- Patke, A. *et al.* (2017) Mutation of the human circadian clock gene CRY1 in familial delayed sleep phase disorder. *Cell*, **169**, 203–215.e13.
- Perea, J.A. *et al.* (2015) SW1PerS: sliding windows and 1-persistence scoring; discovering periodicity in gene expression time series data. *BMC Bioinformatics*, **16**, 257.
- Puttonen, S. *et al.* (2010) Shift work and cardiovascular disease – pathways from circadian stress to morbidity.
- Robin, X. *et al.* (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, **12**, 77.
- Roenneberg, T. *et al.* (2007) Epidemiology of the human circadian clock. *Sleep Med. Rev.*, **11**, 429–438.
- Salkovic, E. *et al.* (2020) OutPyR: Bayesian inference for RNA-Seq outlier detection. *J. Comput. Sci.*, **47**, 101245.
- Straume, M. (2004) DNA microarray time series analysis: automated statistical assessment of circadian rhythms in gene expression patterning. *Methods Enzymol.*, **383**, 149–166.
- Takens, F. (1981) Detecting strange attractors in turbulence. In: Rand D., Young L.S. (eds) *Dynamical Systems and Turbulence*, Warwick 1980. Lecture Notes in Mathematics, vol 898. Springer, Berlin, Heidelberg.
- Thaben, P.F. and Westermark, P.O. (2014) Detecting rhythms in time series with RAIN. *J. Biol. Rhythms*, **29**, 391–400.
- Thaben, P.F. and Westermark, P.O. (2016) Differential rhythmicity: detecting altered rhythmicity in biological data. *Bioinformatics*, **32**, 2800–2808.
- Videnovic, A. *et al.* (2014) Circadian melatonin rhythm and excessive daytime sleepiness in Parkinson disease. *JAMA Neurol.*, **71**, 463–469.
- Wichert, S. *et al.* (2004) Identifying periodically expressed transcripts in microarray time series data. *Bioinformatics*, **20**, 5–20.
- Wu, G. *et al.* (2014) Evaluation of five methods for genome-wide circadian gene identification. *J. Biol. Rhythms*, **29**, 231–242.
- Wu, G. *et al.* (2016) MetaCycle: an integrated R package to evaluate periodicity in large scale data. *Bioinformatics*, **32**, 3351–3353.
- Yang, R. and Su, Z. (2010) Analyzing circadian expression data by harmonic regression based on autoregressive spectral estimation. *Bioinformatics*, **26**, i168–i174.
- Youden, W.J. (1950) Index for rating diagnostic tests. *Cancer*, **3**, 32–35.
- Zhang, R. *et al.* (2014) A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc. Natl. Acad. Sci. USA*, **111**, 16219–16224.