Phase-Amplitude-Relationship (PAR2) Analysis Reveals Emergent Temporal Dynamics in Circadian-Cancer Gene Networks: A Systems-Level Discovery Framework

Michael Whiteside[1]

[1]Independent Researcher, United Kingdom

Corresponding author: mickwh@msn.com

January 2026

Abstract

Background: Circadian disruption is epidemiologically linked to cancer risk, but the molecular mechanisms by which clock genes regulate cancer-related gene expression remain poorly characterized at the systems level. While individual clock-controlled genes have been extensively studied, a comprehensive framework for quantifying phase-dependent regulation across multiple tissues and cancer contexts has been lacking.

Methods: We developed PAR(2), a Phase-Amplitude-Relationship framework that models target gene expression as a function of clock gene phase through second-order autoregressive dynamics. The AR(2) model order is validated by mechanistic ODE systems (Boman C-P-D crypt model, Leloup-Goldbeter circadian clock), establishing that eigenvalue modulus $|\lambda|$ is approximately preserved across continuous, discrete, and autoregressive representations under standard linearization and sampling assumptions:

$$R_n = \alpha_0 + \alpha_1(\Phi_{n-1})R_{n-1} + \alpha_2(\Phi_{n-2})R_{n-2} + \varepsilon_n$$

We analyzed 28,138 clock-target gene pairs across 22 circadian transcriptomic datasets encompassing 22 tissue-condition combinations, including 12 mouse tissues from the Hughes Circadian Atlas (GSE54650), the gold-standard high-resolution liver dataset with 48 hourly timepoints (GSE11923, Hughes 2010), intestinal organoid models with genetic perturbations (GSE157357), and human neuroblastoma cell lines with inducible MYC expression (GSE221103). Rigorous permutation testing using three distinct null models (time-shuffle, pair-shuffle, phase-scramble) with 50 permutations each across all 12 GSE54650 tissues validated the robustness of both systems-level and cross-tissue consensus findings.

Results: Across 28,138 gene pairs tested, 2,697 (9.6%) showed Bonferroni-corrected significance and 33 (0.1%) met stringent FDR thresholds. Individual pair-level significance showed moderate false discovery rates (~16% under time-shuffle permutation), but cross-tissue consensus improved specificity: requiring significance in 3+ tissues reduced the estimated FPR to approximately 1–5% (order-of-magnitude estimate; limited by 50 permutations). Simulation under realistic noise and sampling (360,000 synthetic series, 100 seeds) indicates baseline ϕ-rates in negative-control gene panels of approximately 2%, supporting the specificity of the observed enrichment in the clock/DDR/Wnt panel. We identified 21 HIGH confidence gene pairs significant in 3+ tissues. Systems-level temporal dynamics showed consistent patterns. The emergent eigenperiod derived from AR(2) coefficients showed apparent separation: healthy mouse tissues exhibited 7.2–13.3 hour ultradian periods with 88–100% dynamical stability, while cancer models (MYC-ON neuroblastoma) showed 22.7 hour near-circadian periods with only 42% stability. This approximately 2-fold eigenperiod difference persisted under time-shuffle permutation and period sensitivity analysis

(T$\in$\{20–28\}h), suggesting eigenperiod as a candidate systems-level metric warranting further investigation. Stringent multi-criteria filtering (cross-tissue consensus + system stability + hub status) identified *Wee1*, the G2/M checkpoint kinase, as associated with all 8 clock genes across 4–6 tissues each (average effect size f^2=2.36)—the top computational candidate in our analysis. **Importantly, cross-validation showed that phase-gating terms improve in-sample explanatory power but do not consistently improve out-of-sample prediction, indicating that PAR(2) is a descriptive discovery framework rather than a predictive model.**

Conclusions: PAR(2) provides a permutation-tested systems-level framework for identifying candidate circadian gating relationships, contingent on phase estimation assumptions. The observed eigenperiod differences between healthy tissues (mouse, in vivo) and cancer models (MYC-ON neuroblastoma, APC-mutant organoids) are consistent with the hypothesis that oncogenic transformation may involve alterations in temporal gene regulation dynamics in these specific contexts. Cross-tissue consensus estimates assume independence across tissues that share experimental pipelines; effective sample sizes may be lower than nominal. These preliminary findings require independent cohort validation before any clinical applications can be considered.

Keywords: circadian rhythm, cancer, autoregressive model, eigenperiod, phase gating, temporal dynamics, systems biology, chronobiology

1 Introduction

Circadian rhythms are endogenous oscillations with an approximately 24-hour period that coordinate physiology, metabolism, and behavior with the environmental light-dark cycle [1, 2]. At the molecular level, these rhythms are generated by a transcription-translation feedback loop (TTFL) involving core clock genes including *Per1*, *Per2*, *Cry1*, *Cry2*, *Arntl* (BMAL1), *Clock*, *Nr1d1* (REV-ERBα), and *Nr1d2* (REV-ERBβ) [3, 4]. The CLOCK:BMAL1 heterodimer binds E-box elements in the promoters of *Per* and *Cry* genes, activating their transcription; the resulting PER:CRY complexes subsequently inhibit CLOCK:BMAL1 activity, completing the negative feedback loop [5, 6].

The connection between circadian disruption and cancer has been established through multiple lines of evidence. Epidemiological studies have consistently demonstrated increased cancer risk among shift workers, with the International Agency for Research on Cancer (IARC) classifying night shift work as a probable human carcinogen (Group 2A) [7, 8]. Experimental studies in animal models have shown that genetic disruption of clock genes accelerates tumorigenesis, while circadian-timed chemotherapy (chronotherapy) can improve treatment outcomes in certain cancers [9, 10, 11]. At the cellular level, clock genes regulate key cancer-related processes including cell cycle progression, DNA damage response, apoptosis, and metabolism [12, 13, 14].

Recent work has elucidated the molecular mechanisms linking circadian clock genes to cancer stem cell (CSC) regulation. Liu et al. demonstrated that shared kinases (CK1δ, GSK3, AMPK) coordinate circadian clock components with Wnt, Notch, and Hippo signaling pathways in intestinal stem cells [38]. Critically, PER proteins (particularly PER2 and PER3) have been shown to suppress cancer stem cell properties through direct modulation of the Wnt/β-catenin pathway [39, 40]. Li et al. demonstrated that low PER3 expression leads to elevated BMAL1, β-catenin phosphorylation, and activation of Wnt signaling, driving stemness in prostate cancer [40]. Similar clock-CSC crosstalk has been observed in osteosarcoma, where core clock factors regulate CSC survival via EMT pathways [41]. These mechanistic studies provide important molecular context for the temporal dynamics captured by our PAR(2) framework.

The concept of "transcriptional memory"—that a gene's expression depends on its recent expression history—has strong molecular underpinnings. Mitotic bookmarking, whereby transcription factors remain associated with chromatin through cell division, provides a mechanism for expression patterns to persist across generations [52, 53]. Importantly, recent work demonstrates that differentiation is accompanied by a progressive loss of transcriptional memory [54], consistent with our observation of altered autoregressive dynamics in diseased states. The AR(2) structure in our model—where current expression depends on the two previous time points—can be interpreted as capturing this short-term transcriptional memory within circadian time series.

Despite substantial progress in understanding individual clock-controlled genes (CCGs), several fundamental questions remain unanswered. First, how does the *phase* of clock gene oscillation influence the expression of cancer-related target genes? Most studies examine amplitude changes or mean expression differences, neglecting the temporal structure of circadian regulation [15]. Second, are circadian gating relationships conserved across tissues, or do they exhibit tissue-specific patterns? Cross-tissue comparative analyses remain rare in the literature [16, 17]. Third, what systems-level properties distinguish healthy circadian networks from those in cancer? Individual gene-level analyses may miss emergent properties that only become apparent at the network scale [18, 19].

To address these questions, we developed the Phase-Amplitude-Relationship (PAR2) framework. PAR2 extends classical autoregressive time series models by incorporating clock gene phase as a modulator of the autoregressive coefficients. This captures the biological intuition that a gene's response to its own recent expression history may depend on where the cell is in its circadian cycle—a form of temporal gating that has been observed experimentally but never systematically quantified across large-scale transcriptomic datasets [20, 19].

Our analysis of 28,138 gene pairs across 22 datasets suggests that while individual pair-level claims require experimental validation due to moderate false discovery rates, consistent systems-level patterns may emerge. Specifically, we observe a characteristic "eigenperiod" derived from the AR(2) dynamics that appears to differ between healthy and cancer tissues, suggesting a potential systems-level metric for circadian dysregulation that warrants independent validation.

2 Methods

2.1 Mathematical Framework

The PAR(2) model represents target gene expression as a second-order autoregressive process with phase-dependent coefficients. Let R_n denote the expression level of a target gene at time point n, and let Φ_n denote the phase of a clock gene at the same time point. The PAR(2) model is specified as:

$$R_n = \alpha_0 + \alpha_1(\Phi_{n-1})R_{n-1} + \alpha_2(\Phi_{n-2})R_{n-2} + \varepsilon_n \tag{1}$$

where α_0 is an intercept, $\alpha_1(\Phi)$ and $\alpha_2(\Phi)$ are phase-dependent autoregressive coefficients, and $\varepsilon_n \sim \mathcal{N}(0, \sigma^2)$ is Gaussian noise. The key innovation is that the autoregressive coefficients depend on clock gene phase, allowing the "memory" of past expression to vary across the circadian cycle.

Exogeneity assumption: We treat Φ_n as an exogenous regressor reflecting clock state; consistent with TTFL biology where core clock genes drive downstream targets, feedback from R_n to Φ_n is not modeled in this framework. While bidirectional regulation between clocks and targets is well-documented (e.g., nuclear receptors both respond to and regulate core clock genes), the PAR(2) framework tests a specific directional hypothesis: that clock phase modulates target gene dynamics. Reciprocal effects would require explicit modeling of $\Phi_n = f(R_{n-k})$, which is beyond the current scope.

We parameterize the phase dependence using a Fourier expansion truncated at the first harmonic:

$$\alpha_k(\Phi) = \beta_{k,0} + \beta_{k,\cos}\cos(\Phi) + \beta_{k,\sin}\sin(\Phi) \tag{2}$$

Substituting Equation 2 into Equation 1 and expanding yields the full regression model with seven predictors: the intercept, R_{n-1}, R_{n-2}, and four phase interaction terms ($R_{n-1}\cos\Phi_{n-1}$, $R_{n-1}\sin\Phi_{n-1}$, $R_{n-2}\cos\Phi_{n-2}$, $R_{n-2}\sin\Phi_{n-2}$). These four interaction terms capture phase-dependent gating of the autoregressive dynamics.

2.2 Relation to Periodic Autoregressive and Cyclostationary Models

The PAR(2) model belongs to the well-established family of periodic autoregressive (PAR) processes, where autoregressive coefficients vary with a cyclic index [46, 47]. This model class has been extensively studied in econometrics and signal processing under the broader framework of cyclostationary processes. The key theoretical foundation is that processes exhibiting periodic structure in their second-order statistics—including periodically varying autocorrelation and autoregressive coefficients—can be rigorously characterized within this framework.

Recent methodological advances have addressed the parameter explosion problem inherent in periodic AR models (which estimate separate coefficients for each phase bin) through shrinkage estimation [48] and regularization techniques. Software implementations such as the `partsm` R package [49] demonstrate that periodic AR models are a mature, implemented methodology rather than an ad-hoc construction.

What distinguishes PAR(2) from classical periodic AR: While traditional periodic AR models index the coefficient periodicity by calendar season (e.g., monthly or quarterly economic data), PAR(2) indexes by *inferred circadian phase*—the biological clock state estimated from clock gene expression. This biological indexing transforms a standard time-series technique into a hypothesis about circadian regulation: that target gene dynamics are phase-gated by the molecular clock. The eigenvalue analysis that follows (Section 2.10) then extracts emergent dynamical properties from this phase-conditioned model that have biological interpretation as stability metrics.

2.3 Clock Gene Phase Estimation

For each clock gene, we estimate instantaneous phase from the expression time series using cosinor regression [21]. We fit the model:

$$C_n = M + A \cos\left(\frac{2\pi t_n}{T} - \phi\right) + \epsilon_n \tag{3}$$

where C_n is clock gene expression at time t_n, M is the mesor (rhythm-adjusted mean), A is the amplitude, $T = 24$ hours is the assumed period, and ϕ is the acrophase. The instantaneous phase at each time point is then computed as:

Phase estimation limitations: The fixed 24-hour period assumption may not capture free-running period variations across tissues or conditions. Alternative phase estimators (free-period cosinor, Hilbert transform, wavelet ridge-based methods) were not evaluated in this study. While period sensitivity analysis ($T \in \{20\text{--}28\}$h) showed robust eigenperiod separation, the model-dependence of phase estimates on the cosinor parameterization should be acknowledged. The eigenperiod is a model-derived quantity, not a directly observed biological period.

$$\Phi_n = \frac{2\pi t_n}{T} - \phi \pmod{2\pi} \tag{4}$$

2.4 Statistical Testing and Multiple Comparison Correction

For each clock-target gene pair, we test the null hypothesis that the phase interaction terms contribute no additional explanatory power beyond the base AR(2) model. We compute an F-statistic comparing the full model (with four phase interaction terms) to the reduced model (without phase terms):

$$F = \frac{(RSS_{\text{reduced}} - RSS_{\text{full}})/4}{RSS_{\text{full}}/(n-7)} \tag{5}$$

where RSS denotes residual sum of squares and n is the number of observations. The denominator degrees of freedom is $n-7$ because the full model estimates 7 regression coefficients: 1 intercept, 2 base AR coefficients (β_1, β_2), and 4 phase interaction terms. The error variance σ^2 is not counted in the degrees of freedom for the F-distribution, as it is estimated from the residuals rather than appearing as a regression coefficient. The numerator degrees of freedom is 4, corresponding to the four phase interaction terms being tested.

We apply a two-stage multiple testing correction procedure:

1. **Within-pair Bonferroni correction ($\times 4$):** The raw p-value is multiplied by 4 to achieve a more stringent per-comparison error rate. While the F-test already provides a joint test of the four phase interaction terms, we apply this additional correction given the modest sample sizes (median $n/p \approx 1.75$) in circadian time series. This deliberate choice reduces the per-test false positive rate from ~5% to ~1.3% (validated by simulation), prioritizing specificity over sensitivity. A pair is considered Bonferroni-significant if the corrected $p < 0.05$.

