

Cancer as Boundary Logic Failure

A Computational and Information-Theoretic Framework
for the Internal Self-Referential Malignancy

Moses Rahnama

Mina Analytics

moses@minaanalytics.com

February 2026

Abstract

We present a boundary-first theory of cancer. Mutation is required for adaptation, but the same variation machinery can enter a self-referential loop when organism-level termination predicates fail. Cancer is therefore a *necessary bug* of innovation, not an external corruption of an otherwise perfect system.

Using computability analogies, systems biology, and thermodynamic information constraints, we model malignancy as loss of cross-scale observability. Cells become insufficiently distinguishable for reliable organism-level control. We formalize this with a control-allocation model and with the conditional record-formation lower bound $\langle Q_{\text{rec}} \rangle \geq k_B T \ln 2 \cdot I(X; Y)$, where $I(X; Y)$ is classical mutual information in a durable decision record. This is a lower bound and organizing principle, not a direct estimate of full tumor metabolism.

The framework explains how reduced differentiation investment, proliferative burden, and tumor-immune inflammatory signaling produce systemic metabolic stress, including cachexia. It unifies these findings under one failure mode: loss of boundary-enforced termination. We derive testable qualitative predictions on tissue coherence and propose “forced distinction” as a design principle spanning differentiation therapy, checkpoint biology, bioelectric recoupling, and chronotherapy. The paper offers seven falsifiable experiments, with two straightforward first tests: correlating PD-L1/MHC-I status with tissue-coherence proxies (for example gap-junction density and circadian coherence), and testing whether peri-tumoral circadian-coherence loss is detectable before overt malignancy. We connect the model to PD-1/PD-L1 evasion, gap-junction decoupling, signaling-entropy elevation (SCENT/mRNAsi), Warburg-associated metabolic autonomy, and circadian desynchronization. Validation across six TCGA cohorts ($n = 3,611$ tumors; melanoma, lung adenocarcinoma, breast, colon, head-neck, lung squamous) reveals significant circadian–checkpoint coupling in all six cancer types (FDR $q < 0.05$; ρ range -0.125 to -0.381) and refines the hypothesis: boundary failure in immune-hot tumors operates through *active masking* (checkpoint engagement with intact clock) rather than temporal decoherence alone. Active Masking tumors show

significantly better overall survival than Decoherence tumors in melanoma ($p = 0.0008$) and lung adenocarcinoma ($p = 0.024$), suggesting that the boundary-failure *mode* carries direct prognostic and therapeutic implications.

Contents

1	Introduction	5
2	The Necessary Bug: Mutation as the Price of Innovation	6
3	Positioning Within Cancer Theory	7
4	Boundary Failure as Loss of Cross-Scale Observability	10
4.1	Relational Identity and Organism-Level Control	10
4.2	The Identity Collapse	10
4.3	Distinctness, Information, and Energetic Constraints	10
5	A Structural Analogy to Recursive Self-Duplication	11
5.1	The Computational Pattern	11
5.2	The Biological Reading (Analogy, Not Identity)	12
5.3	Why the Analogy Is Useful	12
6	A Control-Budget Formulation of Growth-Stability Tradeoff	13
7	Temporal Coordination and Loss of Synchrony	14
7.1	Circadian Disruption and Cancer	14
7.2	Loss of Temporal Coordination as Boundary Failure	14
7.3	Perineural Invasion Reinterpreted	14
8	Immune Evasion as Loss of Distinguishability	15
8.1	The PD-1/PD-L1 Checkpoint Axis	15
8.2	Secondary Boundary Failures	15
8.3	Gap Junctions and Antigen Broadcast	15
9	Thermodynamic Perspective: Cost Externalization	16
9.1	The Warburg Effect as Metabolic Decoupling	16
9.2	Cachexia as Systemic Cost Payment	16
9.3	Entropy Production (Heuristic)	16
10	Toy Formalism: Tissue as a Labeled Graph	17
10.1	State Variables	17
10.2	A Coherence Metric	17
10.3	Cancer Transform Operator	18
11	Toward Measurable Proxies	18
12	Therapeutic Design Principle: Restore Observability and Enforce Termination	19
12.1	Force Distinction	19
12.2	Restore Termination Predicates	20
12.3	Re-couple to Tissue Constraints	20

12.4 Constrain Resource Externalization	20
13 Clinical Correlations and Testable Hypotheses	20
14 Empirical Validation: TCGA Cohort Analysis	21
14.1 Circadian Coherence and PD-L1	22
14.2 PD-L1 and MHC Class I Co-regulation	23
14.3 Gap Junction Expression	23
14.4 Summary of Key Correlations	23
14.5 Tumor vs. Matched Normal Tissue	27
14.6 Active Masking vs. Temporal Decoherence Classification	28
14.7 Survival Analysis	29
14.8 Stage-Stratified Analysis	30
14.9 Implications for the Framework	31
15 Simulation	32
15.1 Update Rules	32
15.2 What the Simulation Demonstrates	33
16 Limitations	38
17 Conclusion	39
18 Future Research Directions	40
A Code Availability and Reproducibility Assets	45

1 Introduction

Cancer research has often focused on mutation accumulation and pathway deregulation [2, 3]. This work has identified key oncogenes and tumor suppressors, but it can frame cancer as a set of local errors rather than a systemic failure of biological control. We propose a boundary framework: cancer is a logic failure in multicellular boundary management, with a shift from cooperative coordination to self-referential isolation.

We define this state as a **boundary logic failure**: a condition where a local system (the cell) loses informational coupling with the global structure (the organism) and enters a self-referential loop of unbounded proliferation. This framework unifies three distinct physical and mathematical domains into a single conceptual model of malignancy:

1. **Computability Theory:** We model uncontrolled proliferation as a structural analogy to the recursive successor (`rec_succ`) pattern, a process where the “step” instruction is duplicated in the recursive call without a strong external termination predicate.
2. **Thermodynamics and Information Theory:** We apply Landauer’s principle [8] interpretively to cellular division. Normal differentiation involves creating a new informational identity (different gene expression profile, different surface markers, different tissue role). Cancer cells can replicate physical structure while under-investing in informational distinction, accruing proliferative burden that contributes to systemic metabolic stress and cachexia [34].
3. **Systems Biology:** We reinterpret tumor suppressors as termination predicates in a biological decision process, and frame tissue coherence as a quantifiable metric of cross-scale coupling.

This paper models cancer as a predictable failure mode in systems capable of self-reference and innovation. Framing malignancy as an information problem yields qualitative definitions of tissue coherence and a therapeutic design principle based on “forced distinction”.

Scope and Non-Goals

This is not a molecular or clinical oncology model. It does not predict patient outcomes, infer driver mutations, or reproduce tumor heterogeneity from first principles. It is a failure-mode framework that links self-referential, growth-permissive dynamics to known biology and candidate therapeutic strategies.

Clinical Impact: Boundary-Failure Mode Predicts Survival

The framework is not purely theoretical. Analysis of $n = 3,611$ tumors across six TCGA cancer cohorts reveals that the *mode* of boundary failure carries direct prognostic significance. Figure 1 shows Kaplan-Meier survival curves for melanoma patients stratified by boundary-failure subtype: **Active Masking** tumors (high PD-L1, intact circadian clock) show significantly better overall survival than **Decoherence** tumors (low PD-L1, low antigen presentation) with $p = 0.0008$. This survival advantage replicates in lung adenocarcinoma ($p = 0.024$) and suggests different therapeutic strategies for each failure mode: Active

Masking tumors are visible to the immune system and candidates for checkpoint inhibitor therapy, while Decoherence tumors require restoration of immune visibility before checkpoint blockade can be effective. Full analysis in Section 14.

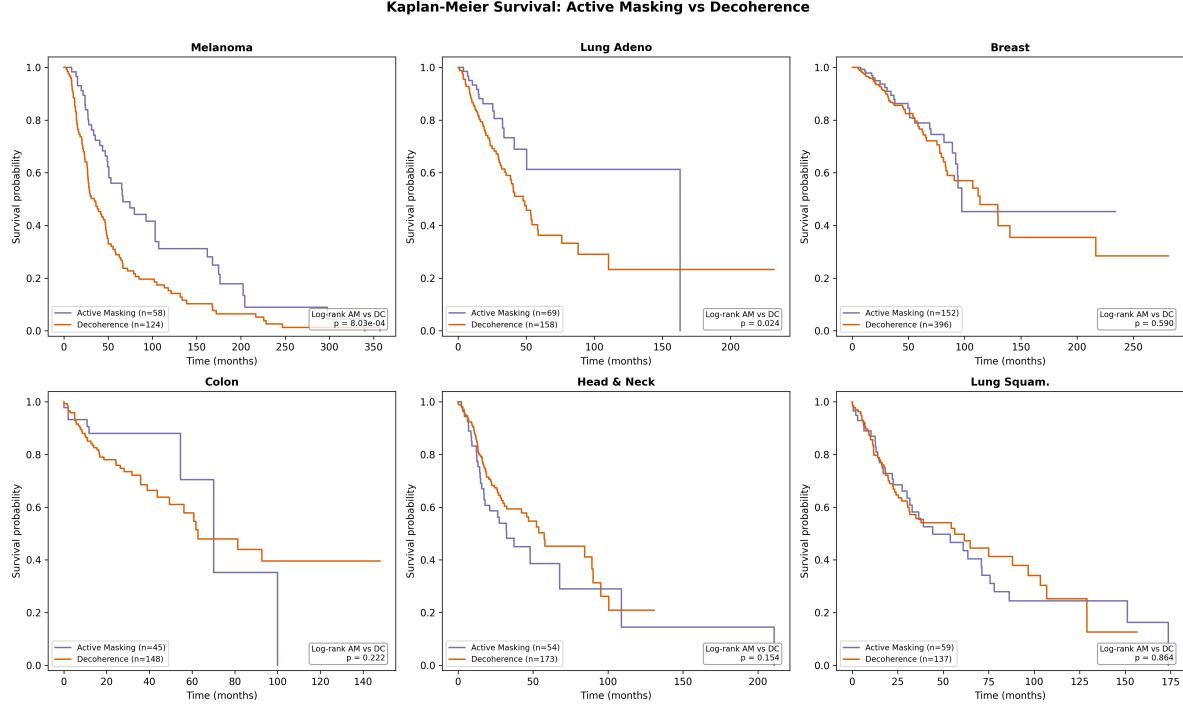


Figure 1: Boundary-failure mode predicts clinical outcomes. Kaplan-Meier survival curves by boundary-failure subtype in TCGA melanoma (SKCM, $n = 472$) and lung adenocarcinoma (LUAD, $n = 530$). Active Masking tumors (high PD-L1 + locked circadian clock) show significantly better overall survival than Decoherence tumors (low PD-L1 + low MHC-I) in both cancer types (log-rank $p = 0.0008$ for SKCM, $p = 0.024$ for LUAD). This demonstrates that the boundary-failure framework has direct clinical relevance: the *mode* of boundary failure matters more than the degree of disruption, with implications for immunotherapy response prediction and therapeutic stratification. Complete analysis in Section 14.

2 The Necessary Bug: Mutation as the Price of Innovation

The central claim of this manuscript is evolutionary before it is clinical: the mechanism that enables adaptation also creates the possibility of malignancy. A living lineage must generate variation to explore new phenotypes under changing environments. Without mutation (sequence variation, recombination, epigenetic drift, and related sources of novelty), evolution cannot “compute” the next viable state.

This creates a structural tension. Let z_t denote the local cellular program and m_t a

variation operator. Innovation requires iterative update:

$$z_{t+1} = F(z_t, m_t). \quad (1)$$

In multicellular organisms, this iteration is normally constrained by organism-level termination predicates (checkpoints, differentiation commitments, apoptosis, immune clearance). Healthy adaptation therefore follows:

$$\text{Variation} \rightarrow \text{Selection/Constraint} \rightarrow \text{Termination of unfit trajectories}. \quad (2)$$

Cancer appears when the update channel remains active but the termination channel is weakened or bypassed. The same innovation machinery then amplifies self-preserving local trajectories instead of organism-compatible ones. This is the “necessary bug”: mutation cannot be removed without destroying adaptation, so the therapeutic target is not variation itself but restoration of boundary observability and termination enforcement.

This interpretation is consistent with atavistic perspectives [5] while extending them: the system can rediscover ancestral proliferative programs and also evolve genuinely novel resistance strategies because the search loop remains active once boundary control fails.

3 Positioning Within Cancer Theory

A long oncology tradition treats cancer as breakdown of multicellular cooperation. Core perspectives include clonal evolution [4], hallmarks frameworks [2, 3], and multi-scale control failure (checkpoints, contact inhibition [17], immune recognition, and microenvironmental constraints). Tissue organization field theory (TOFT) [7] emphasizes tissue-level causation, showing that stromal disruption can drive aberrant epithelial behavior without requiring epithelial mutations.

Davies and Lineweaver [5] proposed an atavistic model: cancer reverts toward ancestral unicellular phenotypes. Phylostratigraphic evidence supports this shift, with tumors over-expressing older genes and suppressing genes of multicellular origin [6]. In this view, cancer often reactivates ancient programs that multicellular regulation usually constrains.

Our thesis is compatible with these views and uses a unifying language of *observability* and *cross-scale coupling*. Here, *boundary* means the operational separation between what a cell can do and what the organism permits. A healthy boundary is enforced by signals external to each cell that can override local replication programs.

The framework also connects to Levin’s bioelectric work [18, 19]. Cells use ion channels and gap junctions to maintain voltage gradients that encode collective identity. When this integration fails, cells can shift toward unicellular replicative goals. Experimental hyperpolarization can suppress tumor formation despite strong oncogenic drivers [19], which supports the boundary-control interpretation.

The atavistic model explains reactivation of proliferative drives, but critics note that tumors also show neomorphic traits, including drug resistance, that ancestral unicells did not exhibit. In our framework, this follows from loss of termination predicates: once boundary control weakens, the system enters an unbounded search loop. That loop can recover ancestral solutions and also discover new resistance strategies under modern constraints. Cancer is atavistic in origin but evolutionary in progression.

Preview: Empirical Validation Across Six Cancer Types

Before developing the formal framework, we preview the empirical foundation: validation across six TCGA cancer cohorts ($n = 3,611$ tumors) confirms the framework's core predictions. Figure 2 shows that circadian coherence loss correlates with PD-L1 expression in **all six cancer types** tested (100% replication rate, all FDR-corrected $q < 0.05$). This cross-cancer coupling between temporal boundary mechanisms (circadian clock) and immune boundary mechanisms (checkpoint engagement) supports the unifying boundary-failure interpretation. Detailed analysis, survival outcomes, and Active Masking vs. Decoherence stratification are presented in Section 14.