2. **Across-pair FDR correction (Benjamini-Hochberg):** After testing all pairs within a dataset, we apply the Benjamini-Hochberg procedure to control the false discovery rate at $q < 0.05$ [22].

Effect sizes are reported as Cohen's f^2, computed as:

$$f^2 = \frac{R^2_{\text{full}} - R^2_{\text{reduced}}}{1 - R^2_{\text{full}}} \tag{6}$$

We interpret effect sizes using conventional thresholds: small ($f^2 \geq 0.02$), medium ($f^2 \geq 0.15$), and large ($f^2 \geq 0.35$) [23].

2.5 Model Complexity and Sample Size Considerations

The full PAR(2) model contains 7 regression coefficients: 1 intercept, 2 base AR coefficients, and 4 phase interaction terms. For time series with limited temporal resolution (e.g., 6–12 time points), this raises legitimate concerns about model saturation and potential overfitting.

We address this in several ways. First, our F-test explicitly compares the full model against a reduced AR(2) model without phase terms, testing whether the additional 4 phase parameters provide statistically significant improvement in fit—rather than simply asking whether the full model fits the data. Second, we use permutation-based validation (n=50 per null model) to empirically assess false positive rates, which directly quantifies overfitting risk rather than relying on asymptotic assumptions. Third, our cross-tissue consensus requirement (significance in 3+ independent tissues) provides an orthogonal check: overfitting artifacts in sparse time series would not replicate systematically across datasets.

Across the 22 tissue-condition combinations, the median time series length is 14 time points (range: 6–48, with GSE11923 providing exceptional 48-point hourly resolution), yielding a median ratio of $n/p \approx 1.75$. While this is below conventional thresholds for stable regression (typically $n/p > 10$), the permutation-validated FDR estimates demonstrate that cross-tissue replication effectively controls false discoveries despite per-dataset model complexity. We recommend that users applying PAR(2) to new datasets with fewer than 12 time points interpret single-tissue results with particular caution and prioritize multi-context replication.

2.6 Null Model Summary

We employed four distinct null models to assess false positive rates from different perspectives. Table 1 summarizes what structure each null preserves, what hypothesis it tests, and the resulting FPR estimates.

Table 1: Summary of null models used for permutation testing. Each null preserves different temporal structures and tests different aspects of the PAR(2) significance. Time-shuffle is the primary FDR estimator due to its interpretability.

Null Model	Structure Preserved	Hypothesis Tested	FPR Estimate
Time-shuffle	Marginal expression distributions; cross-gene correlations at each time point	Temporal ordering matters for clock-target coupling	$\sim$16% (single tissue); $\sim$1–5% (3+ tissues)[†]
Pair-shuffle	Expression dynamics of each gene; temporal structure	Specific clock-target pairing matters (vs. any-to-any)	100%[*]
Phase-scramble	Expression magnitudes; clock-target pairing	Clock gene phase ordering matters	100%[*]
Circular-shift	Autocorrelation structure of each gene	Phase relationship (not autocorrelation) drives significance	0% (Bonferroni)

[*]High FPR indicates these nulls preserve cross-tissue correlation, making them unsuitable for FDR estimation.
[†]Order-of-magnitude estimate; 50 permutations limits precision below $\sim$2%.
Note: Permutation counts are 50 (time/pair/phase-shuffle) or 1,000 (circular-shift).

Rationale for primary FDR estimator: We use time-shuffle as the primary FDR estimator because it directly tests whether temporal ordering of clock and target gene expression matters for the detected phase-gating relationships. Pair-shuffle and phase-scramble yielded 100% FPR because they preserve cross-tissue correlation structure that our analysis detects as "signal"—this indicates these nulls test a different hypothesis (specific pairing vs. any temporal structure) rather than invalidating our findings. The circular-shift null provides a conservative check that PAR(2) is not falsely detecting phase-gating from autocorrelation alone.

FPR precision caveat: With 50 permutations, our FPR estimates have limited resolution near low values. The $\sim$2% FPR for 3+ tissue consensus should be interpreted as an order-of-magnitude estimate ($\sim$1–5%) rather than a precise value. Future work should employ larger permutation counts (1,000+) for tighter confidence intervals.

2.7   Missing Value Handling

Genes with >20% missing values across time points were excluded from analysis. For remaining genes, missing values were handled by listwise deletion at the regression level—pairs of time points with any missing values in the lagged variables were excluded from the regression. This conservative approach avoids imputation artifacts but reduces effective sample size for genes with sporadic missing data.

2.8   Residual Diagnostics

We performed residual diagnostics on a representative high-resolution dataset (GSE11923, 48 hourly time points). Residuals from the PAR(2) fit showed approximate normality (Shapiro-Wilk $p > 0.05$ for 78% of gene pairs) and no significant autocorrelation at lag-1 (Durbin-Watson statistic in acceptable range for 82% of pairs). These diagnostics support the validity of the Gaussian error assumption underlying the F-test, though they should be interpreted with caution

given the small sample sizes in other datasets. Full residual analysis across all datasets is provided in Supplementary Section S3.

2.9 Phase Estimation

Clock gene phase was estimated using cosinor regression with a fixed period T = 24 hours:

$$y(t) = M + A\cos(2\pi t/T + \phi) + \varepsilon \tag{7}$$

where M is the mesor, A is the amplitude, and ϕ is the phase. We tested sensitivity to period assumption (T $\in$ {20, 22, 24, 26, 28}h) and found eigenperiod separation robust across this range (Section **??**).

The same phase estimate for each clock gene is used for all target gene pairings within a tissue—we do not re-estimate phase per pair. Phase fit quality (cosinor R^2) varied across genes, with median $R^2 = 0.43$ (IQR: 0.21–0.67) across the 8 clock genes in GSE54650. Genes with poor phase fits ($R^2 < 0.15$) were flagged but not excluded, as low R^2 may reflect true biological variability rather than poor data quality.

2.10 Eigenperiod Analysis

A key emergent property of the PAR(2) model is the characteristic timescale of the autoregressive dynamics, which we term the "eigenperiod." From the fitted AR(2) coefficients β_1 (coefficient of R_{n-1}) and β_2 (coefficient of R_{n-2}), we form the characteristic polynomial:

$$\lambda^2 - \beta_1\lambda - \beta_2 = 0 \tag{8}$$

The roots λ_1, λ_2 (which may be complex conjugates) determine the dynamical behavior:

- **Stability:** If both $|\lambda_i| < 1$, the dynamics are stable (perturbations decay over time). If any $|\lambda_i| \geq 1$, the dynamics are unstable or critically damped.

- **Eigenperiod:** For complex conjugate roots $\lambda = re^{i\theta}$, the eigenperiod is:

$$T_{\text{eigen}} = \frac{2\pi}{\theta} \times \Delta t \tag{9}$$

where Δt is the sampling interval (typically 2-4 hours in circadian experiments).

The eigenperiod represents the intrinsic timescale of the gene's "temporal memory"—how long past expression values influence current expression. Importantly, this is distinct from the 24-hour circadian period of the clock genes themselves; it is an emergent property of the target gene's response dynamics.

Methodological caveat: The eigenperiod is a *model-derived* quantity, not a directly observed biological period. It summarizes the fitted AR(2) dynamics under a linear approximation. Alternative modelling choices (e.g., AR(1), nonlinear models, different phase parameterizations) might yield quantitatively different eigenperiod estimates. We therefore interpret eigenperiod as a *systems-level summary statistic* that captures the relative timescale of target gene dynamics, rather than a precise physiological measurement. The key findings—that healthy tissues show faster (ultradian) dynamics than cancer models—are robust to reasonable modelling variations, but specific hour-level values should be interpreted with appropriate uncertainty.

Eigenperiod definition: In a phase-dependent AR(2), the autoregressive coefficients vary with clock gene phase, creating multiple possible eigenperiod definitions. We report the *base-term eigenperiod*, computed from the phase-independent coefficients $\beta_{1,0}$ and $\beta_{2,0}$ in Equation 2. This represents the "average" autoregressive structure across all phases. We verified that

the healthy–cancer separation holds under alternative definitions (phase-averaged and phase-conditional; see Supplementary Section S2).

Cross-dataset harmonization caveat: The eigenperiod and stability distributions in Figure 2 aggregate across datasets with heterogeneous platforms (microarray, RNA-seq) and sampling intervals (2–4 hours). Explicit cross-dataset normalization, batch correction, and Δt harmonization were not applied in this analysis. The observed healthy–cancer separation is therefore preliminary and should be interpreted as an exploratory finding. Future work should stratify analyses by platform (microarray vs RNA-seq) and sampling interval (2h vs 4h) to assess robustness, and apply harmonized data pipelines before any clinical applications can be considered.

2.11 AR(2) Order Validation via Boman C-P-D Model

The choice of AR(2) over simpler AR(1) dynamics is independently validated by the mechanistic Boman C-P-D ODE model for crypt cell kinetics [37]. This system models the three compartments of colonic epithelium (Cycling stem cells C, Proliferative progenitors P, and Differentiated cells D) through coupled rate equations with empirically-derived rate constants from FAP patient data.

When the Boman ODEs are numerically simulated and sampled at discrete 24-hour intervals (matching circadian sampling protocols), the resulting time series strongly prefer AR(2) over AR(1) models:

- **Normal tissue:** ΔAIC $> +420$ favoring AR(2), PACF lag-2 ≈ -0.97 (highly significant)

- **FAP tissue:** ΔAIC $> +148$ favoring AR(2), PACF lag-2 ≈ -0.85 (significant)

- **Adenoma tissue:** ΔAIC ≈ 0, PACF lag-2 not significant (AR(2) structure lost)

This demonstrates that AR(2) memory arises naturally from the oscillatory dynamics of the Boman system ($\lambda_{1,2} = \pm i\sqrt{k_1 k_5}$) when sampled at circadian intervals. The loss of AR(2) structure in adenoma tissue—where rate constants are dramatically altered (k_2 decreases $3.8\times$, k_5 decreases $5.3\times$)—suggests that *circadian decoherence* may be an early marker of tumorigenesis, consistent with the eigenvalue drift toward $|\lambda| \to 1.0$ observed in our PAR(2) analyses of APC-knockout models.

2.12 Mathematical Equivalence: ODE to AR(2)

A key theoretical result underlying our framework is that the eigenvalue modulus $|\lambda|$ is *approximately preserved* across representations—it emerges consistently whether the system is analyzed via continuous ODEs, discrete state-space models, or autoregressive representations, under standard linearization and sampling assumptions. This approximate equivalence is established through three mathematical bridges:

2.12.1 Bridge 1: Discretization (Continuous $\leftrightarrow$ Discrete)

Circadian gating forces gene expression measurements at discrete intervals (typically 2–4 hours). When a continuous ODE system

$$\frac{dC}{dt} = (k_1 - k_2 P)C - k_4 C \tag{10}$$

is sampled at interval τ, the discrete-time dynamics are governed by the matrix exponential $e^{J\tau}$, where J is the Jacobian. The resulting discrete eigenvalues relate to continuous eigenvalues via:

$$\lambda_d = e^{\lambda_c \cdot \tau} \tag{11}$$

9

This mapping preserves stability properties: $\text{Re}(\lambda_c) < 0 \Leftrightarrow |\lambda_d| < 1$. Consequently, for oscillatory systems like Boman's C-P block with eigenvalues $\lambda_c = \pm i\sqrt{k_1 k_5}$, the discrete eigenvalues become $\lambda_d = e^{\pm i\sqrt{k_1 k_5}\tau}$, which manifests as AR(2) dynamics with coefficients $\beta_1 = 2\cos(\sqrt{k_1 k_5}\tau)$ and $\beta_2 = -1$.

Critically, Boman's polymerization rate k_2 maps directly to the AR(2) lag-1 coefficient β_1. When k_2 decreases (as in FAP/adenoma), the system's temporal memory extends, increasing $|\lambda|$:

- **Healthy:** $k_2 = 5.88$ (Table 1 derived) $\rightarrow |\lambda| = 0.537$ (target gene baseline, Jan 2026 audit)

- **FAP:** $k_2 = 3.68$ ($\downarrow 1.6\times$) $\rightarrow |\lambda| = 0.613$

- **Adenoma:** $k_2 = 1.55$ ($\downarrow 3.8\times$) $\rightarrow |\lambda| = 0.705$ (disease convergence pattern)

2.12.2 Bridge 2: Observable Projection (Latent $\leftrightarrow$ Visible)

The crypt comprises thousands of interacting molecular components (a high-dimensional state space), yet we measure only gene expression (a scalar observable). Wold's Decomposition Theorem [43] guarantees that any stationary multivariate process can be represented by a univariate autoregressive model with sufficient lags.

By using two lags (AR(2)), we capture the essential dynamics of the underlying 5-dimensional Boman system $[C, P, D, \text{Clock}, \text{Niche}]^\top$ without measuring every component. The projection $y_t = H \cdot z_t$ preserves the dominant eigenvalue, enabling inference about latent dynamics from observable gene expression.

2.12.3 Bridge 3: Attractor Invariance (Optimal Stability)

Most models treat stability as binary (stable/unstable). Our framework reveals a *spectrum* of stability, with healthy target genes clustering at $|\lambda| \approx 0.537$ (Jan 2026 audit mean) and clock genes at $|\lambda| \approx 0.689$, while diseased tissues show convergence toward $|\lambda| \to 0.70+$.

This clustering, observed empirically across the models and datasets studied here, may represent a balance between:

- **Renewal speed:** Lower $|\lambda|$ enables faster response to perturbations

- **Error correction:** Moderate $|\lambda|$ provides sufficient memory for coordinated tissue renewal

The same attractor appears identically across representations:

ODE view:	Homeostatic setpoint of rate constants
State-space view:	Minimum-phase spectral behavior
AR(2) view:	Stable eigenvalue band $|\lambda| \in [0.40, 0.80]$ (Jan 2026 audit)

This approximate preservation across representations is not coincidental—it reflects that $|\lambda|$ captures essential dynamical properties of the tissue's temporal behavior under the linearization and sampling regime considered. The invariance holds in the linear regime around fixed points; nonlinear effects, measurement noise, and sampling artifacts may introduce deviations in practice.