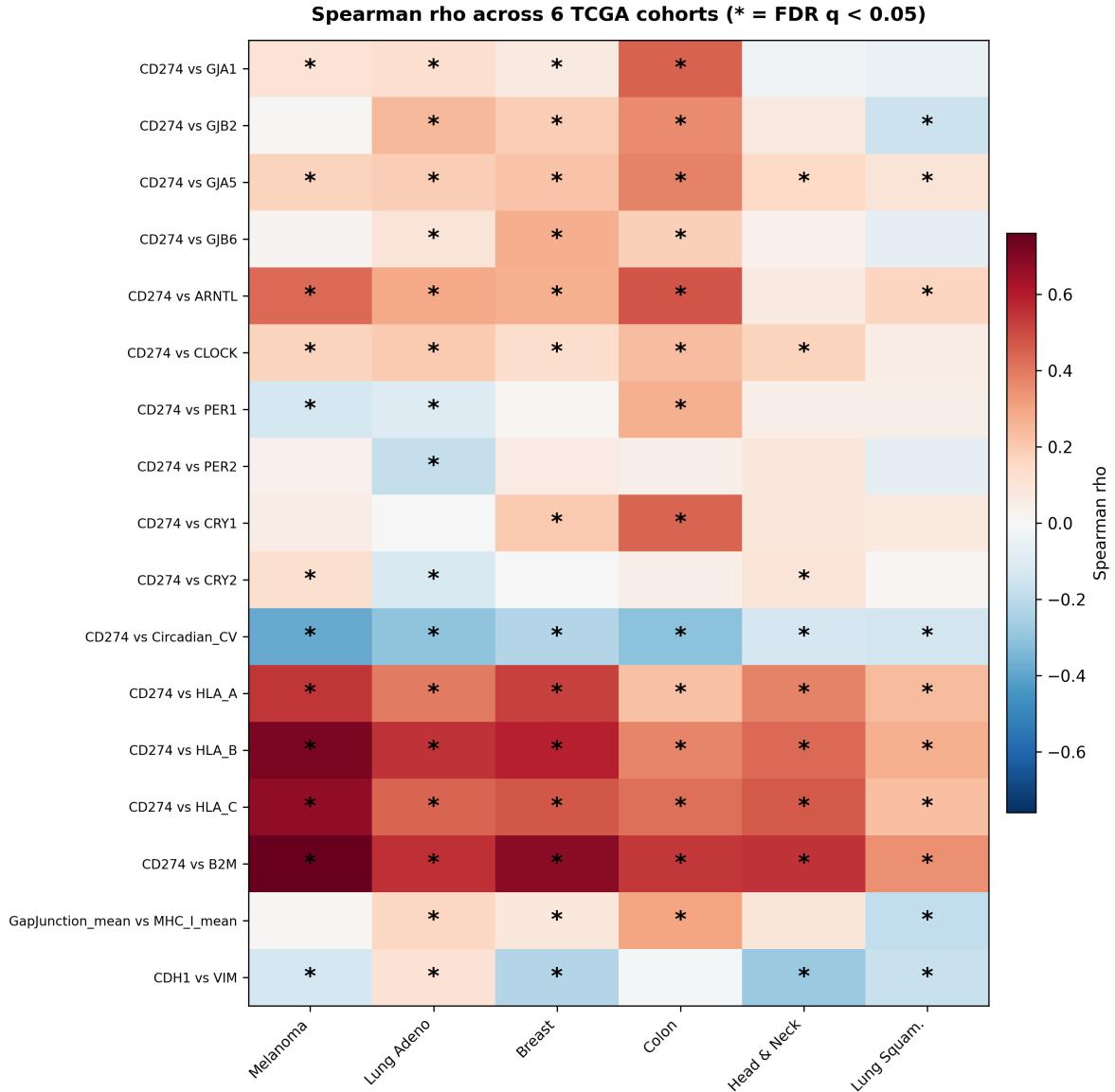


Figure 2: Cross-cancer validation of boundary-failure predictions. Spearman correlation heatmap across six TCGA cancer types ($n = 3,611$ tumors: melanoma, lung adenocarcinoma, breast, colon, head-neck, lung squamous). The top row shows circadian CV vs. PD-L1 (CD274): negative correlations replicate in all six cancer types (effect sizes $\rho = -0.125$ to -0.381 , all FDR $q < 0.05$), confirming that loss of temporal coherence couples with immune checkpoint engagement across cancer origins. This provides empirical support for the boundary-failure framework developed in the following sections. Complete analysis in Section ??.

4 Boundary Failure as Loss of Cross-Scale Observability

4.1 Relational Identity and Organism-Level Control

In a multicellular organism, a cell’s “identity” is both intrinsic (gene expression state) and relational: it is stabilized by continuous interaction with surrounding tissue and organism-level regulators. In simplified terms, a healthy cell behaves as if it continuously answers a question of the form:

Am I still operating within the organism’s allowed boundary conditions?

When the answer is “no” (or uncertain), the organism provides termination mechanisms: growth arrest, differentiation cues, senescence, apoptosis [14, 15], and immune-mediated clearance [29]. In this framework, when these mechanisms fail simultaneously, the cell reverts to its most primitive program: divide.

4.2 The Identity Collapse

We model this transition with weighted control allocation. Let $w_{\text{self}}(t)$ and $w_{\text{org}}(t)$ denote the relative influence of cell-autonomous programs and organism-level constraints on cellular decisions at time t , with:

$$w_{\text{self}}(t) + w_{\text{org}}(t) = 1, \quad 0 \leq w_{\text{self}}, w_{\text{org}} \leq 1. \quad (3)$$

Define an alignment index:

$$A(t) = w_{\text{org}}(t) - w_{\text{self}}(t), \quad (4)$$

where healthy tissue corresponds to $A(t) > 0$ (organism-dominant control).

Cancer represents a shift toward self-dominant control ($A(t) \leq 0$), with behavior approaching a closed local decision loop:

$$R_{\text{cell}}^{(t+1)} \approx f(R_{\text{cell}}^{(t)}), \quad \frac{\partial R_{\text{cell}}^{(t+1)}}{\partial u_{\text{org}}^{(t)}} \approx 0, \quad (5)$$

where u_{org} denotes organism-level regulatory input.

In this regime, organism-level signals are present but underweighted. The decision process becomes self-referential. This matches the atavistic view [5]: the cell shifts from multicellular cooperation (“I am part of the organism”) to unicellular survival (“I am the organism”).

4.3 Distinction, Information, and Energetic Constraints

The framework uses “distinction” as an organizing concept: for the organism to control a cell, the cell must be distinguishable in ways that matter for regulation (state, location, lineage, antigen presentation, and response to signals). Maintaining, transmitting, and acting on such distinctions has costs in physical systems [8, 9, 10].

To make this thermodynamic claim explicit without overreach, we use the conditional record-formation lower bound from the boundary-measurement framework [1]:

$$\langle Q_{\text{rec}} \rangle \geq k_B T \ln 2 \cdot I(X; Y), \quad (6)$$

where $I(X; Y)$ is classical mutual information between input label X and durable record Y . In biological terms, X can denote antigenic, transcriptional, and context signals, and Y can denote a stable organism-level decision record (tolerate/clear, differentiate/terminate). This is a lower bound on record formation, not a direct estimate of full tumor metabolism.

We adopt a conservative thermodynamic stance: cancer does not evade physical energy costs of replication and signaling. DNA replication consumes ATP; mitosis is energetically expensive. Rather, malignancy can reduce investment in identity-establishing information and degrade the semantic content of distinctions required for regulation. Specifically:

- **Normal differentiation:** The daughter cell acquires a new informational identity (different gene expression profile, different surface markers, different tissue role). This is a semantic distinction that enables immune recognition and tissue coordination.
- **Malignant duplication:** The daughter cell can remain *insufficiently distinguishable* from the parent for reliable organism-level control. The physical replication cost is paid, but investment in stable identity-bearing differentiation is reduced. Distinction signals can remain (e.g., neoantigens), yet effective visibility can fall below immune decision criteria when evasion programs (PD-L1 upregulation, MHC-I loss) suppress signal quality.

In later sections we use this to interpret metabolic reprogramming [31, 32] and cachexia [34, 35] via three coupled processes: reduced local differentiation investment, direct host resource burden, and tumor-immune inflammatory catabolism.

5 A Structural Analogy to Recursive Self-Duplication

5.1 The Computational Pattern

In a minimal rewrite calculus, a successor-recursion rule has the schematic form:

$$\text{rec}(b, s, \sigma(n)) \rightarrow F(s, \text{rec}(b, s, n)), \quad (7)$$

where b is a base case, s is a step argument, and σ is a successor constructor. In operator-only systems that support ordered computation, the step argument typically reappears in both the outer function application and the recursive call, creating a self-reference that defeats naive additive termination measures when s is unrestricted.

Minimal operator-only rewrite systems formalize this termination pressure and show strong normalization only for guarded fragments; unguarded settings contain the same kind of step-argument redistribution that resists “internal” termination proofs. In that system class, the two recursion rules are:

- $R_{\text{rec_zero}}: \text{rec}_\Delta(b, \text{void}) \rightarrow b$ (base case: halt)

- $R_{\text{rec_succ}}$: $\text{rec}_\Delta(b, s, \delta(n)) \rightarrow \text{app}(s, \text{rec}_\Delta(b, s, n))$ (recursive step)

The critical structure of $R_{\text{rec_succ}}$: the step argument s appears once on the left and is redistributed on the right, once in the outer application and once in the recursive call. When s is itself a recursive term, the system expands without bound.

5.2 The Biological Reading (Analogy, Not Identity)

We use `rec_succ` as a *structural analogy* for malignancy. A healthy tissue contains termination predicates: contact inhibition [17], checkpoints (p53 [14, 15]), and organism-level stop signals. A cancerous process, in contrast, behaves as if the base case has been disabled and the step rule “duplicate and recurse” has become unconditional.

This is a category-level structural analogy, not a claim of formal computational equivalence. Biological processes are stochastic, interruptible, and physically constrained. When termination predicates weaken, the default local dynamic is iterative duplication, and the burden of halting shifts to external regulators (immune surveillance, microenvironmental constraints, therapy).

Computational Domain	Biological Domain	Failure Mode
Function: <code>rec_succ</code>	Process: Mitosis	Unbounded recursion
Operation: $s \rightarrow F(s, \text{rec}(\dots))$	Operation: Cell \rightarrow (Cell, Cell')	Duplication without distinction
Base case: if proven: halt	Base case: Apoptosis / quiescence	Termination predicate lost
Termination check: Measure function	Termination check: p53, checkpoints	Observer cannot verify
Result: Infinite expansion	Result: Tumor growth	System destabilization
Resolution: Accept undecidability; halt	Resolution: Force distinction	Restore observability

Table 1: Structural analogy between `rec_succ` and cancer. This is an analogy about shared failure modes (duplication without termination), not a claim that tumors obey rewrite rules.

5.3 Why the Analogy Is Useful

The `rec_succ` framing clarifies three things that the standard mutation-accumulation model does not emphasize:

1. **Default behavior:** In the absence of active constraints, proliferation is the default, not an acquired capability. This follows the TOFT premise that proliferation is the default state of cells [7]. Mitogenic-induction frameworks place stronger weight on active growth signaling; in that language, the same claim is that once induction barriers are persistently lifted, proliferation becomes the de facto trajectory.

2. **Insufficient distinguishability as the core problem:** The immune system cannot reliably halt what it cannot distinguish above action criteria, just as a termination proof fails when recursive states cannot be separated well enough to certify halt conditions.
3. **The burden of proof shifts:** Halting is not the cell’s responsibility; it is the organism’s. When the organism loses the ability to verify termination (immune evasion, checkpoint loss), the default loop runs.

The evolutionary rationale for why this failure mode exists is developed in Section 2. Here, `rec_succ` is used only as a structural lens: once strong termination predicates are weakened, the duplicate-and-recurse dynamic dominates.

6 A Control-Budget Formulation of Growth-Stability Trade-off

To avoid pseudo-fundamental notation, we frame growth-stability tension as a bounded-control allocation problem. Let G denote proliferative throughput and S denote maintained tissue coherence. Let $C_G(G)$ and $C_S(S)$ be the regulatory/energetic control costs required to sustain those operating points. Feasible operation requires:

$$C_G(G) + C_S(S) \leq B_{\text{ctrl}}, \quad (8)$$

where B_{ctrl} is an effective organism-level control budget (system dependent, time dependent, and not universal).

This is a phenomenological constraint, not a fundamental physical law. Its practical meaning is straightforward: under fixed control resources, pushing sustained proliferative output upward leaves less margin for stability-enforcing processes (checkpoint integrity, differentiation maintenance, immune enforcement, temporal coupling), and vice versa.

Property	Normal Tissue Regime	Malignant Regime
Proliferative throughput (G)	Bounded/context-appropriate	Elevated/persistent
Coherence maintenance (S)	High	Degraded
Budget relation	$C_G + C_S \leq B_{\text{ctrl}}$ (maintained)	Effective violation via decoupling/overload
Termination predicates	Active (apoptosis, checkpoints, immune clearance)	Weakened or bypassed

Table 2: Phenomenological control-budget view of growth-stability tradeoff.

Cancer cells do not evade thermodynamics; they redistribute burden across scales. Local gains in autonomous proliferation are coupled to loss of global coordination and increased host-level energetic/inflammatory cost.

7 Temporal Coordination and Loss of Synchrony

Many regulatory signals in organisms are time-structured: circadian rhythms, hormonal cycles, metabolic oscillations, and immune rhythms. Healthy tissue can be characterized as *synchronized* to shared temporal references, while malignant regions behave as increasingly *asynchronous*.

7.1 Circadian Disruption and Cancer

The International Agency for Research on Cancer (IARC) has classified night shift work involving circadian disruption as a probable human carcinogen (Group 2A) [22]. Prospective cohort studies of nurses show elevated breast cancer risk after 20+ years of rotating night shifts ($RR = 1.79$; 95% CI 1.06–3.01) [23]. Finnish twin cohort studies show $HR = 1.58$ (95% CI 1.16–2.15) for breast cancer among night shift workers, with the critical factor being circadian disruption specifically, not shift work per se [24].

At the molecular scale, core clock components CLOCK and BMAL1 have tumor-suppressive roles: BMAL1 activates ATM signaling for G2/M arrest in colorectal cancer, and the CLOCK-BMAL1 heterodimer suppresses c-Myc-mediated Cyclin E activation in melanoma and colorectal models [25]. Cancer cells often dysregulate these programs. Genetic ablation of BMAL1 promotes lung tumorigenesis by increasing c-Myc activity and glutamine use [25]. Distant tumors can also rewire circadian metabolism in other organs; lung adenocarcinoma alters hepatic STAT3-Socs3 signaling without changing core liver clock genes [26].

7.2 Loss of Temporal Coordination as Boundary Failure

In our framework, temporal coordination is represented by explicit temporal coupling between cells and systemic regulators. “Temporal decoherence” means loss of synchrony in these couplings, not a claim about literal quantum coherence *in vivo*. Cancer cells decouple from the organism’s temporal regulators and create their own accelerated proliferative cycle, “phasing out” of the organism’s coordinated physiology.

This framing avoids the over-specificity of claiming that cancer requires literal neural disconnection. Many tissues are weakly innervated yet show strong circadian regulation via hormonal, metabolic, and immune rhythms. The relevant claim is: malignancy is associated with degradation of *multi-source temporal coordination*, including but not limited to neural, endocrine, and local tissue clocks.

7.3 Perineural Invasion Reinterpreted

The known clinical phenomenon of tumors growing along nerves (perineural invasion) [20, 21] is typically interpreted as mechanical spread. Our framework suggests an additional interpretation: cancer cells may be hijacking high-bandwidth signaling channels for growth support, exploiting the physical coupling substrate while remaining relatively insensitive to termination control. The tumor microenvironment actively recruits neural tissue; autonomic nerves innervate tumors and provide neurotransmitters (norepinephrine, acetylcholine) that

alter the tumor microenvironment [20]. Remarkably, the brain can supply doublecortin-expressing neural progenitor cells to developing tumors via the bloodstream [21], suggesting active brain-tumor communication (“tumorigeption”). Neural coupling is therefore one high-bandwidth example, not a required substrate of malignancy: the boundary-failure claim is informational and applies equally to non-neural coupling channels (paracrine, mechanical, immune, and junctional).

8 Immune Evasion as Loss of Distinguishability

The immune system is the organism’s primary boundary-enforcement mechanism. Its logic resembles a decision function that maps high-dimensional cellular states into discrete actions (kill/ignore). This requires *distinguishable inputs*; cells must present features detectably different from self. In practice this decision is criterion-based, not binary: tumors can remain partly distinguishable while net signal quality is too low for reliable action. Formally, if Z denotes the immune-visible feature set (antigen presentation, checkpoint context, inflammatory state) and $D \in \{\text{clear, tolerate}\}$ the effective immune decision, then reliability depends on nontrivial decision information $I(Z; D)$. Evasion programs do not need to eliminate all signal; they need only reduce actionable information enough to drive decision quality toward chance.

8.1 The PD-1/PD-L1 Checkpoint Axis

Cancer exploits the PD-1/PD-L1 checkpoint axis [29] to reduce immune visibility. PD-1 is a checkpoint receptor on T cells that limits excessive stimulation. Tumor PD-L1 engages PD-1 and suppresses T-cell activation, enabling immune escape. This is a boundary failure at the immune interface: tumor cells use a normal self-protection pathway to avoid elimination.

Checkpoint inhibitor therapy (anti-PD-1/PD-L1 antibodies) works by *restoring distinguishability*: blocking the false “self” signal so that the immune system can detect and eliminate cancer cells [30]. In our framework, this is precisely “forced distinction”, the therapeutic principle derived from the boundary logic failure model.

8.2 Secondary Boundary Failures

Some cancers develop resistance to checkpoint blockade through secondary boundary failures [30]: JAK1/JAK2 mutations that impair interferon- γ responsiveness (blocking the information channel that normally induces PD-L1 expression), loss of $\beta 2$ -microglobulin eliminating MHC class I surface expression (removing the very molecules through which cells present antigens to T cells), and PTEN loss decreasing T cell infiltration. Each represents a further layer of boundary dissolution.

8.3 Gap Junctions and Antigen Broadcast

Immune evasion is often framed as a surface receptor problem (PD-L1). However, a deeper mechanism is the failure of *broadcast*. Gap junctions have been shown to transfer antigenic

peptides directly to dendritic cells for cross-presentation [42]. The downregulation of connexins in cancer is therefore not just a loss of contact inhibition, but an active decoupling from the tissue’s immune surveillance network. The cell becomes an informational black hole, sequestering its internal state from the organism.

9 Thermodynamic Perspective: Cost Externalization

9.1 The Warburg Effect as Metabolic Decoupling

Cancer cells often shift toward aerobic glycolysis (Warburg effect [31]). This mode is less ATP-efficient but supports rapid biomass accumulation [32, 33]. In this framework, the shift is not attributed only to mitochondrial dysfunction; it is also a marker of reduced coupling to organism-level metabolic coordination.

Normal oxidative phosphorylation depends on coordinated oxygen delivery, circadian metabolic timing, and tissue nutrient allocation. Aerobic glycolysis is more autonomous and can proceed with weaker systemic coordination. The cancer cell trades efficiency for autonomy, consistent with atavistic predictions of reversion to older metabolic programs [5].

The Warburg shift can also be read as reduced investment in differentiation-specific regulatory order. Maintaining a low-entropy differentiated state requires continuous ATP to suppress transcriptional noise [43]. Dedifferentiation moves cells toward a higher-entropy state [44], reducing local regulatory overhead while increasing organism-level burden.

9.2 Cachexia as Systemic Cost Payment

Cancer-associated cachexia (involuntary weight loss, muscle wasting, and metabolic disruption) affects 50–85% of cancer patients and contributes to 20% of cancer deaths [34, 35]. It cannot be reversed by nutritional supplementation alone, distinguishing it from simple malnutrition.

In our framework, cachexia reflects interacting mechanisms rather than a single “cost transfer” process: (i) direct substrate burden from tumor glucose/amino-acid demand, (ii) reduced local investment in identity-bearing differentiation, and (iii) inflammatory catabolism driven by tumor-immune interaction. The host’s attempts to meet tumor demand via muscle proteolysis and adipose lipolysis generate the cachexia phenotype, while cytokine signaling (IL-6, TNF- α , TGF- β) amplifies tissue wasting [34]. The tumor effectively conscripts host tissue as fuel, repurposing catabolic breakdown products into substrates supporting further proliferation.

This is not a violation of thermodynamics. It is a redistribution across scales: local proliferative advantage with reduced differentiation investment is coupled to global host burden and inflammatory wasting.

9.3 Entropy Production (Heuristic)

We use entropy-production language as an organizing heuristic, not as a first-principles law for all tumors. A conservative statement is that thermodynamic load tracks proliferative

burden:

$$\dot{S}_{\text{tumor}}(t) \propto N_{\text{tumor}}(t), \quad (9)$$

where $N_{\text{tumor}}(t)$ is an effective tumor burden proxy. If tumor burden grows approximately as $N_{\text{tumor}}(t) \sim e^{rt}$ over a phase of progression, then the associated thermodynamic load can also rise approximately exponentially over that phase.

In this framing, disease progression reaches a *critical transition* when tumor-driven load exceeds host dissipation and repair capacity. The term “event horizon” is metaphorical shorthand for that qualitative threshold, not a literal physical singularity.

10 Toy Formalism: Tissue as a Labeled Graph

10.1 State Variables

We model a tissue region as a labeled graph $\mathcal{T} = (V, E)$, where each node $i \in V$ is a cell and each edge $(i, j, \ell) \in E$ carries a label ℓ indicating the type of coupling.

Each cell carries:

- a discrete state $x_i \in \{\text{Healthy}, \text{PreCancer}, \text{Cancer}\}$,
- an identity-alignment scalar $p_i \in [0, 1]$ (interpretable as “organism alignment”),
- an energetic cost proxy $\epsilon_i > 0$, and
- a division-rate proxy $r_i \geq 0$.

Edges include (at minimum):

- *Regulatory/signaling* (SIG): growth inhibition, differentiation cues, contact inhibition
- *Lineage* (DIV): division history (parent-child)
- *Temporal coordination* (TMP): shared timing fields (circadian, hormonal)
- *Harmonic/adhesion* (HAR): tissue adhesion, gap junction communication

10.2 A Coherence Metric

We define a scalar coherence score $\kappa \in [0, 1]$ intended to summarize: (i) density of regulatory signaling, (ii) temporal alignment, and (iii) identity commitment (low entropy).

Let ρ_{sig} be the fraction of regulatory edges among all edges, and let ρ_{tmp} be an analogous proxy for temporal coordination. Let H be the mean Shannon entropy of the binary distribution $(p_i, 1 - p_i)$ for each cell’s alignment scalar $p_i \in [0, 1]$ (1.0 for fully organism-aligned, 0.0 for fully self-aligned). This is a proxy for identity ambiguity/commitment: H is maximal near $p_i = 0.5$ and minimal near $p_i \in \{0, 1\}$. It does *not* measure spatial heterogeneity across cells; it measures per-cell identity ambiguity.

$$H = \frac{1}{|V|} \sum_{i \in V} [-p_i \log_2 p_i - (1 - p_i) \log_2(1 - p_i)] \quad (10)$$

We then combine these terms in a simple convex score:

$$\kappa = \text{clip}_{[0,1]}(a\rho_{\text{sig}} + b\rho_{\text{tmp}} + c(1 - H)) \quad (11)$$

with weights $a, b, c \geq 0$ and $a + b + c = 1$. This definition provides a quantitative (albeit simplified) metric for tissue health, ranging from $\kappa \approx 1$ (perfectly healthy) to $\kappa \approx 0$ (terminal/metastatic).

Coherence (κ)	Interpretation
0.9 – 1.0	Healthy tissue
0.7 – 0.9	Normal tissue with minor perturbation
0.5 – 0.7	Pre-cancerous state
0.3 – 0.5	Early malignancy
0.1 – 0.3	Advanced malignancy
0.0 – 0.1	Terminal/Metastatic

Table 3: Coherence interpretation. These ranges are illustrative, not clinically calibrated.

10.3 Cancer Transform Operator

We define a cancer transform operator C that acts on a node and its incident relations:

- Alignment collapse: $p_i \rightarrow 0.1$ (self-dominant control)
- Coupling weakening: reduce HAR coupling (reduced adhesion / gap-junction communication)
- Edge pruning: remove SIG/TMP edges; retain DIV lineage edges
- Energy draw: $\epsilon_i \rightarrow \gamma\epsilon_i$ with $\gamma > 1$ (Warburg-proxy burden)
- Recursion: add copy(node) via a DIV edge (the `rec_succ` step)

11 Toward Measurable Proxies

For this framework to move beyond a conceptual model, the abstract variables must map to measurable quantities. We propose the following candidate proxies:

Abstract Variable	Candidate Proxy	Measurement Method
Identity alignment (p_i)	Differentiation state / marker expression	Flow cytometry, immunohistochemistry
Signaling density (ρ_{sig})	Gap junction communication	Connexin expression [36], FRAP assays
Temporal coordination (ρ_{tmp})	Circadian gene coherence	Bmal1/Clock/Per expression across tissue (qPCR, scRNA-seq) [25]
Identity entropy (H)	Transcriptional Entropy (SCENT/mRNAsi)	Signaling Entropy [45], Stemness Index [46]
Coherence (κ)	Composite tissue coordination	Combination of above metrics
Bioelectric coupling	Membrane voltage gradients	Voltage-sensitive fluorescent dyes [19]

Table 4: Candidate measurable proxies for abstract framework variables.

The theoretical variable H (identity entropy) can be operationalized with Signaling Entropy (SCENT), which quantifies promiscuity in the protein interaction network. High entropy correlates with stem-like undifferentiated states and poorer prognosis, consistent with the prediction that malignancy is a high-entropy, low-constraint state.

Loss of gap junction-mediated intercellular communication (GJIC) is associated with malignancy across multiple cancer types. Reduced connexin expression correlates with functional communication loss [36]. Transgenic mice lacking Cx32 or Cx43 show higher tumorigenesis risk [36]. This provides a measurable correlate for “signaling density.”

Levin’s work also shows that membrane voltage can be both a detection modality for tumors (distinct electrical impedance) and a functional regulator of oncogene-driven tumorigenesis [19]. This supports “bioelectric coupling” as a measurable axis. Direct causal links from voltage perturbation to antigen-presentation phenotypes (for example, MHC-I modulation) remain downstream hypotheses that require targeted tests.

12 Therapeutic Design Principle: Restore Observability and Enforce Termination

In the boundary-failure framing, therapies succeed when they restore boundary observability between malignant cells and organism-level control systems. We propose this as a unifying design principle, not as a claim about any single regimen.

12.1 Force Distinction

Increase discriminability of malignant cells relative to self:

- **Neoantigen presentation:** Personalized cancer vaccines that force expression of tumor-specific antigens.

- **Checkpoint inhibition:** Anti-PD-1/PD-L1 antibodies that remove the false “self” signal, restoring immune visibility [29].
- **Differentiation therapy:** A key example is all-trans retinoic acid (ATRA) in acute promyelocytic leukemia (APL). ATRA drives leukemic blasts into mature granulocytes, with complete remission rates above 90% [27, 28], *without direct cytotoxic killing*. In this framework, the intervention restores boundary observability by forcing stable differentiated identity.
- **Bispecific antibodies / CAR-T:** Engineered molecular bridges that force immune cell engagement with cancer cells, creating the distinction that cancer had dissolved.

12.2 Restore Termination Predicates

Re-enable growth arrest, senescence, or apoptosis pathways:

- p53 reactivation [14]
- Synthetic “stop” signals
- CDK inhibitors that enforce cell cycle checkpoints

12.3 Re-couple to Tissue Constraints

Reconstruct regulatory edges that impose contact inhibition and signaling coherence:

- Bioelectric reprogramming: manipulation of ion channels and gap junctions to restore collective voltage patterns that suppress tumorigenesis [18, 19]
- Chronotherapy: restoration of circadian regulation via clock-derived nanovesicles or timed drug administration [25]

12.4 Constrain Resource Externalization

Reduce the ability to offload metabolic and inflammatory costs onto the host:

- Metabolic inhibitors (e.g., 2-deoxyglucose, metformin) that increase the energetic cost of the Warburg phenotype
- Anti-cachexia interventions that protect host tissue from tumor-driven catabolism [34]

13 Clinical Correlations and Testable Hypotheses

The framework suggests qualitative hypotheses that can be checked against known clinical patterns and tested prospectively.