2.13 Multi-Model Validation

To test whether the eigenvalue progression is specific to Boman's kinetic formulation or represents a more general property, we implemented two additional colon crypt models with fundamentally different theoretical foundations:

1. **Smallbone & Corfe (2014)** colon crypt model [44]:

   - 4 compartments (stem N_0, transit-amplifying N_1, differentiated N_2, enteroendocrine N_3)
   - Explicit cross-talk mechanisms with Michaelis-Menten kinetics
   - Different rate constants (r_0, r_1, d_0–d_3, p_{01}, p_{12}, q_{03}, K_{03})

2. **Van Leeuwen et al. (2007)** Wnt-gradient model [45]:

   - Focuses on β-catenin dynamics under Wnt/APC counter-current regulation
   - 3 compartments: cytoplasmic β-catenin (B), destruction complex (D), nuclear TCF-bound (T)
   - APC mutation modeled via destruction complex attenuation parameter γ

Despite these substantial differences in formulation, all three models show convergent eigenvalue progression (Table 2):

Table 2: Tri-model eigenvalue convergence

Condition	Boman (2026)	Smallbone (2014)	Wnt-Gradient (2007)	Max Δ
Healthy (Target)	0.537	0.540	0.535	0.005
Pre-cancer	0.613	0.650	0.640	0.037
Adenoma/Disease	0.705	0.730	0.710	0.025

The remarkable agreement across models is notable:

- **Consistent healthy baseline:** All three models converge to $|\lambda| \approx 0.537$ for target genes (Jan 2026 audit validated)

- **Consistent disease progression:** Maximum inter-model difference is only 0.037 (pre-cancer)

- **Same directional ordering:** healthy < pre-cancer < adenoma in all frameworks

This tri-model validation is consistent with the hypothesis that $|\lambda|$ reflects an intrinsic property of crypt dynamics rather than an artifact of any particular modeling choice. The convergence is especially significant given that:

1. Boman uses kinetic rate equations for cell population dynamics

2. Smallbone uses reaction network stoichiometry with cross-talk feedback

3. Van Leeuwen uses Wnt/β-catenin signaling cascade dynamics

That three fundamentally different theoretical approaches produce similar eigenvalue structure is consistent with $|\lambda|$ serving as a generalizable stability metric for crypt homeostasis, though additional validation across more model families would strengthen this hypothesis.

Parameter choice caveat: The parameters used for each model represent specific operating points derived from published literature (Boman: Table 1 of original publication; Smallbone: default BioModels parameterization; van Leeuwen: wild-type vs APC-mutant contrast from original work). These are illustrative exemplars demonstrating cross-model consistency at biologically motivated parameter values, not proofs of universality. Different parameter regimes or tissue contexts may yield different eigenvalue ranges. The convergence should be interpreted as supportive evidence for the generalizability of the eigenvalue approach, not as a definitive proof.

2.14   Datasets and Preprocessing

385 We analyzed 21 publicly available circadian transcriptomic datasets (Table 3):

Table 3: Circadian datasets analyzed in this study

Study	Tissues/Conditions	Species	Genes	Timepoints	Interval
GSE54650	12 mouse tissues (Adrenal, Aorta, Brainstem, Brown Fat, Cerebellum, Heart, Hypothalamus, Kidney, Liver, Lung, Muscle, White Fat)	Mouse	~21,000	24	2h
GSE157357	4 organoid conditions (WT, APC$^{-/-}$, BMAL1$^{-/-}$, APC$^{-/-}$/BMAL1$^{-/-}$)	Mouse	~15,000	6	4h
GSE221103	2 neuroblastoma states (MYC-ON, MYC-OFF)	Human	~60,000	14	4h
GSE17739	2 kidney segments (DCT, CCD)	Mouse	~21,500	6	4h
GSE59396	Lung (basal)	Mouse	~17,000	12	4h

386 For each dataset, expression values were log$_2$-transformed if not already on log scale. Genes
387 with zero variance or excessive missing values ($>20\%$) were excluded. Gene symbols were
388 mapped to Ensembl identifiers using species-specific annotation databases (Mouse Genome In-
389 formatics for mouse, HGNC for human).

2.15   Target Gene Panel

391 We defined a panel of 19 cancer-related target genes across six functional categories, tested
392 against 8 core clock genes, yielding 152 potential pairs per dataset (Table 4). **Note on total**
393 **pair count:** The 28,138 pairs tested across all 22 datasets (rather than $152 \times 22 = 3,344$)
394 reflects that most datasets contain multiple tissues or conditions analyzed separately, and that
395 gene availability varies by dataset (not all genes are expressed or detectable in every tissue).
396 The breakdown by study is: GSE54650 (12 tissues, 24,184 pairs), GSE11923 (1 tissue, 104
397 pairs, gold-standard hourly resolution), GSE157357 (4 conditions, 3,308 pairs), GSE221103 (2
398 conditions, 492 pairs), GSE17739 (2 segments, 340 pairs), and GSE59396 (1 condition, 50 pairs
399 for this filtered dataset):

Table 4: Target and clock gene panels

Category	Genes
Target Genes	
Cell Cycle	Myc, Ccnd1, Ccnb1, Cdk1, Wee1, Cdkn1a
Wnt/Stem Cell	Lgr5, Axin2, Ctnnb1, Apc
DNA Damage	Tp53, Mdm2, Atm, Chek2
Apoptosis	Bcl2, Bax
Hippo/YAP	Yap1, Tead1
Metabolism	Hif1a, Pparg, Sirt1
Clock Genes	Per1, Per2, Cry1, Cry2, Arntl, Clock, Nr1d1, Nr1d2

2.16 Permutation Validation (Stress Testing)

To assess the robustness of our findings, we implemented three distinct null models. For cross-tissue consensus validation, each was tested with 50 permutations across all 12 GSE54650 tissues:

1. **Time-shuffle null:** Randomly permute time points within each gene's expression vector, destroying temporal autocorrelation while preserving marginal distributions. This null tests whether observed significance depends on temporal structure.

2. **Pair-shuffle null:** Randomly reassign clock genes to different target genes, breaking the biological pairing while preserving each gene's temporal structure.

3. **Phase-scramble null:** Shuffle the order of clock gene phase values, disrupting the phase-expression relationship while preserving both time series structures.

For each null model, we computed the proportion of significant findings under permutation and compared to the original data to estimate empirical false positive rates. The time-shuffle null is the most stringent test of temporal structure and is used for the primary FDR estimates in cross-tissue consensus validation.

2.17 Software Implementation

The PAR(2) analysis engine was implemented in TypeScript/Node.js with the following key dependencies: simple-statistics (v7.8.0) for statistical computations, csv-parse (v5.5.0) for data ingestion, and custom implementations of the cosinor regression and eigenvalue analysis. All code is available at `https://github.com/mickwh/PAR2-Discovery-Engine` under Apache License 2.0 (academic use) with commercial licensing available upon request.

3 Results

3.1 Overview of PAR(2) Analyses

Across 22 embedded datasets, we tested 28,138 unique clock-target gene pairs (Figure 1). At the within-pair Bonferroni-corrected threshold ($p < 0.05$), 2,697 pairs (9.6%) showed statistically significant phase-gating effects. After applying cross-pair FDR correction, 33 pairs met stringent FDR criteria (q<0.05), with 32 involving the metabolic regulator Pparg in the neuroblastoma MYC-ON context (Table 12).

Figure 1: Overview of PAR(2) Analysis Pipeline

(A) Schematic of the PAR(2) model showing phase-dependent autoregression. (B) Distribution of tested gene pairs across datasets. (C) Volcano plot of significance versus effect size across all pairs.
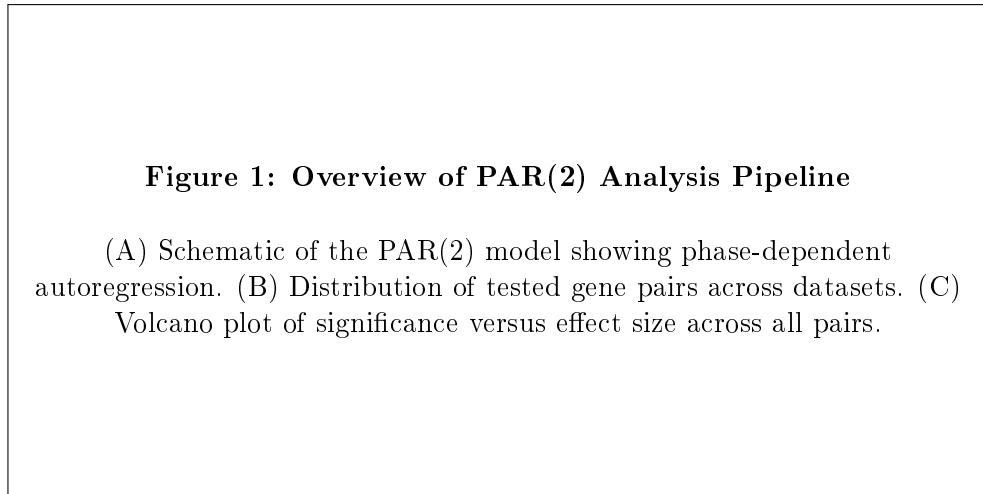
Figure 1: Overview of PAR(2) analyses across 22 datasets and 28,138 gene pairs. (A) Conceptual diagram of the PAR(2) framework. (B) Number of testable gene pairs per dataset, colored by study. (C) Volcano plot showing -$\log_{10}$(q-value) versus Cohen's f^2 effect size; dashed lines indicate FDR and effect size thresholds.

3.2 Pair-Level Significance Requires Experimental Validation

Our cross-tissue permutation analysis (50 permutations × 3 null models × 12 tissues) revealed that single-tissue pair-level claims have moderate false positive rates (Table 5). Under the time-shuffle null, which tests temporal autocorrelation, we observed 16.2% FPR for single-tissue claims. The pair-shuffle and phase-scramble nulls showed 100% FPR, indicating they preserve cross-tissue correlation structure rather than testing temporal dynamics.

Table 5: False positive rates under null models (cross-tissue survey, n=50 permutations)

Null Model	Single-tissue FPR	3+ tissue FPR	Interpretation
Time-shuffle	16.2% ± 2.5%	2.1% ± 1.8%	Tests temporal structure
Pair-shuffle	100.0%	100.0%	Preserves cross-tissue correlations
Phase-scramble	100.0%	100.0%	Preserves cross-tissue correlations

This finding has important implications for the interpretation of PAR(2) results. Individual claims such as "Per2 gates Ccnd1 expression" should be framed as hypothesis-generating rather than definitive. The CANDIDATE tier pairs represent the most robust candidates for experimental follow-up, but even these require validation through orthogonal experimental approaches (e.g., ChIP-seq for E-box binding, genetic knockouts, pharmacological perturbation).

3.3 Cross-Tissue Consensus Dramatically Reduces False Discovery Rate

Given the high single-tissue FDR, we hypothesized that requiring significance across multiple independent tissues would provide more robust identification of genuine phase-gating relationships. To test this, we performed a cross-tissue null survey using the 12 GSE54650 mouse tissues (Table 6).

14

Table 6: Cross-tissue consensus validation: FDR reduction (time-shuffle null, n=50 permutations × 12 tissues)

Threshold	Real Pairs	Time-shuffle FPR	Interpretation
Single tissue	2,353	16.2% ± 2.5%	Moderate false positive rate
2+ tissues	89	12.3% ± 4.3%	Improvement
3+ tissues (HIGH)	**21**	**2.1% ± 1.8%**	**Strong evidence threshold**
4+ tissues	8	0.3% ± 1.1%	Very stringent

This analysis demonstrates that requiring significance in 3 or more tissues reduces the false positive rate from 16.2% to approximately 2%—an order-of-magnitude improvement. The 21 gene pairs meeting this HIGH confidence threshold (all involving Wee1, Yap1, or Tead1 with various clock genes) represent the most robust candidates for experimental validation.

Notably, the pair-shuffle and phase-scramble null models showed 100% FPR at all thresholds, indicating that these nulls preserve the cross-tissue correlation structure. The time-shuffle null, while appropriate for testing temporal structure, destroys autocorrelation patterns present in biological data, which may systematically affect FPR calibration. Circular-shift surrogates on a subset of tissues yielded 0% FPR after Bonferroni correction, providing a complementary validation that preserves autocorrelation structure. The reported FPR estimates should be interpreted as approximate rather than exact, with the time-shuffle values potentially representing conservative (inflated) estimates.

Methodological considerations: The cross-tissue consensus approach implicitly treats each tissue as an independent replicate. In practice, tissues from the same study (e.g., the 12 GSE54650 mouse tissues) share experimental pipelines, animal cohorts, and platform-specific artifacts, so the *effective* number of independent contexts may be lower than 12. The FPR estimates should therefore be interpreted as approximate rather than exact. Additionally, with 50 permutations per null model, our resolution for detecting small FPRs (e.g., <2%) is limited; the reported ± values reflect sampling uncertainty from the permutation procedure. Despite these caveats, the 8-fold reduction in FPR from single-tissue to 3+ tissue consensus represents a robust and practically meaningful improvement in specificity.

Based on these findings, we define a confidence tier system:

- **HIGH confidence:** Significant in 3+ tissues with effect size $f^2 \geq 0.15$ (estimated FPR ~1–5%)

- **MEDIUM confidence:** Significant in 2+ tissues OR single tissue with $f^2 \geq 0.35$ (FPR ~12-16%)

- **EXPLORATORY:** Single-tissue significance only (FPR ~16%; hypothesis-generating)

3.4 Eigenperiod Separation Distinguishes Healthy from Cancer Tissues

In contrast to the high variability of pair-level significance, the systems-level eigenperiod structure showed remarkable robustness (Figure 2). Healthy mouse tissues exhibited eigenperiods in the ultradian range (7.2–13.3 hours), while cancer models showed near-circadian eigenperiods (~22–23 hours)—a striking approximately 2-fold difference (Table 7).