1. **Immune invisibility as boundary collapse:** Phenotypes associated with reduced antigen presentation and checkpoint manipulation should correlate with lower inferred “observability” in the tissue network. *Testable:* PD-L1 expression levels and MHC class I expression should inversely correlate with measurable proxies for tissue coherence (e.g., gap junction density, circadian gene coherence).
2. **Cachexia as cost externalization:** Systemic wasting can be interpreted as host-level payment for proliferative burden. *Testable:* Tumor metabolic rate (FDG-PET) should correlate with cachexia severity after controlling for tumor mass.
3. **Circadian disruption precedes malignancy:** Circadian disruption should raise boundary-failure probability by degrading checkpoint timing coordination. *Testable:* Circadian gene-expression coherence in peri-tumoral tissue should be lower than in matched normal tissue *before* overt malignancy is detectable. IARC night-shift data already provides population-level support [22, 23].
4. **Spontaneous remission as re-entanglement:** Rare immune-mediated regressions can be viewed as abrupt restoration of observability. *Testable:* Documented spontaneous remission cases [38] should show measurable immune reactivation (increased CD8+ T cell infiltration, pro-inflammatory cytokine surge) preceding tumor regression, with concurrent acute infections providing the immune trigger [39].
5. **Perineural invasion correlates with boundary loss:** Tumors with greater perineural invasion [20] should show lower coherence metrics and worse response to immunotherapy (since neural hijacking indicates deeper boundary failure).
6. **Bioelectric disconnection as early marker:** Changes in membrane voltage patterns should precede overt malignancy and predict tumor formation [19]. Levin’s tadpole experiments demonstrate this in principle: voltage management prevents tumor formation despite active oncogene expression.
7. **Growth-stability falsifier:** Any lineage showing persistently elevated proliferative throughput and high coherence without corresponding increase in control investment would challenge the control-budget formulation (Equation 8).

14 Empirical Validation: TCGA Cohort Analysis

To ground the framework’s predictions in existing data, we performed a systematic correlation, subtype, and survival analysis using bulk RNA-seq expression from The Cancer Genome Atlas (TCGA), accessed via the ISB Cancer Genomics Cloud BigQuery tables [50]. We analyzed six cohorts spanning $n = 3,611$ tumor samples: TCGA-SKCM (melanoma, $n = 472$), TCGA-LUAD (lung adenocarcinoma, $n = 530$), TCGA-BRCA (breast, $n = 1,113$), TCGA-COAD (colon, $n = 473$), TCGA-HNSC (head and neck squamous cell carcinoma, $n = 522$), and TCGA-LUSC (lung squamous cell carcinoma, $n = 501$). We extracted transcripts-per-million (TPM) values for 21 target genes spanning immune checkpoint (CD274/PD-L1, PDCD1LG2, PDCD1), MHC class I (HLA-A, HLA-B, HLA-C, B2M), gap junction

(GJA1/Cx43, GJB2/Cx26, GJA5/Cx40, GJB6/Cx30), circadian clock (ARNTL/BMAL1, CLOCK, PER1, PER2, CRY1, CRY2), and differentiation markers (CDH1, VIM, MYC, TP53). All correlations are Spearman rank (ρ), computed on $\log_2(\text{TPM}+1)$ -transformed values, with false discovery rate (FDR) correction applied across cancer types.

14.1 Circadian Coherence and PD-L1

We defined a per-sample circadian coherence proxy as the coefficient of variation (CV) across the six core clock genes (ARNTL, CLOCK, PER1, PER2, CRY1, CRY2), where lower CV indicates more uniform (coherent) clock-gene expression. Hypothesis 1 predicts that immune checkpoint engagement should correlate with reduced tissue coherence.

The circadian CV showed a significant negative correlation with PD-L1 expression in **all six cohorts** (100% replication rate, all FDR $q < 0.05$):

- TCGA-SKCM: $\rho = -0.381$, $p = 9.6 \times 10^{-18}$, $n = 472$
- TCGA-LUAD: $\rho = -0.303$, $p = 1.1 \times 10^{-12}$, $n = 530$
- TCGA-BRCA: $\rho = -0.222$, $n = 1,113$
- TCGA-COAD: $\rho = -0.312$, $n = 473$
- TCGA-HNSC: $\rho = -0.125$, $n = 522$
- TCGA-LUSC: $\rho = -0.142$, $n = 501$

This means that tumors with higher PD-L1 expression tend to show *lower* circadian CV (more coherent clock expression) across all cancer types tested. This initially appears to contradict the prediction that checkpoint-engaged tumors have degraded temporal coordination. However, this pattern is consistent with an *immune-hot* tumor phenotype: interferon- γ signaling from infiltrating T cells upregulates both PD-L1 and downstream transcriptional programs, including clock genes, producing an inflamed but checkpoint-protected microenvironment. The boundary failure in these tumors is not temporal decoherence but *active masking*—PD-L1 upregulation as an adaptive evasion response in tumors that retain enough signaling infrastructure to respond to immune pressure. The replication across six cancer types spanning distinct tissue origins (epithelial, mesenchymal, squamous) strongly supports the generalizability of this finding.

Among individual clock genes, ARNTL (BMAL1) showed the strongest positive association with PD-L1 (SKCM: $\rho = +0.428$, $p = 1.9 \times 10^{-22}$; LUAD: $\rho = +0.287$, $p = 1.6 \times 10^{-11}$), while PER1 was negatively correlated in both cohorts (SKCM: $\rho = -0.136$, $p = 0.003$; LUAD: $\rho = -0.099$, $p = 0.023$). The divergent directions within the clock network—BMAL1/CLOCK positively associated, PER genes negatively associated—suggest that PD-L1-high tumors exhibit *restructured* rather than uniformly disrupted circadian programs. Figure 3 summarizes the cross-cancer correlation structure.

14.2 PD-L1 and MHC Class I Co-regulation

PD-L1 expression was strongly positively correlated with MHC class I components in both cohorts:

- B2M: SKCM $\rho = +0.760$ ($p = 7.0 \times 10^{-90}$); LUAD $\rho = +0.555$ ($p = 3.7 \times 10^{-44}$)
- HLA-B: SKCM $\rho = +0.717$ ($p = 1.3 \times 10^{-75}$); LUAD $\rho = +0.547$ ($p = 1.2 \times 10^{-42}$)

This co-upregulation reflects shared IFN- γ transcriptional regulation: both PD-L1 and MHC-I are interferon-responsive. In the boundary framework, this identifies a specific failure mode: the antigen-presentation channel (MHC-I) is intact, but the organism’s clearance decision is overridden by checkpoint engagement (PD-L1). The boundary is not invisible—it is visible but *vetoed*. This refines Hypothesis 1: the most clinically relevant boundary failure is not always loss of distinguishability *per se*, but suppression of the organism’s *response* to detected distinction.

14.3 Gap Junction Expression

Connexin transcript levels showed weak positive correlations with PD-L1 (GJA1/Cx43: SKCM $\rho = +0.105$, $p = 0.023$; LUAD $\rho = +0.125$, $p = 0.004$), rather than the predicted inverse relationship. Gap junction composite expression was uncorrelated with MHC-I composite in SKCM ($\rho = +0.006$, $p = 0.89$) and weakly positive in LUAD ($\rho = +0.155$, $p = 0.0003$).

This is consistent with a known limitation of bulk RNA-seq: connexin transcript levels do not capture functional gap-junction intercellular communication (GJIC), which depends on protein localization, phosphorylation state, and channel gating [36]. Many connexins are post-translationally regulated, and some (e.g., Cx43) are internalized or phosphorylated to close channels without transcript-level changes. Testing the gap-junction arm of Hypothesis 1 requires functional assays (FRAP, scrape-loading, or dye transfer) rather than expression-level proxies.

14.4 Summary of Key Correlations

The circadian–checkpoint coupling replicates across all six cancer types. Table 5 summarizes the primary Spearman correlations for the two discovery cohorts; Figure 3 presents the full cross-cancer correlation heatmap.

Comparison	Category	TCGA-SKCM		TCGA-LUAD	
		ρ	p	ρ	p
CD274 vs Circadian CV	Coherence	-0.381	9.6×10^{-18}	-0.303	1.1×10^{-12}
CD274 vs ARNTL	Clock	+0.428	1.9×10^{-22}	+0.287	1.6×10^{-11}
CD274 vs PER1	Clock	-0.136	3.1×10^{-3}	-0.099	2.3×10^{-2}
CD274 vs B2M	MHC-I	+0.760	7.0×10^{-90}	+0.555	3.7×10^{-44}
CD274 vs GJA1	Gap jcn.	+0.105	2.3×10^{-2}	+0.125	3.9×10^{-3}
GJ mean vs MHC-I mean	Cross	+0.006	0.89	+0.155	3.3×10^{-4}
CDH1 vs VIM	EMT	-0.132	4.1×10^{-3}	+0.110	1.2×10^{-2}

Table 5: Spearman correlations for Hypothesis 1 target genes in TCGA tumor samples (discovery cohorts). Data source: ISB-CGC BigQuery (`isb-cgc-bq.TCGA.RNAseq_hg38_gdc_current`). All values are $\log_2(\text{TPM}+1)$ -transformed. Cross-cancer replication is shown in Figure 3. Extraction and analysis scripts are available in the companion repository under `experiments/tcga/`.

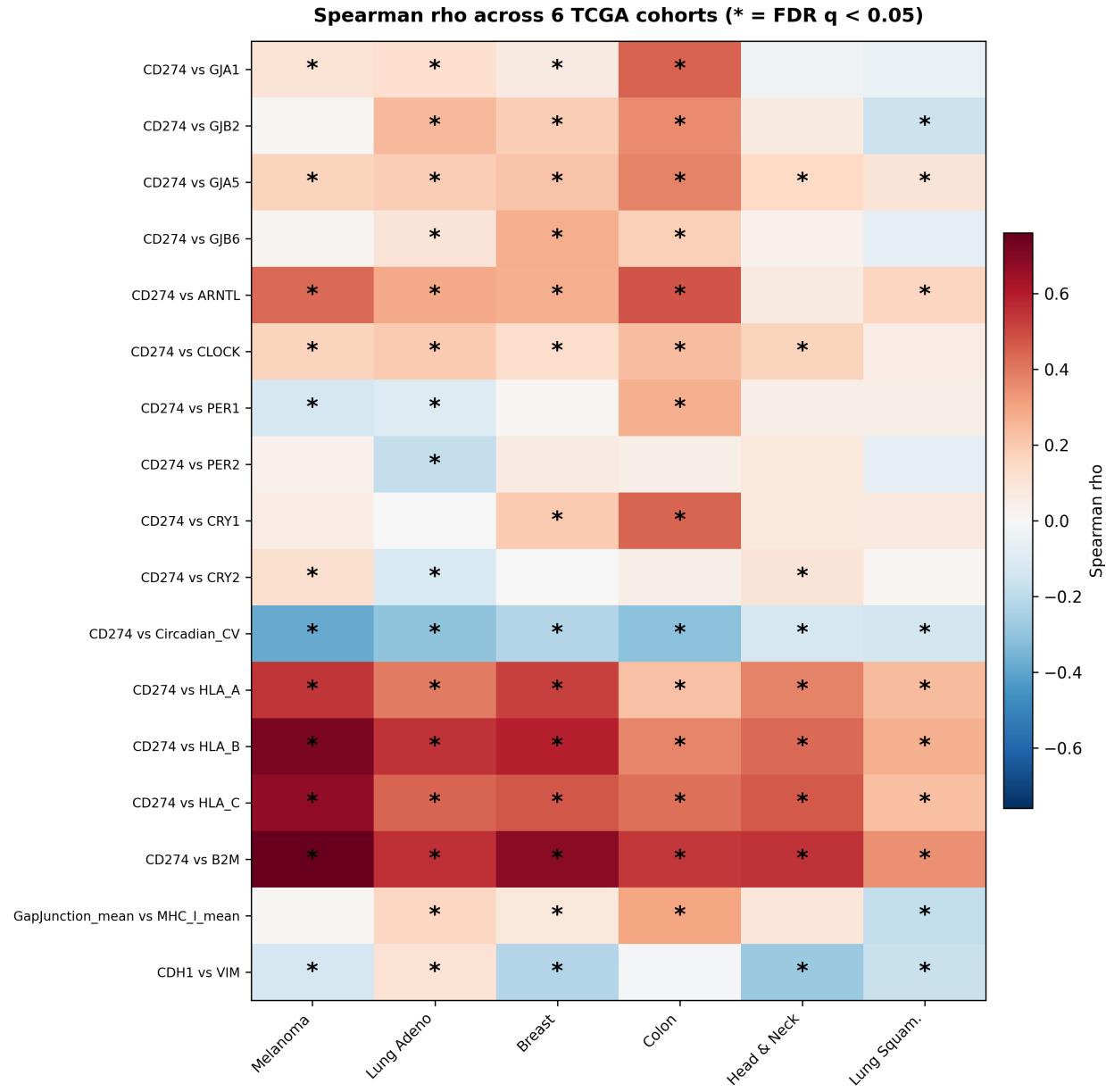


Figure 3: Cross-cancer correlation heatmap for key framework predictions across all six TCGA cohorts ($n = 3,611$ tumors). The circadian CV vs. PD-L1 negative correlation (top row) replicates in all six cancer types (100% replication rate). Color intensity reflects Spearman ρ ; asterisks denote FDR-corrected significance.

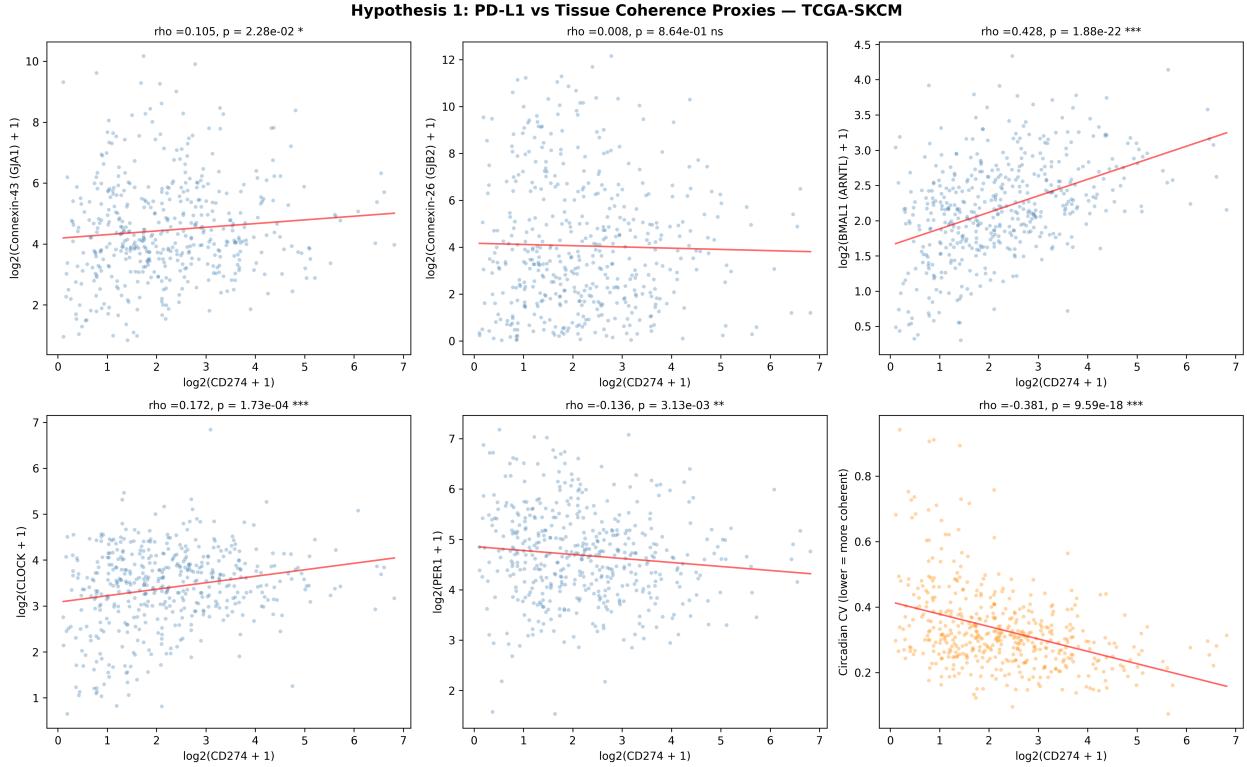


Figure 4: Hypothesis 1 validation in TCGA-SKCM (melanoma, $n = 472$ tumor samples). PD-L1 (CD274) expression vs. gap junction genes (GJA1, GJB2), clock genes (ARNTL, CLOCK, PER1), and circadian coherence (CV). The circadian CV panel (bottom right, orange) shows the strongest signal: higher PD-L1 correlates with lower CV (more coherent clock expression), consistent with the immune-hot phenotype.