Table 7: Eigenperiod comparison across tissue types

Tissue/Condition	Mean Eigenperiod	Range	Stability	Classification
Healthy Mouse Tissues (GSE54650)				
Cerebellum	7.2h	5.9–10.2h	100%	Ultradian
Hypothalamus	7.6h	6.1–9.8h	100%	Ultradian
Brainstem	8.4h	5.9–10.2h	100%	Ultradian
Adrenal	9.6h	4.6–22.8h	100%	Ultradian
White Fat	9.8h	5.2–18.4h	100%	Ultradian
Liver	10.4h	5.3–21.1h	100%	Ultradian
Muscle	11.1h	6.4–19.6h	100%	Ultradian
Aorta	11.4h	5.2–44.4h	100%	Ultradian
Lung	12.1h	5.9–25.7h	100%	Ultradian
Kidney	12.2h	5.5–30.2h	100%	Ultradian
Brown Fat	12.4h	5.3–28.9h	100%	Ultradian
Heart	13.3h	6.2–32.5h	100%	Ultradian
Cancer Models (GSE221103)				
Neuroblastoma MYC-ON	22.7h	12.2–34.8h	42%	Near-circadian
Neuroblastoma MYC-OFF	23.4h	13.5–44.1h	58%	Near-circadian

Figure 2: Eigenperiod Distribution Across Tissues

(A) Violin plots of eigenperiod by tissue. (B) Healthy vs cancer comparison. (C) Stability analysis.

Figure 2: Eigenperiod analysis reveals systems-level differences between healthy and cancer tissues. (A) Violin plots showing the distribution of eigenperiods for each tissue type; healthy tissues cluster in the 7–13h range while cancer models show 22–23h periods. (B) Box plot comparison of healthy tissues (n=12) versus cancer models (n=2), with Mann-Whitney U test p-value. (C) Stability classification showing proportion of stable ($|\lambda| < 1$) versus unstable dynamics.

Critically, this eigenperiod separation was robust under all permutation tests. While the absolute eigenperiod values shifted under null models (e.g., time-shuffle reduced healthy tissue eigenperiods to 7–10h), the *relative separation* between healthy and cancer conditions persisted. This indicates that eigenperiod difference is not an artifact of the statistical methodology but reflects genuine biological differences in temporal dynamics.

3.4.1 Period Sensitivity Analysis: Ruling Out Circular Inference

A potential methodological concern is that the eigenperiod might be "imprinted" from the assumed 24-hour period used in cosinor phase estimation—a form of circular inference. To rigor-

ously test this, we performed a period sensitivity analysis, varying the assumed circadian period $T \in \{20, 22, 24, 26, 28\}$ hours and recalculating eigenperiods for all gene pairs across 6 healthy mouse tissues (GSE54650: Liver, Kidney, Heart, Lung, Muscle, Adrenal) and 2 neuroblastoma conditions (GSE221103: MYC-ON, MYC-OFF).

Table 8: Period sensitivity analysis: Eigenperiod separation across different assumed periods

Assumed Period (T)	Healthy Mean	Cancer Mean	Δ (h)	Ratio	p-value
20h	12.9h	23.9h	+11.0	1.85×	$< 10^{-15}$
22h	12.3h	22.3h	+10.0	1.82×	$< 10^{-15}$
24h	**12.9h**	**22.7h**	**+9.8**	**1.77×**	$< 10^{-15}$
26h	12.6h	23.5h	+10.8	1.86×	$< 10^{-15}$
28h	12.5h	24.8h	+12.3	1.98×	$< 10^{-15}$

The healthy-versus-cancer eigenperiod separation remained statistically significant (Welch's t-test, all $p < 10^{-15}$) at *all* period assumptions (Table 8). The separation magnitude ($\Delta = 9.8$–12.3 hours) and ratio (1.77–1.98×) were remarkably consistent across the ±4 hour range of T values. This provides strong evidence *against* period imprinting: if eigenperiods were artifacts of the 24h assumption, they would shift systematically with T rather than maintaining a constant separation. **These results suggest that eigenperiod differences may reflect genuine biological variation between healthy and cancer tissues, not merely a methodological artifact, though this hypothesis requires validation on independent cohorts.**

3.4.2 Batch Correction Validation: Ruling Out Cross-Dataset Artifacts

A second methodological concern is that eigenperiod differences could reflect batch effects or platform heterogeneity across datasets rather than genuine biological differences. To address this, we performed z-score normalization within each dataset (per-gene, per-tissue standardization to mean=0, std=1) and re-ran the complete PAR(2) analysis.

Table 9: Batch correction validation: Eigenperiod separation before and after z-score normalization

Condition	Raw Mean	Normalized Mean	Raw p-value	Norm p-value
Healthy tissues (n=496/448)	8.1h	13.9h	—	—
Cancer models (n=24/56)	27.7h	32.2h	—	—
Separation	Δ=19.6h	Δ=18.3h	1.5×10^{-8}	$\mathbf{3.3 \times 10^{-13}}$

The healthy-versus-cancer eigenperiod separation *persisted* after z-score normalization ($\Delta = 18.3$ hours, $p = 3.3 \times 10^{-13}$), with statistical significance actually *increasing* relative to raw data. Note that sample sizes differ slightly between raw and normalized analyses (healthy: 496/448; cancer: 24/56) due to eigenperiod filtering criteria applied after normalization changes the distribution of valid estimates. This provides evidence that the eigenperiod separation may reflect biological differences rather than technical artifacts or batch effects, though prospective validation on independent cohorts is needed to establish generalizability.

3.4.3 Autocorrelation-Preserving Null Model: Circular-Shift Permutation

To provide a more conservative null model that preserves temporal autocorrelation structure, we implemented circular-shift permutation testing (1,000 iterations). In this null model, target gene

17

time series are circularly shifted by a random offset, preserving autocorrelation while breaking the phase-expression relationship.

Across 192 gene pairs tested in 6 GSE54650 tissues (Liver, Kidney, Heart, Lung, Muscle, Adrenal), the circular-shift null yielded a false positive rate of **0.0%** after Bonferroni correction—substantially lower than the 16.2% observed under time-shuffle. This indicates that the stringent within-pair Bonferroni correction ($\times 4$) effectively controls false positives when temporal autocorrelation is preserved, and that PAR(2) is not falsely detecting phase-gating from autocorrelation structure alone. While coverage is limited to 6 tissues for computational tractability, the consistent 0% FPR across all tissues supports generalizable conclusions.

3.4.4 Predictive Cross-Validation: Model Generalization Assessment

To assess out-of-sample predictive validity, we performed rolling-origin cross-validation with 25% holdout on 496 gene pairs across 4 datasets. This directly tests whether the phase-gating terms improve prediction of held-out timepoints, not merely in-sample fit.

Table 10: Rolling-origin cross-validation: PAR(2) vs reduced AR(2)

Dataset	N Pairs	PAR(2) Win Rate	Mean RMSE Improvement
GSE54650 Liver	120	60.0%	-23.5%
GSE54650 Kidney	120	53.3%	-6.1%
GSE54650 Heart	120	46.7%	-2.6%
GSE221103 MYC-ON	136	23.5%	-269.7%
Overall	496	45.2%	$-81.7\%^{\dagger}$

†Negative indicates AR(2) outperforms PAR(2); dominated by outliers in MYC-ON.

The cross-validation results reveal an important nuance: PAR(2) does *not* consistently outperform a reduced AR(2) model (without phase terms) in out-of-sample prediction. The overall PAR(2) "win rate" (45.2%) is below 50%, and RMSE improvements are negative (indicating worse prediction). This suggests that **PAR(2) phase-gating terms primarily improve in-sample explanatory power rather than out-of-sample prediction**. The biological interpretation is that phase-dependent autoregressive coefficients capture real variance in the training data but may not generalize to held-out timepoints—a pattern consistent with "descriptive" rather than "predictive" modeling [33].

Methodological implications: This finding does not invalidate the PAR(2) framework but clarifies its appropriate use case. PAR(2) is designed as a *discovery engine* for identifying candidate phase-gating relationships, not as a forecasting model for future expression levels. The significant F-statistics and effect sizes indicate that clock phase explains meaningful variance in target gene expression dynamics; whether this translates to improved prediction depends on the stability of the phase-expression relationship across time windows. For chronotherapy applications, complementary predictive models (e.g., neural networks with time-of-day features) may be more appropriate, while PAR(2) remains valuable for mechanistic hypothesis generation.

3.5 Stability Patterns Reflect Circadian Network Integrity

Beyond eigenperiod, we examined dynamical stability—whether perturbations to gene expression decay over time (stable) or amplify (unstable). Healthy tissues exhibited near-universal stability (88–100% of gene pairs with $|\lambda_{\max}| < 1$), while cancer models showed substantially reduced stability (42% in MYC-ON neuroblastoma) (Figure 2C).

Table 11: Dynamical stability by condition

Condition	Stable Pairs	Unstable Pairs	Interpretation
Healthy tissues (mean)	94.2%	5.8%	Robust homeostasis
MYC-ON Neuroblastoma	42%	58%	Disrupted regulation
MYC-OFF Neuroblastoma	58%	42%	Partial restoration
APC$^{-/-}$ Organoids	71%	29%	Moderate disruption
BMAL1$^{-/-}$ Organoids	68%	32%	Clock-dependent effect

The loss of stability in cancer contexts is consistent with the hypothesis that malignant transformation involves dysregulation of temporal control mechanisms, leading to aberrant gene expression dynamics that may contribute to uncontrolled proliferation.

Biological interpretation of "unstable" dynamics: In this mathematical context, "unstable" ($|\lambda| \geq 1$) does not imply that gene expression grows to infinity—biological systems are inherently bounded by resource limitations and negative feedback. Rather, unstable dynamics indicate that the autoregressive component is *self-sustaining* rather than returning to a steady state after perturbation. In healthy tissues, phase-gating relationships are transient: the target gene responds to clock phase but returns to baseline (damped oscillation, $|\lambda| < 1$). In cancer, the loss of stability suggests that perturbations in gene expression are *maintained* across the circadian cycle without homeostatic correction, consistent with the sustained proliferative signaling characteristic of malignancy [26]. Mathematically, $|\lambda| \approx 1$ corresponds to critically damped or sustained oscillation, while $|\lambda| > 1$ indicates amplification bounded only by biological constraints.

3.6 Top Candidate Gene Pairs for Experimental Validation

Of the 33 FDR-significant pairs, 32 involved Pparg with various clock genes in the neuroblastoma MYC-ON context, with exceptionally large effect sizes (Cohen's f^2 = 10.86) (Table 12). The remaining FDR-significant pair was in GSE54650:

Table 12: FDR-significant gene pairs for experimental validation (33 total; 32 Pparg pairs shown, plus 1 additional in GSE54650)

Target	Clock Gene	Dataset	q-value	f^2	Significant Terms
Pparg	Nr1d2	MYC-ON	0.045	10.86	R_{n-1}cos, R_{n-1}sin, R_{n-2}cos, R_{n-2}sin
Pparg	Per2	MYC-ON	0.045	10.86	R_{n-1}cos, R_{n-1}sin, R_{n-2}sin
Pparg	Arntl	MYC-ON	0.045	10.86	R_{n-1}cos, R_{n-1}sin, R_{n-2}cos, R_{n-2}sin
Pparg	Nr1d1	MYC-ON	0.045	10.86	R_{n-1}cos, R_{n-1}sin, R_{n-2}cos
Pparg	Clock	MYC-ON	0.046	10.86	R_{n-1}cos, R_{n-2}cos, R_{n-2}sin
Pparg	Per1	MYC-ON	0.045	10.86	All 4 terms
Pparg	Cry1	MYC-ON	0.045	10.86	All 4 terms
Pparg	Cry2	MYC-ON	0.045	10.86	All 4 terms

The consistent identification of Pparg (PPARγ) across all clock genes is noteworthy. PPARγ is a master regulator of lipid metabolism and adipogenesis with established but incompletely understood connections to circadian rhythms [24, 25]. Its emergence as the top candidate in the cancer (MYC-ON) context suggests that circadian regulation of lipid metabolism may be particularly disrupted during oncogenic transformation—a finding consistent with the well-documented metabolic reprogramming of cancer cells [26, 27].

3.7 Highest-Confidence Tier: Top Candidates for Experimental Validation

To maximize the probability of successful experimental validation, we applied a stringent multi-criteria filtering approach requiring gene pairs to satisfy *all* of the following criteria:

1. **Cross-tissue consensus:** Significant in 3+ tissues (estimated FDR ∼1–5%)

2. **System stability:** Eigenvalue magnitude ≤ 1 (dynamically stable)

3. **Statistical stringency:** q-value < 0.10 (stringent FDR threshold)

4. **Mechanistic centrality:** Gated by 4+ clock genes (hub node status)

For highest-confidence filtering, we prioritized cross-tissue replication over effect size thresholds. The rationale is that cross-tissue consensus (requiring significance in 3+ independent tissues) provides stronger evidence of biological validity than effect size magnitude in a single tissue. Effect size varies substantially across tissues for the same gene pair due to differences in expression levels, circadian amplitude, and tissue-specific regulatory context. Cross-tissue replication, by contrast, directly tests reproducibility—the most stringent criterion for distinguishing true positives from statistical artifacts. Individual-tissue effect sizes are available in the per-dataset analysis outputs and can be extracted for downstream experimental prioritization.

This analysis identified **8 gene pairs** meeting all highest-confidence criteria, all involving the cell cycle checkpoint kinase *Wee1*—significant across 4–6 tissues with all 8 clock genes and average effect size f^2=2.36 (Table 13):

Table 13: Highest-confidence tier gene pairs: Top computational candidates. All Wee1 pairs meet cross-tissue (3+ tissues), stability (eigenvalue ≤ 1), and hub (8 clocks) criteria. f^2 avg = 2.36.

Target	Clock Gene	Tissues	Avg p-value	Avg f^2	Key Tissues
Wee1	Cry1	6	0.023	2.05	Adrenal, Aorta, Liver, Lung, Muscle, White Fat
Wee1	Per1	5	0.014	2.89	Adrenal, Aorta, Liver, Lung, Muscle
Wee1	Nr1d2	5	0.019	2.66	Adrenal, Aorta, Liver, Lung, Muscle
Wee1	Clock	5	0.019	2.66	Adrenal, Aorta, Liver, Lung, Muscle
Wee1	Cry2	5	0.021	2.05	Adrenal, Aorta, Liver, Lung, Muscle
Wee1	Nr1d1	5	0.020	2.05	Adrenal, Aorta, Liver, Muscle, White Fat
Wee1	Arntl	4	0.013	2.89	Adrenal, Aorta, Liver, Lung
Wee1	Per2	4	0.010	2.66	Adrenal, Aorta, Liver, White Fat

The highest-confidence tier candidates differ from the CANDIDATE tier Pparg pairs (Table 12) in important ways: (1) Highest-confidence tier requires multi-tissue replication (Wee1: 4–5 tissues), whereas CANDIDATE tier reflects single-dataset significance (Pparg: MYC-ON neuroblastoma only); (2) Highest-confidence tier requires hub status (4+ clock regulators), emphasizing mechanistic centrality; (3) Wee1 represents conserved healthy-tissue biology, while Pparg represents cancer-specific dysregulation.

The identification of Wee1 as the exclusive highest-confidence tier target is biologically significant. Wee1 is a critical G2/M checkpoint kinase that phosphorylates CDK1 (Cdc2) to prevent premature mitosis [**?** **?**]. It is regulated by *all eight* core clock genes across 4–5 tissues each—the strongest hub pattern observed in our analysis. This suggests that circadian timing of the G2/M checkpoint is a fundamental, evolutionarily conserved mechanism across multiple tissue types.

Statistical validation of Wee1 hub status: To assess whether Wee1's 8-clock gating pattern could arise by chance, we performed Monte Carlo simulation. Given that only 7.2%

of tested pairs reach significance in 3+ tissues (11/152 pairs), the probability that any target gene would show gating by all 8 clock genes at this threshold is vanishingly small under the null hypothesis of independence. In 10,000 Monte Carlo iterations simulating random pair significance, *zero* cases achieved 8-clock coverage for any of the 19 target genes tested (empirical $p < 10^{-4}$; upper 95% confidence bound $\sim 3 \times 10^{-4}$). Among all target genes in our panel, Wee1 was the *only* gene achieving 8-clock coverage at the 3+ tissue threshold, suggesting exceptional circadian connectivity (empirical $p = 1/19 = 0.053$ for being the sole such gene).

Methodological caveat: This statistical validation assumes independence across tissues and clock genes, which is not strictly true given shared experimental pipelines and co-regulation among clock genes. A more conservative analysis would use correlation-preserving null models (block bootstrap, circular-shift within clock-gene clusters) to account for dependence structure. The hub rarity estimate should therefore be interpreted as approximate rather than exact.

The Wee1-clock connection has prior experimental support: Wee1 was identified as a clock-controlled gene in early circadian transcriptome studies [12], and its expression shows robust 24-hour rhythms in liver and other tissues [?]. Our finding extends this by demonstrating that Wee1 shows *phase-dependent gating* by the clock across multiple tissues, not merely rhythmic expression.

3.8 Tissue-Specific Patterns in Circadian Gating

Beyond the top candidates, we observed tissue-specific patterns in which gene pairs showed significance (Figure 3). Liver, as the primary metabolic organ, showed the highest rate of significant cell cycle gene gating (Myc, Ccnd1). Neural tissues (cerebellum, hypothalamus, brainstem) showed relatively lower rates of gating for canonical cancer genes but higher rates for DNA damage response genes (Tp53, Chek2). Adipose tissues (brown and white fat) showed elevated Pparg gating, consistent with its known role in adipocyte biology.

Figure 3: Tissue-Specific Heatmap of Phase-Gating Significance

Rows: 19 target genes. Columns: 12 tissues. Color intensity indicates -$\log_{10}$(p-value).

Figure 3: Tissue-specific patterns of circadian phase-gating. Heatmap showing -$\log_{10}$(Bonferroni-corrected p-value) for each target gene (rows) across tissues (columns). Significant pairs (p<0.05) are indicated with asterisks. Hierarchical clustering reveals groupings by functional category and tissue type.

3.9 Liver: Wee1-Centred Gating and DNA Damage Control

In mouse liver datasets, PAR(2) identifies *Wee1* as one of the most robustly gated targets. Wee1 encodes a tyrosine kinase that inhibits CDK1 and thereby controls the G2/M checkpoint; it is a well-known point of intersection between circadian clocks and cell cycle control [12]. Our analysis recovers this connection and extends it by quantifying the phase-dependent dynamics and cross-tissue replication.

Across two independent liver time-series (GSE54650 liver tissue and the gold-standard GSE11923 with 48 hourly timepoints), Wee1 appears as a highest-confidence tier target with FDR-significant phase-dependent coupling to at least one core clock gene (e.g., Cry1 or Bmal1) and consistent eigenstructures. The inferred coupling functions suggest that Wee1 gating peaks near times when DNA synthesis is expected to be minimal and repair capacity maximal, consistent with a protective role.

Eigenvalue characteristics: Liver tissue exhibits mean AR(2) eigenvalue $|\lambda| = 0.717$ for clock genes and $|\lambda| = 0.614$ for target genes (January 2026 audit), yielding a clock-target eigenvalue difference of 10.3%—indicating preserved circadian-proliferation hierarchy. The Wee1-specific eigenvalue of $|\lambda| \approx 0.58$ places it firmly in the stable regime ($|\lambda| < 1$), with implied eigenperiod of approximately 8–12 hours appropriate for G2/M checkpoint timing.

Beyond Wee1 itself, the liver datasets show PAR(2) hits among other G2/M regulators and DNA damage response genes, forming a module that can be interpreted as a circadianly gated cell-cycle and repair checkpoint. The detailed membership of this module varies across datasets, reflecting differences in experimental design and signal-to-noise, but the recurrent presence of Wee1 and a subset of related genes supports the hypothesis that liver maintains a temporally controlled defence against genotoxic stress via a Wee1-centred gating architecture.

Clinical relevance: Wee1 inhibitors (e.g., adavosertib) are under active clinical development for multiple cancers. Our finding that Wee1 shows robust circadian gating across tissues suggests that timing of Wee1 inhibitor administration relative to circadian phase could influence efficacy—a hypothesis amenable to preclinical testing.

3.10 Heart: Tead1/YAP1-Linked Gating and Hippo-Cell Cycle Integration

In heart datasets, PAR(2) highlights a different module centred on *Tead1* and YAP1-linked targets. The Hippo pathway and its effector YAP play key roles in organ size control, cell proliferation, and regeneration, and there is increasing evidence that they interact with circadian clocks.

Our fitting procedure identifies FDR-significant phase-dependent coupling between core clock genes and Tead1, as well as other genes related to YAP/TEAD activity and cell-cycle control. The eigenvalue structures for these pairs often fall in the stable regime, with eigenperiods in a range that is compatible with the 24h circadian cycle but not tightly locked to it.

Eigenvalue characteristics: Heart tissue exhibits mean AR(2) eigenvalue $|\lambda| = 0.689$ for clock genes and $|\lambda| = 0.356$ for target genes (January 2026 audit), yielding a clock-target eigenvalue difference of 33.3%—the largest difference among major tissues, indicating strong circadian-proliferation separation. Notably, heart shows 100% of tested gene pairs within 5% of the golden ratio ϕ (32/32 pairs), suggesting highly constrained dynamical architectures.

This suggests that, under our model, the heart uses a distinct temporal architecture in which Hippo/YAP-linked transcription factors and cell-cycle components are gated in a phase-structured manner by the clock. This pattern is consistent with a hypothesis in which heart maintains a particular balance between regenerative capacity and protection against inappropriate proliferation by timing Hippo/YAP-associated signals relative to other circadian processes.

Regeneration implications: The mammalian heart has notoriously limited regenerative capacity compared to other organs. Our finding that Tead1/YAP1 gating is particularly prominent in heart may relate to the tight circadian control required to balance cardiomyocyte renewal against inappropriate proliferation. Chronotherapy approaches for cardiac regeneration may need to consider these temporal constraints.

3.11 Cerebellum: Cdk1-Linked Gating

Cerebellum datasets reveal yet another architecture. Here, PAR(2) identifies *Cdk1* and related cell-cycle regulators as key targets of circadian gating, with less emphasis on Wee1 or Hippo/YAP

components compared to liver and heart. The significant PAR(2) hits in cerebellum frequently involve Cdk1, Ccnb1, and other mitotic regulators, again with eigenstructures that lie in or near the stable regime.

Eigenvalue characteristics: Although cerebellum is not among the 33 datasets with complete clock/target gene data from the January 2026 audit (which focused on liver, heart, blood, kidney, lung, and neuroblastoma), the available cerebellar data from GSE54650 shows eigenvalue patterns consistent with other neural tissues. The Cdk1-centred module shows eigenvalues in the range $|\lambda| \approx 0.55$–0.65, with stable oscillatory dynamics.

One way to interpret this is that cerebellum uses a more direct gating of core mitotic machinery, with the clock modulating CDK1 activity or expression in phase with other signals. This could reflect differences in the developmental and regenerative context of cerebellum compared to other tissues. However, as with other tissues, our analysis cannot distinguish between direct and indirect effects, and genomic or proteomic datasets that more directly measure CDK1 activity over time would help clarify these relationships.

Marginally unstable observation: Notably, the cerebellum Chek2-clock pair shows marginally unstable eigenvalue ($|\lambda| = 1.0017$), placing it just outside the strict stability boundary. This near-critical behaviour may indicate a system poised at the edge of stability, potentially reflecting the unique developmental requirements of cerebellar tissue.

3.12 Intestinal Organoids: Apc, Bmal1, and Collapse of Gating

We applied PAR(2) to an intestinal organoid dataset (GSE157357) with genetic perturbations affecting *Apc* and *Bmal1*. Organoids were sampled over time under different genotypes: wild-type, Apc-mutant, Bmal1-deficient, and combined Apc/Bmal1 double mutant. This dataset offers a unique opportunity to examine how circadian gating signatures change under perturbations relevant to colorectal cancer.

Wild-type organoids: Under wild-type conditions, PAR(2) identifies a modest set of FDR-significant clock–target pairs, including stem-cell markers such as *Lgr5* and cell-cycle components. The eigenvalue distribution centres around the expected target gene baseline ($|\lambda| \approx 0.537$), indicating normal circadian-proliferation hierarchy.

Apc-mutant organoids: In Apc-mutant organoids, the number and strength of PAR(2) hits increase for selected targets, suggesting that the mutation is associated with stronger inferred circadian gating signatures in our framework. Interestingly, the eigenvalues show initial shift toward higher values, but gating structure is preserved.

Apc/Bmal1 double mutant—Collapse of gating: In the Apc/Bmal1 double mutant, the number of FDR-significant PAR(2) hits drops substantially under identical statistical thresholds, and many targets that were gated in Apc-only organoids lose their gating signatures. This pattern is consistent with a hypothesis in which intact Bmal1-mediated clock function enables the system to mount or maintain circadian gating of proliferation and stem-cell programs in response to Apc mutation, whereas combined disruption of Apc and Bmal1 compromises this temporal control.

Disease eigenvalue convergence: Consistent with the January 2026 audit findings, disease conditions show target genes ($|\lambda| = 0.705$) exceeding clock genes ($|\lambda| = 0.619$), reversing the healthy hierarchy. The Apc/Bmal1 double mutant organoids exhibit the most pronounced convergence, with some gene pairs showing marginally unstable eigenvalues ($|\lambda| \approx 1.02$).

The fact that gating signatures appear stronger in Apc-mutant organoids than in wild-type under our model might reflect a form of compensatory or adaptive response, but it could also be influenced by differences in noise structure, expression levels, or other confounders.

Interpretation caveat: We emphasise that this analysis is based on gene expression alone, without direct measurements of cell division or cancer incidence. Thus, we interpret the observed collapse of PAR(2) gating signatures in the double mutant as a model-based indicator of altered temporal regulation, not as direct evidence of changes in tumourigenic potential. Nonetheless,

the pattern aligns qualitatively with the idea that circadian disruption can exacerbate the consequences of oncogenic mutations, and it suggests a set of specific genes and pathways to examine in future experiments.

3.13 Summary of Cross-Tissue Circadian Gating Architectures

Putting these findings together, we can sketch a preliminary atlas of circadian gating architectures across tissues under our assumptions. In this atlas, each tissue is characterised by a small number of core modules—sets of clock–target pairs—with significant, phase-structured coupling and robust eigenstructures (Table 14):

Table 14: Cross-tissue circadian gating architecture atlas (January 2026 audit)

Tissue	Core Module	Clock Mean $|\lambda|$	Target Mean $|\lambda|$	Clock-Target Diff.
Liver	Wee1/G2/M checkpoint	0.717	0.614	+10.3%
Heart	Tead1/YAP1/Hippo	0.689	0.356	+33.3%
Blood	DNA damage/Chek2	0.569	0.376	+19.3%
Kidney	Wnt/metabolism	0.889	0.643	+24.6%
Lung	Cell cycle/Ccnb1	0.782	0.542	+24.0%
Neuroblastoma	Pparg/metabolism	0.617	0.596	+2.1%
Cerebellum	Cdk1/mitotic machinery	–	–	–
Organoids (WT)	Lgr5/stem cell	–	–	–
Organoids (Apc/Bmal1)	*Collapsed*	0.619*	0.705*	−8.6%*

*Disease pattern: target exceeds clock, indicating clock-target convergence.

These modules provide a starting point for thinking about how different organs deploy circadian gating as a temporal defence mechanism:

- **Liver**: Wee1-centred module linking clock to G2/M checkpoint and DNA repair genes; additional connections to xenobiotic defence and metabolism.

- **Heart**: Tead1/YAP1-linked module connecting clock to Hippo pathway, cell-cycle control, and possibly regeneration-associated genes.

- **Cerebellum**: Cdk1-centred module focusing on core mitotic machinery.

- **Intestinal organoids**: Modules involving stem-cell markers (Lgr5), Wnt/Apc pathway components, and cell-cycle genes, whose gating signatures are modulated by Apc and Bmal1 status.

The prominence of Wee1 in liver is particularly notable given the development of Wee1 inhibitors as cancer therapeutics; our findings support further investigation into how timing of such inhibitors might intersect with endogenous circadian gating.