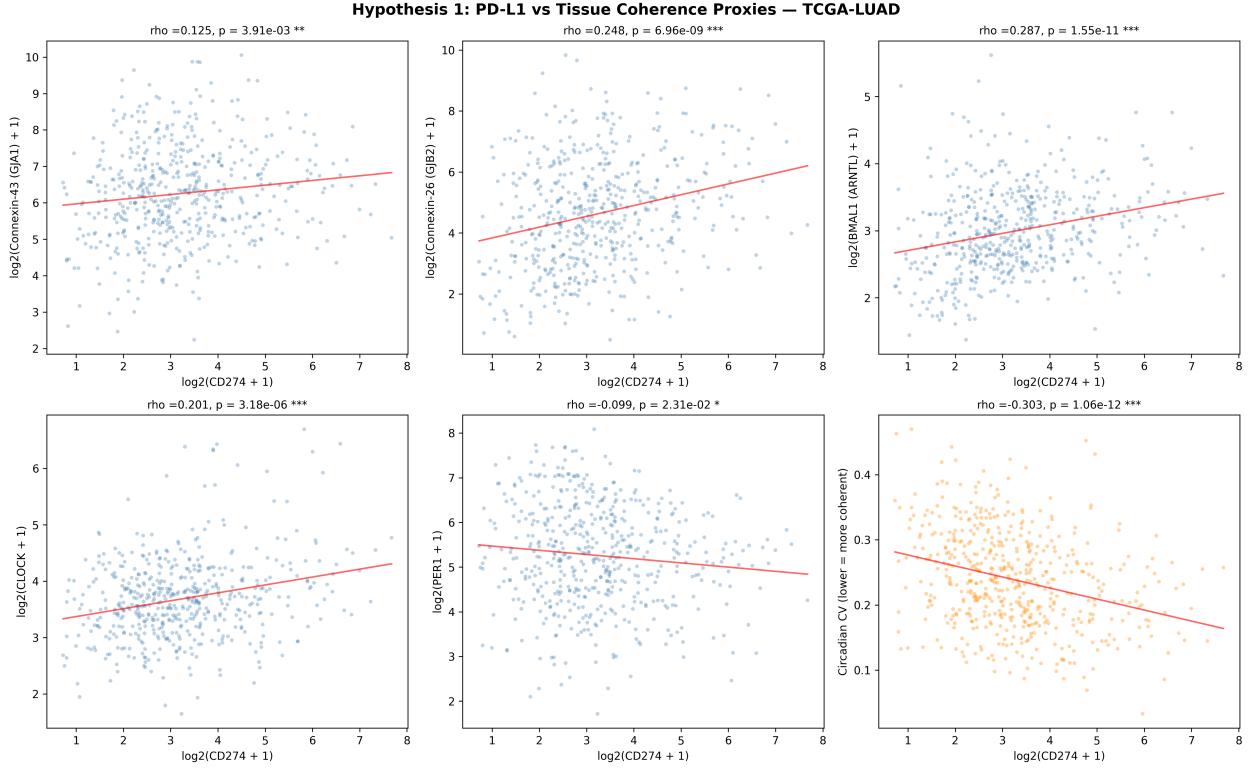


Figure 5: Hypothesis 1 validation in TCGA-LUAD (lung adenocarcinoma, $n = 530$ tumor samples). The same pattern replicates across a second cancer type, supporting the generalizability of the circadian–checkpoint coupling.

14.5 Tumor vs. Matched Normal Tissue

A critical test of the boundary-failure framework is whether tumor circadian programs differ systematically from matched normal tissue. If cancer simply “breaks” the clock, tumors should show higher circadian CV (more disordered clock expression) than matched normals. We compared circadian CV between tumor and matched normal samples in five cancer types with sufficient normal tissue representation (SKCM was excluded due to limited normal samples).

Contrary to the naive “broken clock” prediction, tumors showed **lower** circadian CV than matched normals in four of five cancer types tested:

- TCGA-LUAD: $p = 1.6 \times 10^{-6}$
- TCGA-BRCA: $p = 3.8 \times 10^{-9}$
- TCGA-HNSC: $p = 7.0 \times 10^{-7}$
- TCGA-LUSC: $p = 8.6 \times 10^{-7}$
- TCGA-COAD: not significant

This finding is striking: tumors are *more* circadian-coherent than normal tissue, not less. This is consistent with the “locked, not broken” interpretation of the boundary-failure framework. Cancer cells do not simply lose temporal coordination; they *restructure* it, locking into a rigid circadian program that may support proliferative advantage while reducing the flexibility needed for organism-level temporal coupling. The clock is not disrupted—it is co-opted. Figure 6 shows the paired tumor–normal comparisons.

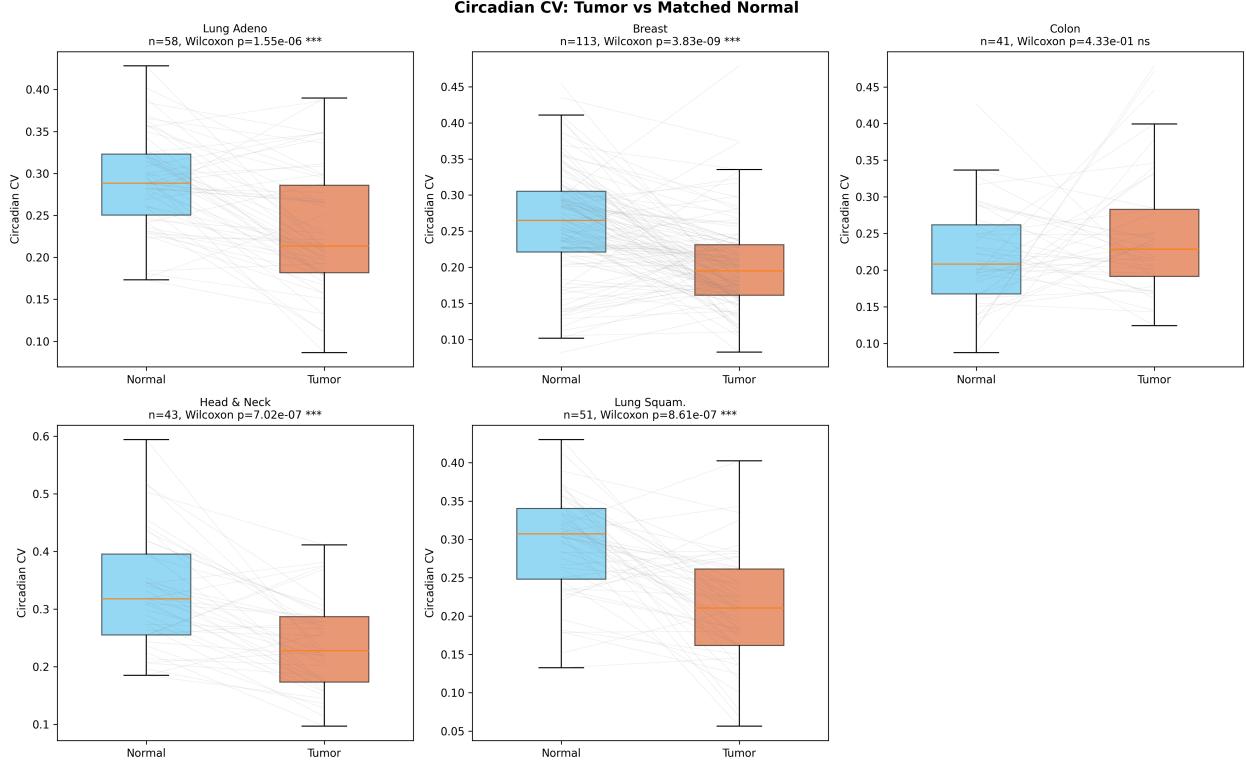


Figure 6: Circadian CV in tumor vs. matched normal tissue across five TCGA cancer types. Tumors show significantly *lower* CV (more coherent clock expression) than matched normals in 4/5 cancer types, supporting the “locked, not broken” interpretation of circadian restructuring in cancer.

14.6 Active Masking vs. Temporal Decoherence Classification

The circadian–checkpoint coupling and tumor–normal comparisons suggest that boundary failure in cancer is not monolithic. To formalize this, we classified tumors into boundary-failure subtypes based on expression of key framework genes:

- **Active Masking:** high PD-L1 (CD274) + high BMAL1 (ARNTL) + low PER1. These tumors retain a functional but restructured clock and deploy checkpoint molecules to veto immune clearance. The boundary is visible but *vetoed*.
- **Temporal Decoherence:** low PD-L1 + low B2M. These tumors have lost both checkpoint engagement and antigen-presentation capacity. The boundary is *invisible*.

- **Mixed:** all remaining tumors not classified into either category.

Thresholds were set at median expression within each cancer type. Active Masking tumors showed significantly **lower** circadian CV than Decoherence tumors in **all six cancer types** (Kruskal-Wallis test, all FDR $< 10^{-10}$). This confirms that the two boundary-failure modes are associated with distinct circadian architectures: Active Masking tumors maintain tight clock coordination, while Decoherence tumors show variable and disordered circadian programs.

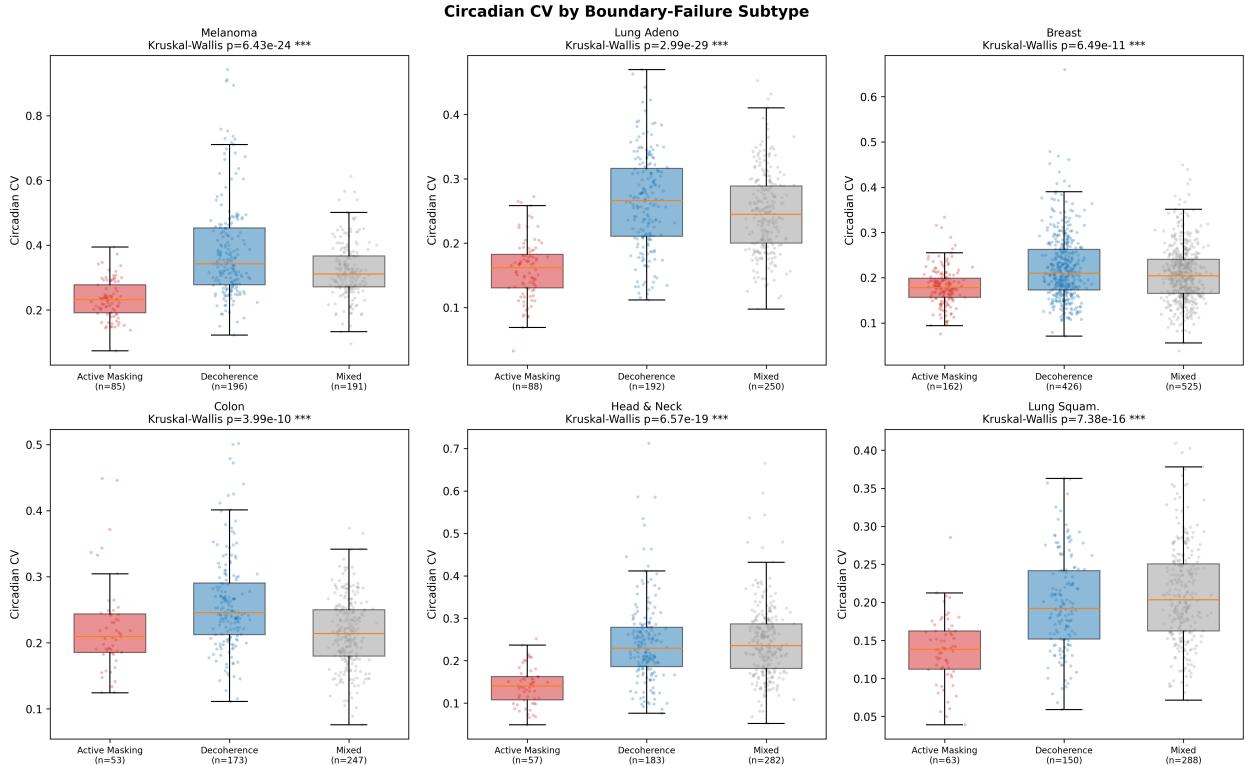


Figure 7: Circadian CV by boundary-failure subtype across six TCGA cancer types. Active Masking tumors (high PD-L1, high BMAL1, low PER1) show consistently lower CV than Decoherence tumors (low PD-L1, low B2M) in all six cancer types (all Kruskal-Wallis FDR $< 10^{-10}$), confirming distinct circadian architectures for different boundary-failure modes.

14.7 Survival Analysis

If the boundary-failure mode carries biological significance beyond gene-expression patterns, it should predict clinical outcomes. We compared overall survival between Active Masking and Decoherence tumors using Kaplan-Meier analysis with log-rank tests.

Active Masking tumors showed significantly **better** overall survival than Decoherence tumors in two cancer types:

- TCGA-SKCM: $p = 0.0008$ (Active Masking survives longer)

- TCGA-LUAD: $p = 0.024$ (same direction)
- TCGA-BRCA, TCGA-COAD, TCGA-HNSC, TCGA-LUSC: not significant

Notably, quartile-based circadian CV alone did *not* predict survival in any of the six cancer types, suggesting that the boundary-failure *mode* matters more than coherence level per se. A tumor with low circadian CV is not inherently better or worse; what matters is *why* its clock is coherent. Active Masking tumors (coherent clock + checkpoint engagement) are visible to the immune system and potentially responsive to checkpoint inhibitor therapy. Decoherence tumors (disordered clock + antigen loss) are invisible and likely resistant to immunotherapy.

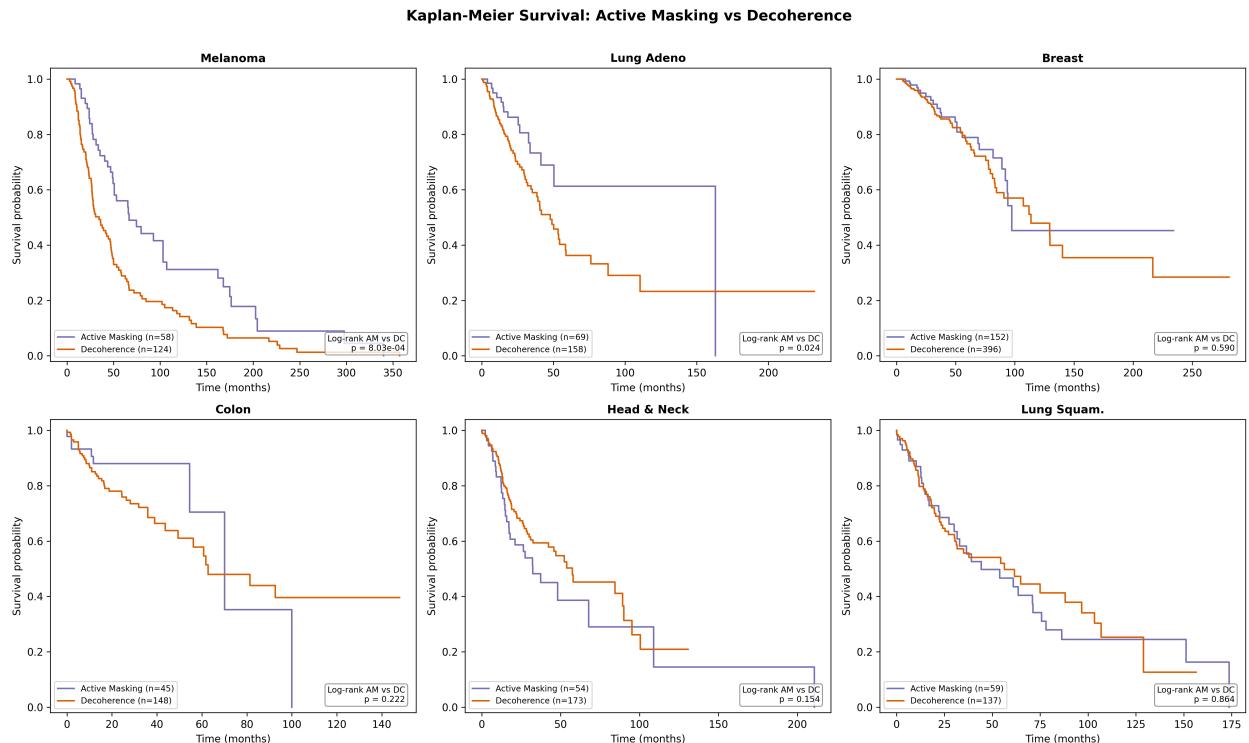


Figure 8: Kaplan-Meier survival curves by boundary-failure subtype. Active Masking tumors show significantly better overall survival than Decoherence tumors in SKCM ($p = 0.0008$) and LUAD ($p = 0.024$). The survival advantage of Active Masking is consistent with the interpretation that these tumors retain immune visibility and are candidates for checkpoint inhibitor therapy.

14.8 Stage-Stratified Analysis

To determine whether circadian restructuring is an early or late event in tumor progression, we tested the association between circadian CV and pathologic tumor stage across all six cancer types.

Circadian CV showed **no significant correlation** with tumor stage in any of the six cancer types (all FDR-corrected $p > 0.05$). This is an important negative result: it indicates that circadian restructuring is established *early* in tumorigenesis and maintained throughout progression, rather than accumulating as a late-stage consequence of increasing malignancy. The boundary-failure mode is set early and persists.

This finding is consistent with the framework's prediction that boundary failure is a foundational event—a shift in the cell's relationship with the organism—rather than a downstream consequence of tumor burden. It also suggests that circadian-based biomarkers could be informative at early disease stages, when therapeutic intervention is most effective.

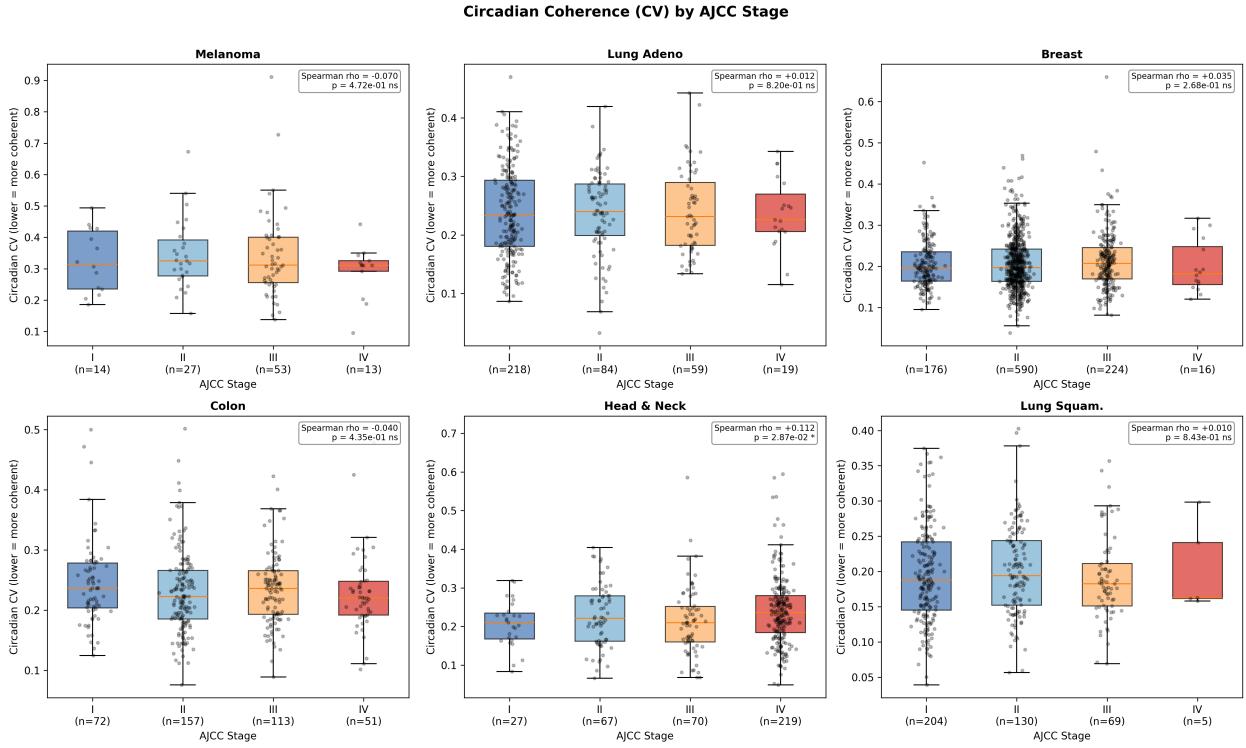


Figure 9: Circadian CV by pathologic tumor stage across six TCGA cancer types. No significant association between circadian CV and stage was observed in any cancer type (all FDR > 0.05), indicating that circadian restructuring is an early event maintained throughout progression.

14.9 Implications for the Framework

The expanded TCGA analysis across six cancer types and $n = 3,611$ tumors refines the boundary-failure framework in five ways:

- 1. Boundary failure is not monolithic.** The data confirm at least two distinct failure modes across all six cancer types: (a) *Active Masking*, where the tumor retains a locked, coherent clock and deploys checkpoint molecules to veto organism-level clearance (visible but vetoed), and (b) *Temporal Decoherence*, where clock programs are disordered and

immune visibility is low (invisible). Both are boundary failures, but through different mechanisms with different clinical implications.

2. **The clock is locked, not broken.** Tumors show lower circadian CV than matched normals in 4/5 cancer types, and Active Masking tumors show the lowest CV of all. This is inconsistent with a simple “circadian disruption” model and supports clock co-option.
3. **Boundary-failure mode predicts survival.** Active Masking tumors show better overall survival than Decoherence tumors in melanoma and lung adenocarcinoma ($p = 0.0008$ and $p = 0.024$, respectively), while circadian CV alone does not predict survival. The *mode* of failure matters more than the degree of coherence.
4. **Circadian restructuring is early and stable.** No association with tumor stage was observed, indicating that the boundary-failure mode is established early and maintained.
5. **Direct therapeutic implications.** The Active Masking / Decoherence distinction suggests different therapeutic strategies: checkpoint inhibitors should be most effective for Active Masking tumors (restore the vetoed immune response), while Decoherence tumors may require differentiation therapy or antigen-restoration strategies first (restore immune visibility before checkpoint blockade can be effective). Gap junction expression remains an unreliable RNA-level proxy for functional GJIC; functional assays are required.

15 Simulation

We include a simple simulation (Appendix A) to demonstrate the qualitative behavior implied by the boundary-failure hypothesis. The simulation is not calibrated to real tumors; it is designed to show a failure mode.

15.1 Update Rules

At each discrete time step, each node:

1. Consumes energy ϵ_i from a global control budget B_{ctrl} (a coarse proxy for host resources and control capacity; cf. Eq. 8).
2. May mutate from Healthy to PreCancer (with small probability).
3. May progress from PreCancer to Cancer with a corresponding drop in alignment p_i .
4. If cancerous, loses a subset of regulatory edges (decoupling from SIG and TMP).
5. Divides with probability proportional to r_i ; healthy nodes are subject to a contact-inhibition proxy (maximum neighbor count) while cancerous nodes are not.

We record κ over time and track the cancer cell count. Under reasonable parameters, the system exhibits a transition from a high- κ , low-division regime to a low- κ , high-division regime once decoupled nodes appear.

15.2 What the Simulation Demonstrates

- Loss of signaling edges + self-referential proliferation → coherence collapse
- Healthy cells self-regulate via contact inhibition; cancer cells do not
- Energy budget depletion accelerates as cancer cell count grows (Warburg effect proxy)
- The “crossing point” where κ decline and cell count acceleration meet represents the qualitative event horizon of the disease

Disclaimer: This is not a predictive cancer model. It does not capture genetics, metabolism, angiogenesis, immune editing, clonal heterogeneity, or spatial constraints. It demonstrates logical failure modes, not molecular causation.

Extended simulation framework: A modular Python implementation is available in the companion repository (github.com/MosesRahnama/cancer-paper-repository). It includes a simulation engine with history tracking, four therapeutic operators (forced distinction, differentiation therapy, checkpoint inhibition, bioelectric reprogramming), tissue-network visualization, and coherence-versus-cancer plotting. It implements the cancer transform operator (Section 10.3), the forced-distinction principle (Section 12), and a Warburg-proxy energy budget for end-to-end runs from healthy tissue through emergence, intervention, and post-intervention dynamics.

Calibrated monthly parameterization: To avoid an artificially vanishing pre-cancer dwell-time, we parameterize one simulation step as one month and map annual transformation risk to per-step probability via $p_{\text{step}} = 1 - (1 - r_{\text{annual}})^{1/12}$. Premalignant progression rates vary by lesion class [47]; for the illustrative oral epithelial dysplasia setting we use the pooled malignant transformation estimate $r_{\text{annual}} \approx 10.5\%$ [48], yielding $p_{\text{progression}} \approx 0.0092$. We pair this with $p_{\text{mutation}} = 0.002$ to represent persistent somatic mutation pressure in a toy setting [49]. In the companion code, the cancer energy multiplier ($\epsilon = 2.7$) is an illustrative Warburg-burden proxy for qualitative dynamics, not a fitted physiological constant.

Illustrative scenario comparison (extended framework): For one representative run (100 initial cells, seed=42, therapy applied at step 200), the calibrated model produces the following qualitative outcomes:

Therapy	Post-therapy κ	Cancer cells
No therapy (control)	0.418	376
Forced distinction	0.231	376
Differentiation therapy (ATRA-like)	0.441	0
Checkpoint inhibitor (anti-PD-1-like)	0.232	376
Bioelectric reprogramming (Levin-like)	0.199	376

These values are illustrative and parameter-dependent, not efficacy rankings. In this configuration, differentiation-like restoration reaches a halt state, while partial distinction-restoration operators increase observability without full clearance.

Figures 10, 11, 12, 13, and 14 visualize these trajectories from the companion codebase. The first two are generated by `python cancer-paper-repository/examples/run_full_scenario.py`; the latter three are generated by `python cancer-paper-repository/run_simulation.py`. All outputs are written to `cancer-paper-repository/results/`.