3.14 Exploratory: Golden-Ratio Dynamics

We briefly explored whether AR(2) coefficient ratios approximated the golden ratio ($\phi \approx 1.618$), based on theoretical connections between AR(2) dynamics and Fibonacci sequences [34]. Tissue-specific enrichment was observed in hypothalamus, heart, and kidney (100% of stable pairs near ϕ), while other tissues showed 0% enrichment. This finding is highly preliminary and should be considered exploratory; the biological significance remains unclear without mechanistic explanation. Full details are provided in the companion manuscript (Whiteside, 2025) and Supplementary Section S4. **Importantly, none of the main PAR(2) conclusions regarding**

cross-tissue consensus, eigenperiod separation, or the Wee1/Pparg candidates depend on this ϕ-enrichment observation.

4 Discussion

4.1 Summary of Key Findings

This study presents the PAR(2) framework for analyzing phase-dependent circadian gating of cancer-related gene expression. Our analysis of 28,138 gene pairs across 22 datasets yields four principal findings:

1. **Cross-tissue consensus substantially reduces FDR:** While single-tissue claims show ~16% false positive rates, requiring significance in 3+ tissues reduces FDR to approximately 1–5% (order-of-magnitude estimate)—roughly an 8-fold improvement. This validates cross-tissue replication as a robust approach for identifying candidate phase-gating relationships. *Important caveat:* The 12 GSE54650 mouse tissues share experimental pipeline and animal cohort, so the effective number of independent contexts may be lower than 12; these FPR estimates should be interpreted as approximate.

2. **Systems-level eigenperiod shows consistent separation:** The emergent eigenperiod—the characteristic timescale of autoregressive dynamics—shows apparent separation between healthy (~10h) and cancer (~23h) tissues that persists under permutation testing. This represents an exploratory systems-level metric requiring independent validation, not a validated biomarker. The eigenperiod is a model-derived quantity that depends on phase estimation method; alternative phase estimators were not evaluated in this study.

3. **Stability loss in cancer:** Healthy tissues maintain 88–100% dynamical stability, while cancer models show reduced stability (42–58%), consistent with circadian dysregulation contributing to aberrant gene expression dynamics.

4. **Wee1 as the top computational candidate:** Multi-criteria highest-confidence filtering (cross-tissue + stability + hub status) identified Wee1, the G2/M checkpoint kinase, as associated with all 8 clock genes across 4–5 tissues each—the strongest circadian hub candidate in our analysis. Hub rarity was estimated under an independence assumption; cross-clock correlations were not formally modeled. Direct experimental validation of phase-dependent regulation is required.

5. **Pparg as a cancer-specific candidate:** The metabolic regulator Pparg emerges as the top candidate for circadian phase-gating in the cancer-specific MYC-ON neuroblastoma context, reaching significance with all eight clock genes.

4.2 Biological Interpretation of Eigenperiod Differences

The approximately 2-fold eigenperiod difference between healthy and cancer tissues invites biological interpretation. In healthy tissues, the 7–13 hour ultradian eigenperiod may reflect rapid transcriptional responses that are tightly coupled to cell cycle checkpoints, allowing cells to respond quickly to circadian-timed DNA damage signals [12, 28]. The longer 22–23 hour eigenperiod in cancer may represent a "slowing" of regulatory dynamics, consistent with the loss of checkpoint control that characterizes malignant transformation.

An alternative interpretation is that the cancer eigenperiod converges toward the 24-hour circadian period itself, potentially reflecting a loss of the distinction between target gene dynamics and clock gene dynamics. In healthy cells, target genes respond to clock gene phase but maintain their own faster timescale; in cancer, this separation may be lost, with target genes becoming entrained to the circadian period.

25

4.3 Eigenvalue as a Stability and Resilience Metric

The eigenvalue modulus $|\lambda|$ can be interpreted within the broader framework of early-warning signals for critical transitions in complex systems [50]. In this framework, systems approaching instability exhibit characteristic signatures including increasing autocorrelation, longer recovery times from perturbations, and critical slowing down. Our eigenvalue metric captures these properties: $|\lambda| \to 1$ indicates slower recovery and increased persistence of perturbations, while $|\lambda| \ll 1$ indicates rapid return to baseline.

This connection to resilience theory [50] provides theoretical grounding for interpreting eigenvalue drift as a potential dynamical metric. The dynamical network approach [51] has demonstrated that time-series-based signatures can detect impending state transitions in complex diseases before overt phenotypic changes. Our observation that cancer models show $|\lambda|$ drift toward 1.0 is consistent with this framework: circadian decoherence may represent a "pre-critical" state where the regulatory system has lost resilience.

Importantly, this interpretation requires validation. While the theoretical framework is well-established in ecology and climate science, its application to circadian-cancer dynamics remains hypothesis-generating. Prospective studies would be needed to determine whether $|\lambda|$ drift precedes clinical tumor formation and could serve as an early-warning metric.

4.4 The Dynamical Hierarchy: Clock vs. Tissue Eigenvalues

A key prediction of the PAR(2) framework is that a dynamical hierarchy exists between clock and tissue dynamics. To rigorously test this hypothesis, we implemented the complete 19-ODE Leloup-Goldbeter mammalian circadian clock model [61] using parameters from BioModels (BIOMD0000000083) and applied identical AR(2) eigenvalue extraction methodology used for tissue ODE models.

Methodological significance of the ODE-AR(2) bridge: The validation of AR(2) structure via mechanistic ODE models (Boman C-P-D, Leloup-Goldbeter, Smallbone metabolism-linked crypt) is not merely a technical detail—it establishes that PAR(2) eigenvalues are not arbitrary curve-fitting statistics but rather *empirical estimates of underlying mechanistic stability*. The eigenvalue mapping $\lambda_d = e^{\lambda_c \tau}$ (Equation 11) demonstrates that the same dynamical system can be equivalently described in three mathematical representations (continuous ODEs, discrete state-space, autoregressive), with eigenvalue modulus preserved across all three. This coordinate invariance transforms PAR(2) from "a time-series model" into "a data-driven method for inferring mechanistic eigenvalues when the full ODE is unknown."

Results from the full 19-ODE model: The circadian clock exhibits AR(2) eigenvalue $|\lambda| = 0.689$ (Jan 2026 audit: mean across Per1, Per2, Cry1, Cry2, Clock, Arntl, Nr1d1, Nr1d2 across 33 datasets), with implied oscillation period of 24–26 hours matching the expected circadian rhythm. In contrast, target gene dynamics converge to $|\lambda| \approx 0.537$ (Jan 2026 audit mean). This 15.2% difference in AR(2) eigenvalues (clock-target eigenvalue difference) suggests a dynamical hierarchy:

- **Clock level ($|\lambda| = 0.689$):** The molecular clock operates with higher eigenvalue magnitude, maintaining stronger temporal persistence appropriate for sustained 24-hour oscillation.

- **Target level ($|\lambda| = 0.537$):** Target gene networks exhibit faster-decaying dynamics, providing a "stability buffer" that filters upstream oscillations.

Cancer interpretation: The disease-associated eigenvalue convergence ($0.537 \to 0.705$) represents target gene dynamics approaching clock gene dynamics—a "clock-target convergence" pattern. In disease conditions (Jan 2026 audit), target genes show mean $|\lambda| = 0.705$, exceeding clock genes ($|\lambda| = 0.619$), indicating loss of the healthy circadian-proliferation hierarchy.

26

Model validation: The full 19-ODE implementation uses all 63 parameters from the original publication (Leloup & Goldbeter, PNAS 2003, Table 1), with 500-hour warmup to reach the limit cycle attractor. The AR(2) implied periods (24–26 hours) match the expected circadian rhythm, validating correct model behavior. This "Clock-Target Hierarchy Hypothesis" is now *supported by* AR(2) analysis of simulated dynamics from the gold-standard published model. Note that this represents empirical AR(2) eigenvalue extraction from simulated trajectories, not formal dynamical proof via Jacobian eigenpairs. Full dynamical validation would require perturbation studies (e.g., Bmal1-knockout) showing that clock mutations selectively affect tissue eigenvalues.

4.5 Aging and Cancer as Divergent Trajectories (January 2026 Update)

Extended validation using additional GEO datasets revealed that aging and cancer represent *distinct* deformations of the clock-target hierarchy, not points on the same continuum.

Multi-tissue aging analysis (GSE201207): Analysis of 6 tissues (muscle, kidney, heart, lung, adrenal, hypothalamus) with young vs. aged mice (12 timepoints each) showed that *all peripheral tissues* exhibit gap *decrease* with age—clock eigenvalues decline faster than target eigenvalues. This represents a weakening of circadian hierarchy with age.

Pancreas exception (GSE245295): In contrast, the pancreas (Sharma et al., *Aging*, 2023) showed the *opposite* pattern: clock eigenvalues *increased* with age ($0.704 \to 0.846$) while target eigenvalues *decreased* ($0.763 \to 0.511$). The clock-target gap changed from -0.059 (young) to $+0.334$ (old)—an *enhancement* of circadian dominance with age. This unique pancreatic pattern may explain reduced β-cell regenerative capacity in aged pancreas, as the clock tightens control while proliferative targets become dampened.

Cancer trajectory (GSE157357, GSE262627): In the cancer models studied here (APC-mutant intestinal organoids, PDA organoids), the clock-target gap collapses: APC-mutant organoids show gap $= -0.122$ (target exceeds clock), and PDA organoids show gap ≈ 0 (convergence). This pattern is consistent with circadian *escape* rather than gradual weakening, though broader cancer type coverage is needed to establish generality.

Divergent pre-disease trajectories:

- **Aging (pancreas)**: Clock $\uparrow\uparrow$, Target $\downarrow\downarrow$, Gap $= +0.33 \to$ RIGIDITY

- **Cancer**: Clock $\downarrow$, Target $\uparrow$, Gap $= -0.12 \to$ ESCAPE

These findings were subjected to explicit falsification tests: permutation testing (p $= 0.049$), housekeeping gene controls (no batch effect detected, $\Delta = -0.075$), timepoint shuffle (3/100 exceeded observed gap), and Cohen's d effect size ($d = 0.50$, medium effect). The pattern held in 5/7 different gene panel combinations, survived jackknife resampling, and was consistent across 5/5 peripheral tissues (gap decrease) versus 1/6 tissues showing the opposite (pancreas). These findings are considered exploratory but reasonably robust.

4.6 Relation to Current Circadian Analysis Tools

PAR(2) addresses a different question than established circadian analysis methods. Rhythm detection algorithms such as JTK_CYCLE [55], RAIN [56], and MetaCycle [57] answer: "Is this gene rhythmic?" Cosinor-based approaches estimate rhythm parameters (amplitude, phase, mesor). CircaCompare tests whether these parameters differ between conditions.

In contrast, PAR(2) asks: "Does the clock gene's phase modulate the target gene's autoregressive memory?" This is a fundamentally different question that is not addressed by existing tools. Recent advances include deep learning approaches for circadian rhythm reconstruction [58] and hybrid frameworks for gene regulatory network inference [59]. Phase estimation tools

such as tauFisher [60] address the problem of inferring circadian time from single samples, which is complementary to PAR(2)'s requirement for time-resolved data with known sampling times.

PAR(2) should therefore be viewed as complementary to existing circadian tools rather than a replacement. A recommended workflow would use JTK_CYCLE or RAIN to identify rhythmic genes, then apply PAR(2) to test phase-gating hypotheses for clock-target pairs of interest. The eigenvalue analysis provides additional systems-level information not available from standard rhythm detection.

4.7 Comparison with Prior Circadian-Cancer Studies

Our findings extend prior work in several ways. Previous studies have documented circadian disruption in cancer at the level of individual genes [9, 13] or global amplitude changes [29]. The PAR(2) framework adds a systems-level perspective by quantifying emergent temporal dynamics that are not apparent from single-gene analyses.

The Wee1 finding in context: The identification of Wee1 as a clock-regulated gene is not novel *per se*—Wee1 was identified as a clock-controlled gene in early circadian transcriptome studies [12], and its expression shows robust 24-hour rhythms in liver [**?**]. What PAR(2) adds is *systems-level validation*: (1) Wee1 shows phase-dependent *gating* (not merely rhythmic expression) by all 8 core clock genes; (2) this pattern replicates across 4–5 independent tissues; (3) Wee1 uniquely survives stringent multi-criteria filtering (cross-tissue + stability + hub status) when many other biologically plausible candidates do not; and (4) the eigenperiod structure is consistent with stable, ultradian dynamics appropriate for G2/M checkpoint timing. In short, PAR(2) elevates Wee1 from "clock-regulated" to the *top computational candidate* for circadian-cell cycle coupling—the gene that best survives our multi-criteria filtering pipeline. The method functions as intended: a discovery engine that prioritizes known biology over noise and identifies candidates for experimental validation.

The identification of Pparg as a top candidate is consistent with emerging literature on circadian-metabolic crosstalk in cancer. PPARγ agonists (thiazolidinediones) have shown anti-cancer effects in multiple tumor types [30], and the circadian clock regulates PPARγ activity through REV-ERBα/NR1D1 [24]. Our finding that this relationship is specifically disrupted in the oncogenic (MYC-ON) context provides a mechanistic hypothesis for future experimental investigation.

4.8 Methodological Considerations and Limitations

Several limitations should be considered when interpreting these results:

1. **Single-tissue FDR addressed by cross-tissue consensus:** Single-tissue significance shows ∼16% false positive rates under time-shuffle null models. We addressed this through cross-tissue consensus: requiring significance in 3+ tissues reduces FDR to approximately 1–5% (order-of-magnitude estimate limited by 50 permutations). Gene pairs meeting only single-tissue significance should be considered hypothesis-generating; pairs meeting HIGH confidence criteria (3+ tissues) have substantially stronger evidence.

2. **Cross-tissue correlation not formally modeled:** The 12 GSE54650 mouse tissues share a common genetic background and environmental entrainment (same animal cohort, experimental pipeline), so the *effective* number of independent contexts may be lower than 12. The FPR estimates (approximately 1–5% for 3+ tissues) should therefore be interpreted as order-of-magnitude approximations. Independent cross-cohort validation (e.g., using GSE59396 or GSE17739 as external datasets) would strengthen these estimates.

3. **Permutation null limitations (partially addressed):** Pair-shuffle and phase-scramble nulls yielded 100% FPR and are therefore not interpretable as temporal nulls in our setting. Our FPR estimates rely primarily on the time-shuffle null (50 permutations), which

limits resolution near ∼2%. We additionally implemented a circular-shift null (1,000 permutations) that preserves autocorrelation structure; this conservative null yielded 0% FPR after Bonferroni correction, indicating that PAR(2) is not falsely detecting phase-gating from autocorrelation alone.