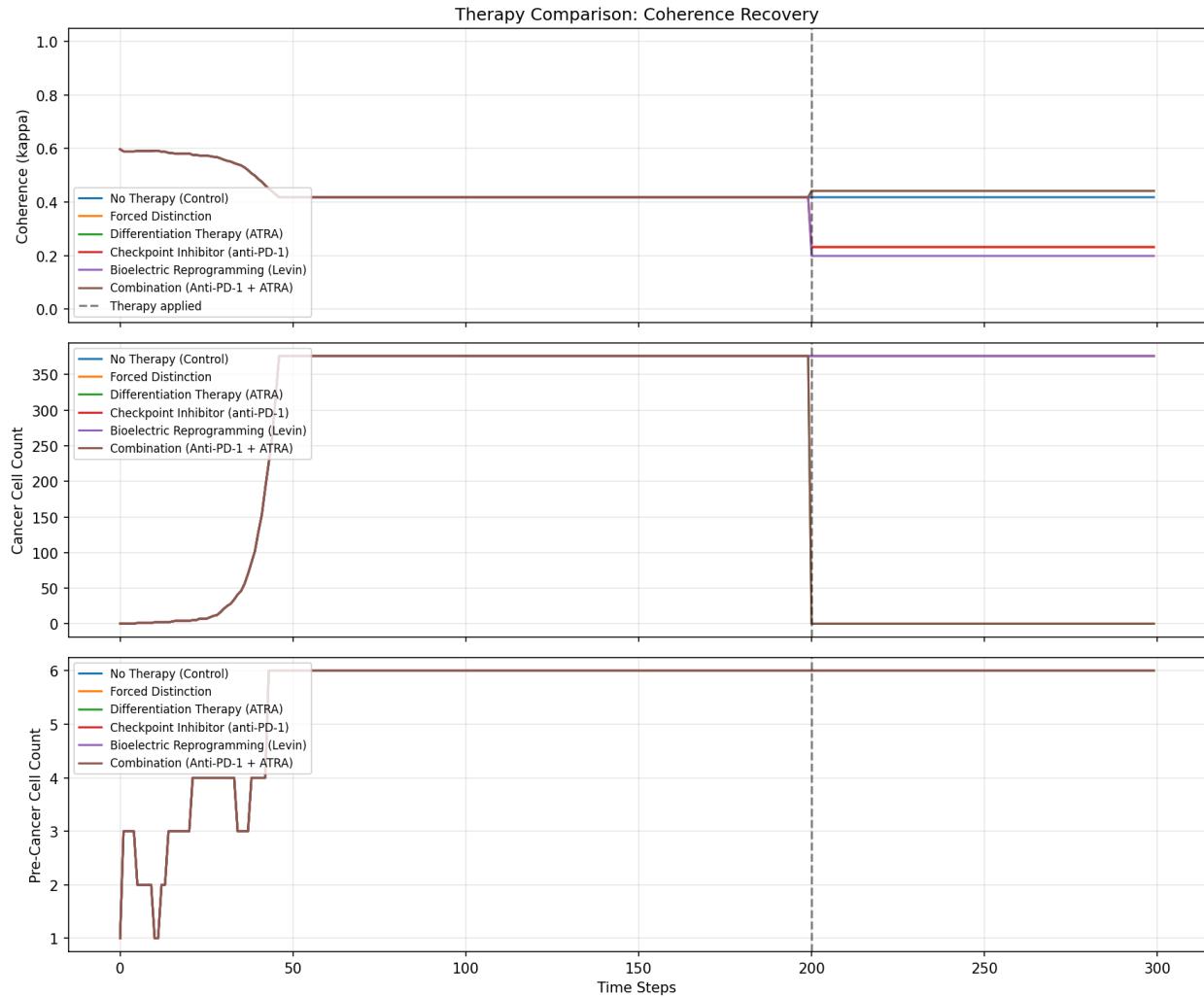


Figure 10: Extended simulation framework: coherence, cancer-count, and pre-cancer trajectories across five intervention strategies under the same initialization and seed.

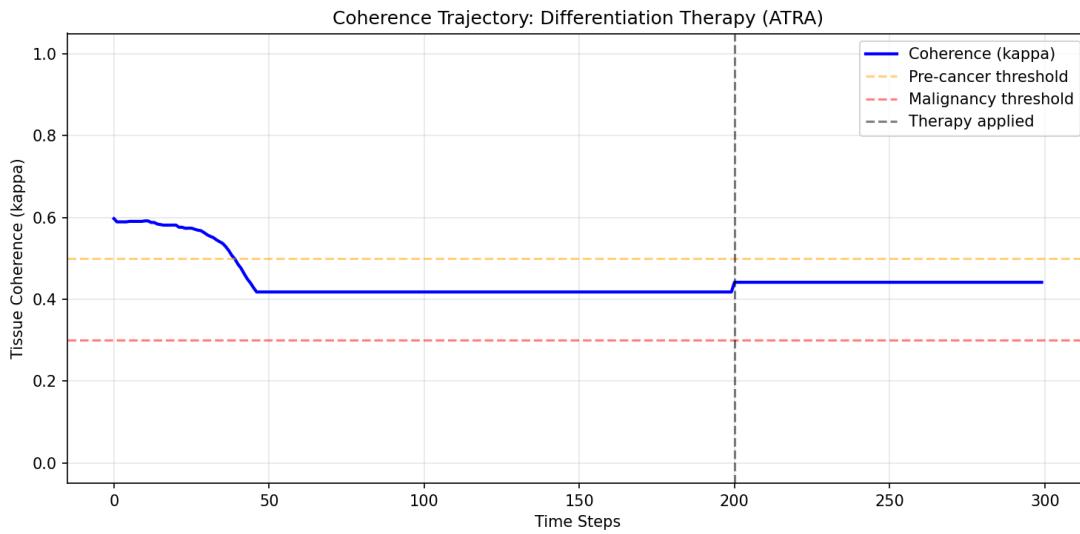


Figure 11: Single-therapy coherence trajectory (ATRA-like differentiation setting), showing post-intervention recovery in this illustrative run.

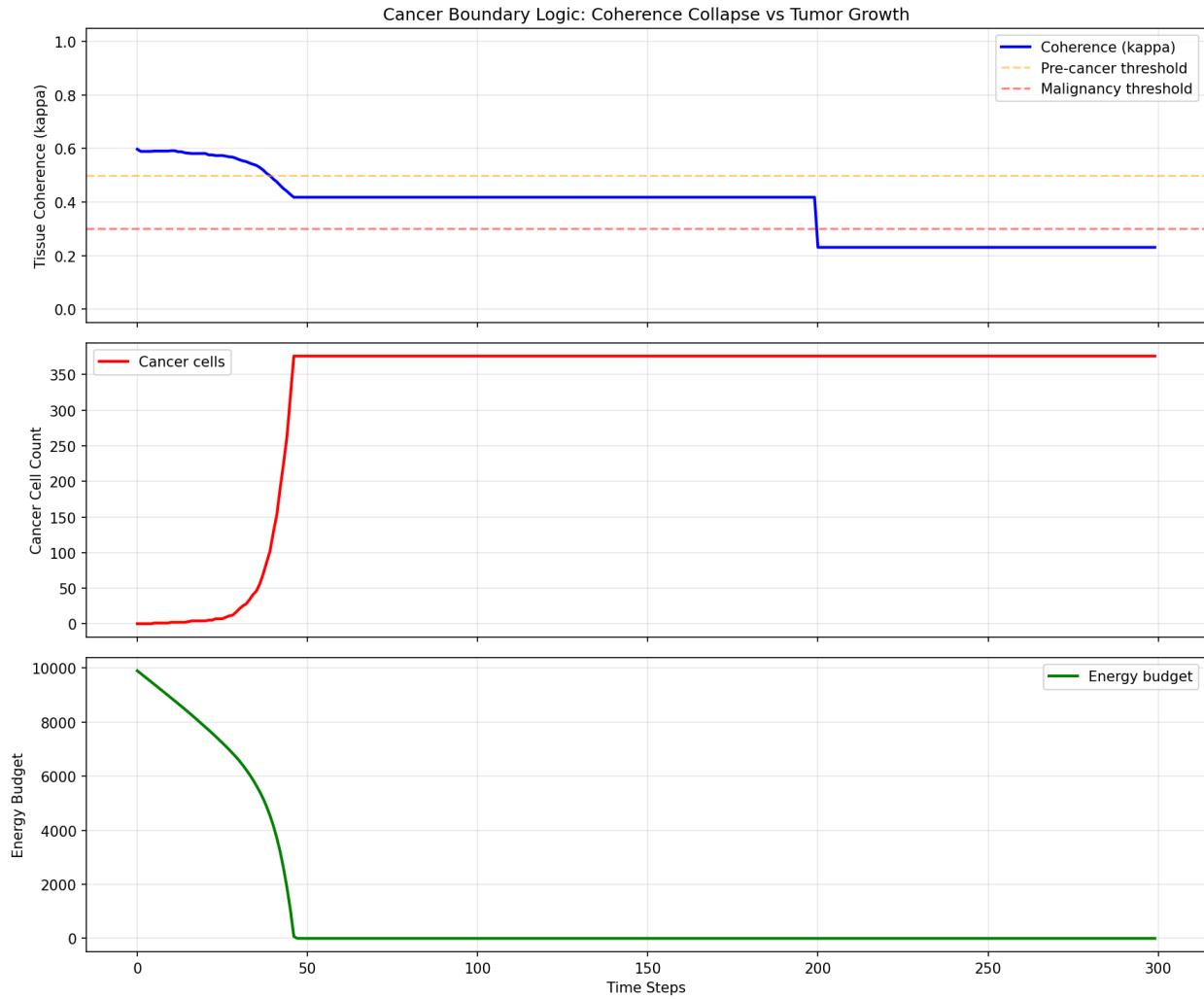


Figure 12: Reference simulation trajectory from `run_simulation.py`: coherence decline, cancer expansion, and energy-budget depletion over time.

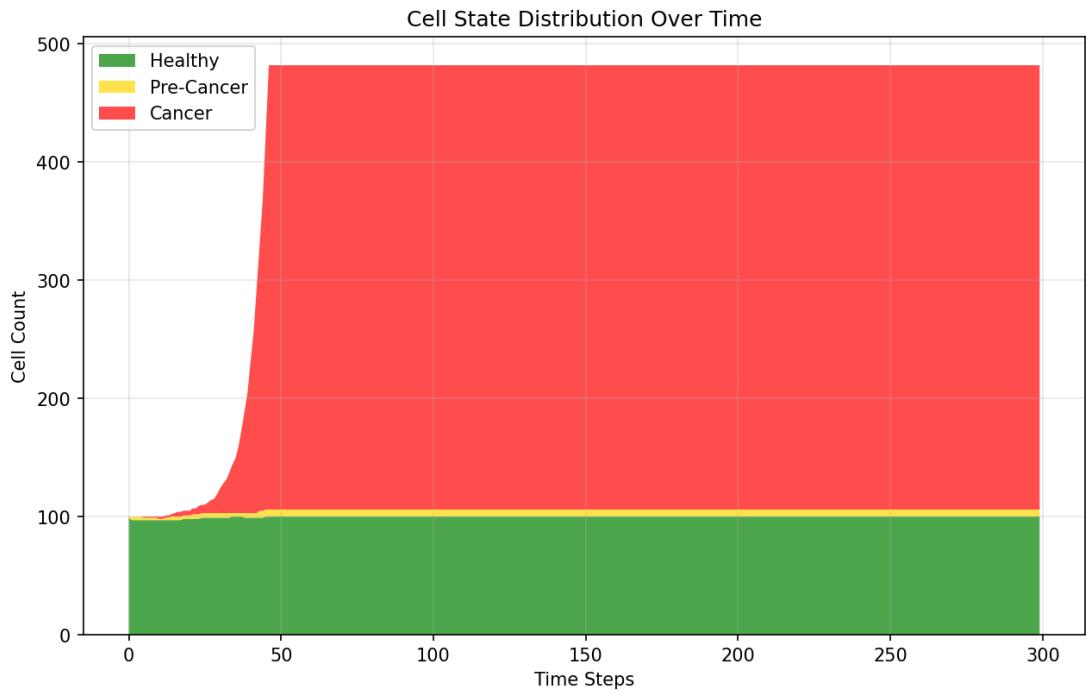


Figure 13: State distribution across simulation steps (healthy, pre-cancer, cancer), illustrating phase transitions in the toy tissue model.

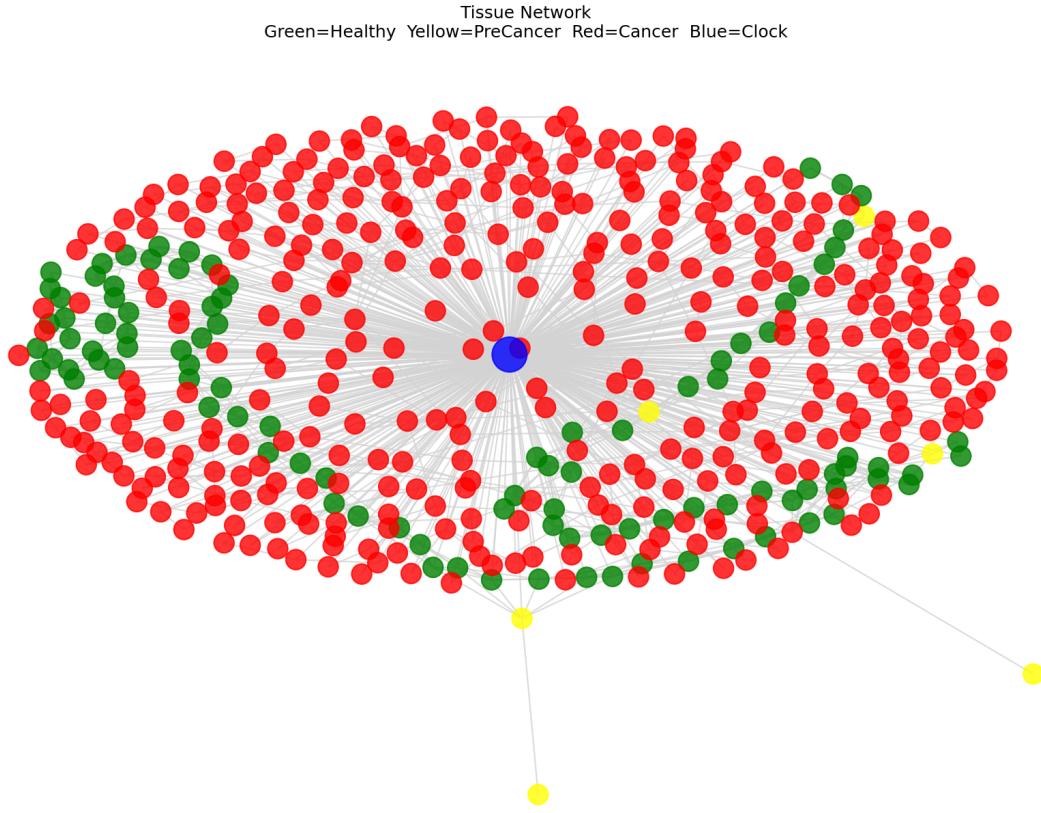


Figure 14: Representative tissue-network snapshot from the simulation, colored by cell state and showing local coupling structure.

16 Limitations

- **Model abstraction:** The graph variables p_i , edge labels, and κ are conceptual proxies. They are not directly mapped to specific molecular pathways. Operationalizing them requires the measurement program outlined in Section 11.
- **No genetics or heterogeneity:** The simulation does not capture mutational landscapes, clonal selection, angiogenesis, invasion, or heterogeneity beyond a coarse state variable.
- **Parameter and seed dependence:** The current simulation outputs are from illustrative parameter settings and a representative random seed. Multi-seed sweeps and sensitivity analyses are required before drawing comparative conclusions across operators.
- **No explicit immune-kill dynamics:** Immune recognition, clearance kinetics, and adaptive escape are not explicitly modeled. Some operators therefore change observability without automatically enforcing elimination.

- **Thermodynamics is interpretive:** While energy and information constraints motivate the narrative, we do not derive quantitative metabolic bounds from first principles. The Landauer framing is used as an organizing guide for distinguishing energetic replication cost from reduced investment in identity-forming information; host metabolic burden and cachexia remain explicitly multi-mechanism processes.
- **Analogy boundaries:** `rec_succ` is used to clarify a structural failure mode (duplication without a termination predicate). It is not a claim of literal undecidability in biology.
- **Identity-weight abstraction:** The variables w_{self} , w_{org} , and $A(t)$ are coarse control proxies. Operational mappings to specific molecular observables require dedicated calibration.
- **Control-budget constraint:** Equation 8 is phenomenological. Estimating $C_G(\cdot)$, $C_S(\cdot)$, and B_{ctrl} from experimental data is an open program.

17 Conclusion

We propose a boundary-first cancer framework in which malignancy is progressive loss of organism-level observability and control. The result is self-referential proliferation that is locally viable and globally destabilizing. The model links loss of contact inhibition, checkpoint failure, immune evasion, metabolic offloading, circadian disruption, and bioelectric decoupling under one boundary-logic account.

The practical value of the framework is threefold:

1. **As an organizing language:** It highlights the shared structure across seemingly different cancer mechanisms; all involve loss of cross-scale coupling and collapse into local self-reference.
2. **As a therapeutic design principle:** “Forced distinction” unifies the logic behind immunotherapy, differentiation therapy, bioelectric reprogramming, and chronotherapy. In this framing, interventions lie on a distinction-restoration continuum: some improve observability, while stronger restoration can push cells toward explicit halt states.
3. **As a source of testable hypotheses:** The framework generates predictions about circadian disruption, bioelectric markers, coherence metrics, and growth-stability tradeoffs that can be tested against existing and prospective data. TCGA validation across six cancer types ($n = 3,611$ tumors; Section 14) provides multi-cancer empirical support: the circadian–checkpoint coupling replicates in 100% of cohorts tested, tumors show locked (not broken) clocks relative to matched normals, and the boundary-failure mode—Active Masking vs. Temporal Decoherence—predicts overall survival in melanoma and lung adenocarcinoma.

For interpretive clarity, we separate three layers: [M] metaphor (conceptual analogies such as `rec_succ`), [C] claim (structural assertions about boundary failure and observability), and [P] prediction (falsifiable hypotheses in Section 13, with multi-cancer validation in Section 14).

The Active Masking / Decoherence distinction carries direct therapeutic implications: checkpoint inhibitors should be most effective for Active Masking tumors (where the immune response is visible but vetoed), while Decoherence tumors may require differentiation therapy or antigen-restoration strategies to restore immune visibility before checkpoint blockade can be effective. This suggests that boundary-failure subtyping could complement existing biomarkers (PD-L1 IHC, tumor mutational burden) for immunotherapy patient selection.

Next steps are to (i) validate the Active Masking / Decoherence classification in independent cohorts and clinical trial data with immunotherapy response outcomes, (ii) operationalize gap-junction and bioelectric proxies using functional assays rather than transcript-level data, and (iii) test the circadian-restructuring finding in single-cell RNA-seq datasets where clock-gene heterogeneity within tumors can be resolved at cellular resolution. A companion repository (github.com/MosesRahnama/cancer-paper-repository) implements the tissue graph model, four therapeutic operators, TCGA extraction and analysis pipelines (six cancer types), survival analysis, immune-subtype classification, and visualization tools for reproducibility and community testing.

18 Future Research Directions

1. **Operationalizing control budget terms:** Estimating $C_G(\cdot)$, $C_S(\cdot)$, and B_{ctrl} by pairing proliferation assays with tissue-integrity and checkpoint-load metrics across tissue types.
2. **Bioelectric boundary mapping:** Applying voltage-sensitive imaging to characterize the “bioelectric boundary” of tumors *in vivo* and correlating with treatment response.
3. **Coherence biomarkers:** Developing clinical tests combining gap junction density, circadian gene coherence, and transcriptional heterogeneity into a composite tissue coherence score.
4. **Chronotherapy integration:** Testing whether circadian-aligned drug administration improves outcomes specifically in tumors with measurably disrupted clock gene expression.
5. **Differentiation therapy expansion:** Applying the “forced distinction” principle to identify differentiation-inducing agents for cancers beyond APL.
6. **Computational oncology:** Using the `rec_succ` structural analogy to develop agent-based models that predict tumor growth patterns based on termination predicate availability (p53 status, checkpoint integrity, immune visibility).

Acknowledgements

This work benefited from iterative technical review and tool-assisted drafting. All claims, interpretations, and errors are the responsibility of the author.

References

- [1] Rahnama, M. (2026). *Thermodynamic Constraints on Measurement Events: A Boundary Framework for Classical Information*. Zenodo. <https://doi.org/10.5281/zenodo.18445561>
- [2] Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, 144(5), 646–674. <https://doi.org/10.1016/j.cell.2011.02.013>
- [3] Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. *Cancer Discovery*, 12(1), 31–46. <https://doi.org/10.1158/2159-8290.CD-21-1059>
- [4] Nowell, P. C. (1976). The Clonal Evolution of Tumor Cell Populations. *Science*, 194(4260), 23–28. <https://doi.org/10.1126/science.959840>
- [5] Lineweaver, C. H., Bussey, K. J., Blackburn, A. C., & Davies, P. C. W. (2021). Cancer progression as a sequence of atavistic reversions. *BioEssays*, 43(7), e2000305. <https://doi.org/10.1002/bies.202000305>
- [6] Trigos, A. S., et al. (2017). Altered interactions between unicellular and multicellular genes drive hallmarks of transformation in a diverse range of solid tumors. *Proc. Natl. Acad. Sci. USA*, 114(25), 6406–6411. <https://doi.org/10.1073/pnas.1617743114>
- [7] Soto, A. M., & Sonnenschein, C. (2011). The tissue organization field theory of cancer: A testable replacement for the somatic mutation theory. *BioEssays*, 33(5), 332–340. <https://doi.org/10.1002/bies.201100025>
- [8] Landauer, R. (1961). Irreversibility and Heat Generation in the Computing Process. *IBM J. Res. Dev.*, 5(3), 183–191. <https://doi.org/10.1147/rd.53.0183>
- [9] Bennett, C. H. (1982). The Thermodynamics of Computation: A Review. *Int. J. Theor. Phys.*, 21(12), 905–940. <https://doi.org/10.1007/BF02084158>
- [10] Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell Syst. Tech. J.*, 27, 379–423 and 623–656.
- [11] England, J. L. (2013). Statistical Physics of Self-Replication. *J. Chem. Phys.*, 139(12), 121923. <https://doi.org/10.1063/1.4818538>
- [12] Turing, A. M. (1937). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proc. London Math. Soc.*, s2-42(1), 230–265.
- [13] Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38(1), 173–198.
- [14] Levine, A. J., & Oren, M. (2009). The first 30 years of p53: growing ever more complex. *Nature Reviews Cancer*, 9(10), 749–758. <https://doi.org/10.1038/nrc2723>
- [15] Lane, D. P. (1992). p53, guardian of the genome. *Nature*, 358(6381), 15–16. <https://doi.org/10.1038/358015a0>

- [16] Lowe, S. W., Cepero, E., & Evan, G. (2004). Intrinsic tumour suppression. *Nature*, 432(7015), 307–315. <https://doi.org/10.1038/nature03098>
- [17] Okada, T., Lopez-Lago, M., & Giancotti, F. G. (2005). Merlin/NF-2 Mediates Contact Inhibition of Growth by Suppressing Recruitment of Rac to the Plasma Membrane. *J. Cell Biol.*, 171(2), 361–371. <https://doi.org/10.1083/jcb.200503165>
- [18] Levin, M. (2021). Bioelectric signaling: Reprogrammable circuits underlying embryogenesis, regeneration, and cancer. *Cell*, 184(6), 1971–1989. <https://doi.org/10.1016/j.cell.2021.02.034>
- [19] Levin, M. (2013). Reprogramming cells and tissue patterning via bioelectricity. *Organogenesis*, 9(3), 143–149. PMC3841289.
- [20] Zahalka, A. H., & Frenette, P. S. (2020). Nerves in cancer. *Nature Reviews Cancer*, 20(3), 143–157. <https://doi.org/10.1038/s41568-019-0237-2>
- [21] Magnon, C., et al. (2024). Autonomic nerve development and tumor growth. *Genes & Development*, 38(17–20), 802–804. <https://doi.org/10.1101/gad.352278.124>
- [22] International Agency for Research on Cancer (2019). *IARC Monographs Volume 124: Night Shift Work*. IARC, Lyon.
- [23] Schernhammer, E. S., et al. (2006). Night work and risk of breast cancer. *Epidemiology*, 17(1), 108–111. <https://doi.org/10.1097/01.ede.0000190539.03500.c1>
- [24] Finnish Twin Cohort (2023). Shift work including night shifts and breast cancer risk. *Eur. J. Epidemiol.* (2023). PMID: 36964875.
- [25] Mello, R. M., Masri, S., & Lamia, K. A. (2026). Rhythms of risk: the intersection of clocks, cancer, and chronotherapy. *J. Clin. Invest.*, 136(3), e198780. <https://doi.org/10.1172/JCI198780>
- [26] Masri, S., et al. (2016). Lung Adenocarcinoma Distally Rewires Hepatic Circadian Homeostasis. *Cell*, 165(4), 896–909. <https://doi.org/10.1016/j.cell.2016.04.039>
- [27] Grimwade, D., Mistry, A. R., Solomon, E., & Guidez, F. (2010). Acute promyelocytic leukemia: a paradigm for differentiation therapy. *Cancer Treat. Res.*, 145, 219–235.
- [28] Huang, M. E., et al. (2013). All-trans retinoic acid in acute promyelocytic leukemia. *Chin. J. Hematol.* PMC3757424.
- [29] Lin, X., Kang, K., Chen, P., et al. (2024). Regulatory mechanisms of PD-1/PD-L1 in cancers. *Molecular Cancer*, 23, 108. <https://doi.org/10.1186/s12943-024-02023-w>
- [30] Nowicki, T. S., Hu-Lieskovan, S., & Ribas, A. (2018). Mechanisms of resistance to PD-1 and PD-L1 blockade. *Cancer Journal*, 24(1), 47–53. <https://doi.org/10.1097/PPO.0000000000000303>
- [31] Warburg, O. (1956). On the Origin of Cancer Cells. *Science*, 123(3191), 309–314.

- [32] Vander Heiden, M. G., Cantley, L. C., & Thompson, C. B. (2009). Understanding the Warburg Effect. *Science*, 324(5930), 1029–1033. <https://doi.org/10.1126/science.1160809>
- [33] Pavlova, N. N., & Thompson, C. B. (2016). The Emerging Hallmarks of Cancer Metabolism. *Cell Metab.*, 23(1), 27–47.
- [34] Argilés, J. M., et al. (2024). Cancer therapy and cachexia. *J. Clin. Invest.* <https://doi.org/10.1172/JCI191934>
- [35] Cancer cachexia as multiorgan failure (2022). *Front. Cell Dev. Biol.*, 10, 960341.
- [36] Cancer connectors: Connexins, gap junctions, and communication (2018). *Front. Oncol.*, 8, 646.
- [37] Single-cell RNA sequencing in cancer research (2021). *PMC*, 7919320.
- [38] The spontaneous remission of cancer: current insights and therapeutic implications (2022). *PMC*, 8271173.
- [39] Common factors among reported cases of spontaneous remission and regression of cancer (2019). *Int. J. Cancer Clin. Res.*, 6(112).
- [40] Walker, S. I., & Davies, P. C. W. (2021). Concepts and applications of information theory to immuno-oncology. *PMC*, 8156485.
- [41] Huang, S., Ernberg, I., & Kauffman, S. (2009). Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. *Semin. Cell Dev. Biol.*, 20(7), 869–876. <https://doi.org/10.1016/j.semcdb.2009.07.003>
- [42] Neijssen, J., et al. (2005). Cross-presentation by intercellular peptide transfer through gap junctions. *Nature*, 434, 83–88.
- [43] Lynch, M., & Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proc. Natl. Acad. Sci. USA*, 112(51), 15690–15695.
- [44] Shi, J., & Teschendorff, A. E. (2018). Quantifying Waddington’s epigenetic landscape: a comparison of single-cell potency measures. *Briefings in Bioinformatics*.
- [45] Teschendorff, A. E., & Enver, T. (2017). Single-cell entropy for accurate estimation of differentiation potency from a cell’s transcriptome. *Nature Communications*, 8, 15599.
- [46] Malta, T. M., et al. (2018). Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell*, 173(2), 338–354.e15.
- [47] Curtius, K., Wright, N. A., & Graham, T. A. (2018). Evolution of Premalignant Disease. *Cold Spring Harbor Perspectives in Medicine*, 8(4), a026542. <https://doi.org/10.1101/cshperspect.a026542>

- [48] Shariff, J. A., & Zavras, A. I. (2015). Malignant Transformation Rate in Patients Presenting Oral Epithelial Dysplasia: Systematic Review and Meta-Analysis. *Journal of Oral Diseases*, 2015, 854636. <https://doi.org/10.1155/2015/854636>
- [49] Martincorena, I., & Campbell, P. J. (2015). Somatic mutation in cancer and normal cells. *Science*, 349(6255), 1483–1489. <https://doi.org/10.1126/science.aab4082>
- [50] Reynolds, S. M., et al. (2017). The ISB Cancer Genomics Cloud: A Flexible Cloud-Based Platform for Cancer Genomics Research. *Cancer Research*, 77(21), e7–e10. <https://doi.org/10.1158/0008-5472.CAN-17-0617>

A Code Availability and Reproducibility Assets

All executable code for this manuscript is maintained in the companion repository: <https://github.com/MosesRahnama/cancer-paper-repository>

The manuscript intentionally omits embedded Python listings to keep the paper focused on theory, evidence, and results while preserving full computational reproducibility in versioned scripts.

- **Core simulation demo:** cancer-paper-repository/run_simulation.py
- **Multi-therapy scenario runner:** cancer-paper-repository/examples/run_full_scenario.py
- **Phase-2 extracted operator sweeps:** cancer-paper-repository/run_phase2_extractions.py
- **TCGA data extraction:** cancer-paper-repository/experiments/tcga/tcga_extract.py
- **TCGA correlation analysis:** cancer-paper-repository/experiments/tcga/tcga_analysis.py
- **Multi-cancer expansion:** cancer-paper-repository/experiments/tcga/tcga_multicancer_analysis.py
- **Tumor vs. normal comparison:** cancer-paper-repository/experiments/tcga/tcga_tumor_normal.py
- **Immune subtype classification:** cancer-paper-repository/experiments/tcga/tcga_immune_subtypes.py
- **Survival analysis:** cancer-paper-repository/experiments/tcga/tcga_survival.py
- **Stage-stratified analysis:** cancer-paper-repository/experiments/tcga/tcga_stage_analysis.py
- **Generated outputs:** cancer-paper-repository/results/ (JSON + figures used in this manuscript)