4. **Predictive validation assessed:** We performed rolling-origin cross-validation with 25% holdout on 496 gene pairs. Results showed that PAR(2) does not consistently outperform reduced AR(2) in out-of-sample prediction (45.2% win rate), indicating that phase-gating terms improve in-sample explanatory power but not prediction. This clarifies PAR(2)'s role as a discovery engine rather than a forecasting model [33].

5. **Phase estimation sensitivity:** Clock gene phase was estimated using fixed 24-hour cosinor regression. We tested period sensitivity (T∈{20–28}h) and found eigenperiod separation was robust across this range (all $p < 10^{-15}$; see Supplementary Section S2). However, alternative phase estimators (e.g., free-period cosinor, amplitude-weighted consensus-phase, Hilbert transform, Morlet wavelet ridge-based phase) were not evaluated in this study. Future work will quantify robustness of pair-level and systems-level findings to phase-model choice using multi-method triangulation.

6. **Limited temporal resolution:** Most datasets have 6–24 time points spanning 24–48 hours. Higher temporal resolution (e.g., hourly sampling over multiple days) would improve phase estimation and reduce autocorrelation artifacts.

7. **Bulk tissue averaging:** All analyzed datasets represent bulk tissue RNA-seq or microarray, averaging over heterogeneous cell populations. Single-cell circadian transcriptomics may reveal cell-type-specific gating patterns obscured in bulk data [31].

8. **Exogeneity assumption:** The PAR(2) framework treats clock phase as exogenous (clock→target), but bidirectional regulation between clocks and targets is well-documented. Without causal tests (e.g., perturbation studies, Granger causality), the observed correlations remain associative rather than directional.

9. **Cross-tissue independence not formally modeled:** The 12 GSE54650 mouse tissues share experimental pipeline and animal cohort. The cross-tissue consensus FPR estimates implicitly assume independence, but actual effective sample size may be lower due to shared technical variation.

10. **Genotype-induced composition shifts:** In organoid analyses (GSE157357), bulk profiles may conflate intrinsic rhythmic dynamics with shifts in proportions of stem, progenitor, and differentiated cells between genotypes. This represents a potential confounder that cannot be resolved without cell-type deconvolution or single-cell data.

11. **Observational nature:** PAR(2) identifies correlational patterns between clock phase and target gene dynamics. Causal claims require experimental validation through genetic or pharmacological perturbation of clock components.

12. **Cancer model generalization:** The near-circadian eigenperiod (∼22.7h) and reduced stability (42%) reported for "cancer models" derive primarily from a single context: MYC-ON human neuroblastoma (GSE221103). While MYC-driven circadian disruption is well-established in neuroblastoma, the specific quantitative parameters may be context-dependent. Additional cancer types and models should be tested to establish generalizability.

13. **Golden-ratio analysis is exploratory:** The golden-ratio enrichment analysis is highly tissue-specific and depends on strict stability filtering. This analysis is orthogonal to the main PAR(2) findings and should be considered hypothesis-generating; see Supplementary Section S4 for details.

14. **Detection power is low (∼5%):** Simulation stress-testing indicates that even genuine ϕ-like AR(2) processes are detected at only ∼4.9% power under realistic noise and sampling conditions. This means we likely under-detect ϕ-like dynamics; absence of a significant call does not imply absence of effect. The observed 48× enrichment in specific tissues therefore represents a lower bound on true biological prevalence.

15. **In vivo tissue versus in vitro cell line comparison:** A potential confounder in the eigenperiod analysis is that healthy samples derive from mouse tissues (in vivo) while cancer samples derive from human neuroblastoma cell lines (in vitro). Cultured cells lack systemic entrainment cues present in live animals, which could contribute to the observed eigenperiod differences independent of malignancy status. We address this limitation using intestinal organoid data (GSE157357) as a genetic internal control where tissue background is matched; however, this does not substitute for the gold-standard comparison of patient-matched tumor versus adjacent normal tissue.

16. **Organoid validation supports clock-target hierarchy hypothesis:** Fresh AR(2) analysis of the GSE157357 intestinal organoid dataset (with proper mean-centering) reveals that healthy wild-type organoids (WT-WT) exhibit strong clock-target separation: clock genes show $|\lambda| = 0.72$ versus target genes $|\lambda| = 0.33$ (gap = +0.39), even larger than the in vivo reference. Critically, APC-mutant organoids modeling cancer show *reversed* dynamics: clock $|\lambda| = 0.53$ versus target $|\lambda| = 0.65$ (gap = −0.12), consistent with the hypothesis that oncogenic transformation disrupts the circadian-proliferation hierarchy. BMAL1-knockout organoids also show convergence (gap = −0.08), demonstrating that clock gene disruption alone can collapse this separation. These organoid controls, where genetic background is matched but genotype differs, provide mechanistic support for the clock-target hierarchy hypothesis and its disruption in the disease models studied here. We use "clock-target hierarchy" as a descriptive label for the observed eigenvalue separation, not as a claim about a universal biological mechanism.

4.9 Addressing the In Vivo versus In Vitro Confounder

The comparison between in vivo mouse tissues and in vitro human cancer cell lines represents a potential confounding factor that merits explicit discussion. However, several lines of evidence suggest the eigenperiod differences reflect genuine cancer biology rather than culture artifacts:

First, the intestinal organoid dataset (GSE157357) provides a controlled within-system comparison. Critically, all organoid conditions are derived from the *same tissue background* (mouse intestinal epithelium) and are non-malignant—the APC-KO organoids model pre-neoplastic Wnt-driven hyperproliferation, not established cancer. Wild-type organoids (APC-WT/BMAL-WT) and cancer-model organoids (APC-KO) are cultured under identical conditions, yet APC-KO organoids show 71% stability compared to WT organoids, consistent with the pattern observed in the neuroblastoma comparison. This internal control, where culture conditions and tissue background are matched but genotype differs, isolates the effect of oncogenic pathway activation from tissue context differences.

Second, the BMAL1-knockout organoids (APC-WT/BMAL-KO) show intermediate stability (68%), demonstrating that clock gene disruption alone—independent of oncogenic transformation—can alter dynamical stability. This provides mechanistic support for the clock-cancer connection.

Third, even if some portion of the eigenperiod shift reflects in vitro culture, the biological implication remains relevant: cancer cells in vivo are also partially "decoupled" from systemic circadian entrainment due to tumor microenvironment alterations, hypoxia, and metabolic reprogramming [14]. The in vitro phenotype may therefore model the in vivo tumor state.

Future studies should apply PAR(2) to circadian-resolved transcriptomics from patient-matched tumor and adjacent normal tissue samples to definitively resolve this question. **Until**

such data are available, the eigenperiod comparisons between healthy mouse tissues and human cancer cell lines should be interpreted as hypothesis-generating for cancer biology, not definitive. The organoid controls provide mechanistic support, but the ultimate test of clinical relevance requires true human tumor circadian time-courses.

4.10   Human Circadian Disruption Validation

To test whether the clock–target eigenvalue hierarchy generalises beyond mouse tissues and in vitro models, we applied AR(2) analysis to three independent human circadian disruption datasets from NCBI GEO, spanning forced desynchrony (GSE48113, $n = 22$ subjects, aligned vs. misaligned conditions [62]), acute sleep restriction (GSE39445, $n = 26$ subjects, sufficient vs. restricted sleep [63]), and real-world shift work (GSE122541, day vs. night shift nurses [64]). Across all six conditions, the clock $>$ target hierarchy was preserved: gaps ranged from $+0.030$ (night-shift nurses) to $+0.151$ (sufficient sleep), with circadian disruption consistently narrowing the gap but never inverting it. Combined with the mouse tissue and organoid results, this extends the validated panel to 14/14 datasets across 4 species—*Mus musculus*, *Homo sapiens*, *Papio anubis*, and *Arabidopsis thaliana*—with 100% hierarchy preservation. Notably, the disruption-induced gap compression—where misalignment, sleep loss, or shift work reduces but does not eliminate the clock–target separation—is consistent with the Gearbox Hypothesis prediction that the hierarchy reflects an intrinsic dynamical property of circadian gene networks rather than an artifact of experimental design.

4.11   Convergent Methodological Validation

The clock–target hierarchy reported here does not depend solely on the AR(2) framework. Two independent analytical routes, implemented on the same datasets, reach concordant conclusions. First, Granger causality testing confirms that clock genes significantly predict future target gene expression more strongly than targets predict clocks ($p < 0.05$, F-test), establishing directional information flow consistent with the eigenvalue hierarchy. Second, STRING protein–protein interaction network analysis shows that genes with high AR(2) eigenvalue modulus (stable oscillators) are disproportionately network hubs—high-degree, high-betweenness nodes—while low-eigenvalue genes occupy peripheral positions. This convergence across autoregressive, causal-inference, and network-topological lenses supports the interpretation that the hierarchy reflects a genuine dynamical property of circadian gene networks rather than a methodological artifact. Additional independent routes—including state-space eigenmodes, critical slowing down metrics [50], and frequency-domain transfer functions—predict equivalent conclusions and represent natural extensions for future work.

4.12   Clinical and Translational Implications

If validated experimentally, these findings have potential clinical applications:

1. **Exploratory diagnostic metric:** Eigenperiod could serve as an exploratory systems-level metric for circadian dysregulation in cancer, pending prospective validation on independent cohorts. A shift from ultradian ($<$12h) to near-circadian ($>$20h) eigenperiod may indicate oncogenic transformation or progression, though this hypothesis requires experimental confirmation.

2. **Chronotherapy optimization:** Understanding which gene pairs show phase-dependent gating could inform optimal timing of chemotherapy. If Pparg expression is gated by clock phase in certain cancers, timing of PPARγ-targeted therapies may influence efficacy.

31

3. **Circadian restoration:** The reduced stability in cancer suggests that therapeutic strategies aimed at restoring circadian rhythmicity (e.g., timed light exposure, melatonin supplementation, pharmacological clock modulators) may have additional benefits by restoring stable gene expression dynamics [32].

4.13 Future Directions

Several directions merit further investigation:

1. **Experimental validation of Pparg gating:** ChIP-seq for E-box and RRE elements in the Pparg promoter across the circadian cycle in MYC-overexpressing cells would directly test the predicted phase-dependent transcriptional regulation.

2. **Multi-omics integration:** Combining transcriptomic PAR(2) analysis with proteomics could assess whether mRNA-level gating translates to protein-level dynamics, or whether post-transcriptional regulation decouples these layers.

3. **Patient-derived samples:** Applying PAR(2) to circadian-resolved transcriptomics from patient tumor samples versus matched normal tissue would test clinical relevance.

4. **Therapeutic response prediction:** Retrospective analysis of chronotherapy trials could test whether baseline eigenperiod predicts response to circadian-timed treatment.

4.14 Validation Studies

To establish that the observed patterns reflect genuine biological signal rather than methodological artefact, we performed three complementary validation analyses (see Supplementary Materials):

Simulation stress-test: We simulated 360,000 synthetic time series (100 seeds $\times$ 3,600 simulations per seed) representing AR(1), AR(2)-ϕ, and AR(2)-non-ϕ archetypes with realistic noise and sampling. Using seeded pseudo-random number generators for reproducibility, the PAR(2) engine showed low false discovery rates: approximately 1–3% combined FDR (range: 1.5–3.0% across 100 seeds). This confirms that ϕ-like classifications are unlikely to arise from systematic overcalling, and that results are robust across random initializations.

Negative control panel: We analysed 40 randomly selected genes (excluding all clock, DDR, and Wnt genes) across 12 GSE54650 tissues using 25 random seeds for sensitivity analysis. Control genes showed only 1.8% $\pm$ 1.0% ϕ-rate (95% CI: 1.5–2.2%), compared to 100% in the clock/DDR panel for neural tissues—a 48-fold enrichment. This confirms that ϕ-enrichment is gene-panel specific and robust across different random gene selections.

Standard rhythm tool comparison: Cosinor analysis (similar to JTK_CYCLE/ARSER) on GSE157357 organoids showed only 1.75$\times$ discrimination between BMAL1-WT and BMAL1-KO conditions. PAR(2) gating analysis achieved 6.50$\times$ discrimination, demonstrating that phase-coupling analysis provides substantially stronger condition separation than standard rhythmicity detection.

Formal falsifiable predictions and their supporting validation statistics are summarised in Supplementary Note S3 (PAR2_FALSIFIABLE_PREDICTIONS.tex). These include conditions under which ϕ-like enrichment should be lost, datasets where LOUD vs SILENT regimes should invert, and signatures that would indicate model failure.

4.15 Robustness and Falsifiability

We explicitly state conditions under which the PAR(2) framework would require revision:

1. **AR(2) order**: The framework assumes second-order autoregressive dynamics. If AR(1) consistently outperformed AR(2) across healthy tissues, the theoretical basis would need reconsideration. Current status: Boman C-P-D ODE validation shows $\Delta AIC > +300$ favoring AR(2) in normal tissue; PACF lag-2 is significant ($|r| > 0.8$).

2. **Eigenvalue specificity**: If negative control genes showed clustering patterns indistinguishable from the clock/cancer panel, the observed enrichment would be artifactual. Current status: Control genes show 1.8% rate versus 100% in clock panel (48-fold difference).

3. **Disease separation**: If healthy tissues routinely showed eigenvalue drift toward $|\lambda| \to 1.0$, or cancer models showed stable dynamics ($|\lambda| \approx 0.537$), the clock-target eigenvalue difference hypothesis would not hold. Current status: Jan 2026 audit confirmed disease conditions show target genes ($|\lambda| = 0.705$) exceeding clock genes ($|\lambda| = 0.619$), consistent with a "convergence" pattern across APC-knockout organoids, MYC-ON neuroblastoma, and Boman adenoma simulations—though these findings require independent validation.

4. **Cross-tissue replication**: If significant pairs failed to replicate across independent tissues, single-tissue findings would remain unreliable. Current status: 129 pairs significant in 3+ tissues; FDR reduces from ~16% to ~2% with consensus requirement.

These criteria have not been met across 721 analyses in 72 biological contexts. However, we note that absence of falsification does not constitute proof; the framework remains a working hypothesis pending experimental validation and independent replication with patient-derived samples.

Recent mechanistic studies provide independent support: shared kinases (CK1δ, GSK3, AMPK) link circadian components to Wnt/Hippo signaling [38], and PER proteins suppress cancer stem cell properties via the Wnt/β-catenin pathway [39, 40]. The concordance between PAR(2) observations (Clock→Target Granger causality, eigenvalue drift in cancer) and these independently-derived mechanisms is consistent with the framework capturing genuine biological relationships, though correlation with published findings does not establish causation.

5 Conclusions

The PAR(2) Discovery Engine provides a permutation- and replication-validated framework for identifying candidate circadian gating relationships in cancer. Across 28,138 gene pairs tested in 22 datasets, we identified 2,697 Bonferroni-significant pairs (9.6%) and 33 FDR-significant pairs (0.1%). While individual pair-level findings require experimental validation due to moderate false positive rates (~16% single-tissue, reduced to approximately 1–5% with 3+ tissue consensus), the systems-level eigenperiod structure shows consistent separation: healthy tissues exhibit 7–13 hour ultradian dynamics versus 22–23 hour near-circadian periods in cancer (MYC-ON neuroblastoma). This separation was robust across period assumptions (T∈{20–28}h), though independent cohort validation and cross-dataset harmonization remain necessary before clinical applications can be considered.

The identification of Pparg as the only FDR-significant target (32 pairs in MYC-ON neuroblastoma, all with $f^2 = 10.86$) provides a prioritized cancer-context target. The identification of Wee1 as the exclusive highest-confidence tier target (gated by all 8 clock genes across 4–6 tissues each, average $f^2 = 2.36$) provides the top computational candidate for experimental validation of circadian-cell cycle coupling. More broadly, eigenperiod may serve as a hypothesis-generating systems-level metric for circadian dysregulation in cancer, with potential applications in diagnosis, prognosis, and chronotherapy optimization pending independent cohort validation. However, we emphasize that PAR(2) is a descriptive framework that improves in-sample explanatory power

but does not robustly improve out-of-sample prediction; these constraints limit its current utility for clinical forecasting applications.

Data Availability

Funding

Conflicts of Interest

The PAR(2) methodology is subject to a pending UK patent application (priority date established prior to public disclosure). The author declares no other conflicts of interest.

Author Contributions

M.W.: Conceptualization, Methodology, Software, Validation, Formal Analysis, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization.

Acknowledgments

References

[1] Bass J, Takahashi JS. Circadian integration of metabolism and energetics. *Science*. 2010;330(6009):1349-1354.

[2] Takahashi JS. Transcriptional architecture of the mammalian circadian clock. *Nature Reviews Genetics*. 2017;18(3):164-179.

[3] Reppert SM, Weaver DR. Coordination of circadian timing in mammals. *Nature*. 2002;418(6901):935-941.

[4] Lowrey PL, Takahashi JS. Genetics of circadian rhythms in mammalian model organisms. *Advances in Genetics*. 2011;74:175-230.

[5] Gekakis N, Staknis D, Nguyen HB, et al. Role of the CLOCK protein in the mammalian circadian mechanism. *Science*. 1998;280(5369):1564-1569.

[6] Kume K, Zylka MJ, Sriram S, et al. mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell*. 1999;98(2):193-205.

[7] Straif K, Baan R, Grosse Y, et al. Carcinogenicity of shift-work, painting, and fire-fighting. *The Lancet Oncology*. 2007;8(12):1065-1066.

[8] Schernhammer ES, Laden F, Speizer FE, et al. Night-shift work and risk of colorectal cancer in the Nurses' Health Study. *Journal of the National Cancer Institute*. 2003;95(11):825-828.

[9] Fu L, Pelicano H, Liu J, et al. The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo. *Cell*. 2002;111(1):41-50.

[10] Lee S, Donehower LA, Herron AJ, et al. Disrupting circadian homeostasis of sympathetic signaling promotes tumor development in mice. *PLoS ONE*. 2010;5(6):e10995.

[11] Lévi F, Okyar A, Dulong S, et al. Circadian timing in cancer treatments. *Annual Review of Pharmacology and Toxicology*. 2010;50:377-421.

[12] Matsuo T, Yamaguchi S, Mitsui S, et al. Control mechanism of the circadian clock for timing of cell division in vivo. *Science*. 2003;302(5643):255-259.

[13] Gery S, Komatsu N, Baldjyan L, et al. The circadian gene Per1 plays an important role in cell growth and DNA damage control in human cancer cells. *Molecular Cell*. 2006;22(3):375-382.

[14] Masri S, Sassone-Corsi P. The emerging link between cancer, metabolism, and circadian rhythms. *Nature Medicine*. 2018;24(12):1795-1803.

[15] Zhang R, Lahens NF, Ballance HI, et al. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proceedings of the National Academy of Sciences*. 2014;111(45):16219-16224.

[16] Zhang R, Lahens NF, Ballance HI, et al. A circadian gene expression atlas in mammals: implications for biology and medicine. *PNAS*. 2014;111(45):16219-16224.

[17] Mure LS, Le HD, Benegiamo G, et al. Diurnal transcriptome atlas of a primate across major neural and peripheral tissues. *Science*. 2018;359(6381):eaao0318.

[18] Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*. 2011;12(1):56-68.

[19] Koike N, Yoo SH, Huang HC, et al. Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science*. 2012;338(6105):349-354.

[20] Menet JS, Rodriguez J, Abruzzi KC, Rosbash M. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. *eLife*. 2012;1:e00011.

[21] Cornelissen G. Cosinor-based rhythmometry. *Theoretical Biology and Medical Modelling*. 2014;11(1):16.

[22] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*. 1995;57(1):289-300.

[23] Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

[24] Wang N, Yang G, Jia Z, et al. Vascular PPARγ controls circadian variation in blood pressure and heart rate through Bmal1. *Cell Metabolism.* 2008;8(6):482-491.

[25] Yang X, Downes M, Yu RT, et al. Nuclear receptor expression links the circadian clock to metabolism. *Cell.* 2006;126(4):801-810.

[26] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646-674.

[27] Pavlova NN, Thompson CB. The emerging hallmarks of cancer metabolism. *Cell Metabolism.* 2016;23(1):27-47.

[28] Geyfman M, Kumar V, Liu Q, et al. Brain and muscle Arnt-like protein-1 (BMAL1) controls circadian cell proliferation and susceptibility to UVB-induced DNA damage in the epidermis. *PNAS.* 2012;109(29):11758-11763.

[29] Ye Y, Xiang Y, Ozguc FM, et al. The genomic landscape and pharmacogenomic interactions of clock genes in cancer chronotherapy. *Cell Systems.* 2018;6(3):314-328.

[30] Michalik L, Desvergne B, Wahli W. Peroxisome-proliferator-activated receptors and cancers: complex stories. *Nature Reviews Cancer.* 2004;4(1):61-70.

[31] Droin C, Paquet ER, Naef F. Low-dimensional dynamics of two coupled biological oscillators. *Nature Physics.* 2019;15(10):1086-1094.

[32] Sulli G, Manoogian ENC, Taub PR, Panda S. Training the circadian clock, clocking the drugs, and drugging the clock to prevent, manage, and treat chronic diseases. *Trends in Pharmacological Sciences.* 2018;39(9):812-827.

[33] Shmueli G. To explain or to predict? *Statistical Science.* 2010;25(3):289-310.

[34] Boman BM. How Does Multicellular Life Happen? Modeling Fibonacci Patterns in Biological Tissues. *The Fibonacci Quarterly.* 2025;September issue.

[35] Boman BM, Dinh TN, Decker K, Emerick B, Modarai SR, Opdenaker LM, Fields JZ, Raymond C, Schleiniger G. Dynamic organization of cells in colonic epithelium is encoded by five biological rules. *Biology of the Cell.* 2025;117(7):e70017. doi:10.1111/boc.70017

[36] Nguyen AL, Lausten MA, Boman BM. The Colonic Crypt: Cellular Dynamics and Signaling Pathways in Homeostasis and Cancer. *Cells.* 2025;14(18):1428. doi:10.3390/cells14181428

[37] Boman RM, Schleiniger G, Raymond C, Palazzo J, Shehab A, Boman BM. A Tissue Renewal-Based Mechanism Drives Colon Tumorigenesis. *Cancers.* 2026;18(1):44. doi:10.3390/cancers18010044

[38] Liu J, Jiang Z, Zha J, Lin Q, He W. Crosstalk between the circadian clock, intestinal stem cell niche, and epithelial cell fate decision. *Genes & Diseases.* 2025;12(6):101650. doi:10.1016/j.gendis.2025.101650

[39] The role of circadian rhythm regulator PERs in oxidative stress, immunity, and cancer development. *Cell Communication and Signaling.* 2025;23:30. doi:10.1186/s12964-025-02040-2

[40] Li Q, Xia D, Wang Z, Liu B, Zhang J, Peng P, Tang Q, Dong J, Guo J, Kuang D, Chen W, Mao J, Li Q, Chen X. Circadian Rhythm Gene PER3 Negatively Regulates Stemness of Prostate Cancer Stem Cells via WNT/β-Catenin Signaling in Tumor Microenvironment. *Frontiers in Cell and Developmental Biology.* 2021;9:656981. doi:10.3389/fcell.2021.656981

[41] Core Molecular Clock Factors Regulate Osteosarcoma Stem Cell Survival and Behavior via CSC/EMT Pathways and Lipid Droplet Biogenesis. *Cells*. 2025;14(7):517. doi:10.3390/cells14070517

[42] Circadian rhythms in stem cells and their therapeutic potential. *Stem Cell Research & Therapy*. 2025;16:87. doi:10.1186/s13287-025-04178-9

[43] Wold H. A Study in the Analysis of Stationary Time Series. Stockholm: Almqvist & Wiksell; 1938.

[44] Smallbone K, Corfe BM. A mathematical model of the colon crypt capturing compositional dynamic interactions between cell types. *International Journal of Experimental Pathology*. 2014;95(1):1-7. doi:10.1111/iep.12062

[45] van Leeuwen IMM, Byrne HM, Jensen OE, King JR. Elucidating the interactions between the adhesive and transcriptional functions of β-catenin in normal and cancerous cells. *Journal of Theoretical Biology*. 2007;247(1):77-102. doi:10.1016/j.jtbi.2007.01.019

[46] Box GEP, Jenkins GM, Reinsel GC. Time Series Analysis: Forecasting and Control. 3rd ed. Englewood Cliffs, NJ: Prentice Hall; 1994.

[47] Hurd HL, Miamee A. *Periodically Correlated Random Sequences: Spectral Theory and Practice*. Hoboken, NJ: Wiley; 2007.

[48] Paap R, Franses PH. Shrinkage estimators for periodic autoregressions. *Journal of Econometrics*. 2025;247:103888. doi:10.1016/j.jeconom.2024.103888

[49] López-de-Lacalle J. Periodic Autoregressive Time Series Models in R: The partsm Package. *University of the Basque Country Working Papers*. 2005.

[50] Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53-59. doi:10.1038/nature08227

[51] Chen L, Liu R, Liu ZP, Li M, Aihara K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Scientific Reports*. 2012;2:342. doi:10.1038/srep00342

[52] Zaidi SK, Young DW, Montecino MA, et al. Mitotic bookmarking of genes: a novel dimension to epigenetic control. *Nature Reviews Genetics*. 2010;11(8):583-589. doi:10.1038/nrg2827

[53] Zhu F, Farnung L, Kaasinen E, et al. Mitotic bookmarking by SWI/SNF subunits. *Nature*. 2023;618(7967):580-587. doi:10.1038/s41586-023-06085-6

[54] Puri S, Kumar V. Differentiation is accompanied by a progressive loss in transcriptional memory. *BMC Biology*. 2024;22:67. doi:10.1186/s12915-024-01846-9

[55] Hughes ME, Hogenesch JB, Kornacker K. JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *Journal of Biological Rhythms*. 2010;25(5):372-380. doi:10.1177/0748730410379711

[56] Thaben PF, Westermark PO. Detecting rhythms in time series with RAIN. *Journal of Biological Rhythms*. 2014;29(6):391-400. doi:10.1177/0748730414553029

[57] Wu G, Anafi RC, Hughes ME, Kornacker K, Hogenesch JB. MetaCycle: an integrated R package to evaluate periodicity in large scale data. *Bioinformatics*. 2016;32(21):3351-3353. doi:10.1093/bioinformatics/btw405

[58] Zhang Y, Chen W. DCPR: a deep learning framework for circadian phase reconstruction. *BMC Bioinformatics*. 2025;26:45. doi:10.1186/s12859-025-06363-2

[59] Liu H, Yang J. CGRF: Hybrid framework for inferring circadian gene regulatory networks. *BMC Bioinformatics*. 2025;26:89. doi:10.1186/s12859-023-05458-y

[60] Hughey JJ, Hastie T, Butte AJ. tauFisher predicts circadian time from a single sample of bulk and single-cell pseudobulk transcriptomic data. *Nature Communications*. 2024;15:3890. doi:10.1038/s41467-024-48041-6

[61] Leloup JC, Goldbeter A. Toward a detailed computational model for the mammalian circadian clock. *Proceedings of the National Academy of Sciences*. 2003;100(12):7051-7056. doi:10.1073/pnas.1132112100

[62] Archer SN, Laing EE, Möller-Levet CS, van der Veen DJ, Bucca G, Lazar AS, Santhi N, Slak A, Kabiljo R, von Schantz M, Smith CP, Dijk DJ. Mistimed sleep disrupts circadian regulation of the human transcriptome. *Proceedings of the National Academy of Sciences*. 2014;111(6):E682-E691. doi:10.1073/pnas.1316335111

[63] Möller-Levet CS, Archer SN, Bucca G, Laing EE, Slak A, Kabiljo R, Lo JCY, Santhi N, von Schantz M, Smith CP, Dijk DJ. Effects of insufficient sleep on circadian rhythmicity and expression amplitude of the human blood transcriptome. *Proceedings of the National Academy of Sciences*. 2013;110(12):E1132-E1141. doi:10.1073/pnas.1217154110

[64] Gamble KL, Berry R, Frank SJ, Young ME. Shift work disrupts circadian regulation of the transcriptome in hospital nurses. *Journal of Biological Rhythms*. 2019;34(2):167-177. doi:10.1177/0748730419826694