Patterns

CellPress
OPEN ACCESS

Review

A review of dynamical systems approaches for the detection of chaotic attractors in cancer networks

Abicumaran Uthamacumaran[1,*]
[1]Concordia University, Montreal, QC, Canada
*Correspondence: a_utham@live.concordia.ca
https://doi.org/10.1016/j.patter.2021.100226

THE BIGGER PICTURE Cancers are complex diseases orchestrated by the nonlinear feedback loops between networks of genes and proteins. Cancer stem cells (CSCs) are a subset of cancer populations driving their self-renewal, therapy resistance, heterogeneity, phenotypic plasticity, and recurrence. Identifying the minimal set of molecular networks driving cancer stemness remains an intractable problem in oncology. The time series analysis of protein flows and gene expression dynamics in cancer cells may provide a rich repertoire of patterns, which currently remain uninvestigated. These patterns are universally defined causal structures observed in the state space of dynamical systems. We refer to these causal patterns as attractors. It is proposed herein that CSCs may be chaotic attractors in their signaling state space. If CSCs are chaotic attractors, a causally deterministic pattern drives their emergence, instability, and irregular signaling dynamics. This review discusses various algorithms for identifying such chaotic attractors in cancer time series datasets. The implications of these algorithmic approaches in assessing targeted therapies and precision oncology are demonstrated by implementation on pediatric brain cancer datasets. Furthermore, the detection of chaotic attractors in the state space of cancer stemness networks are suggested to pave the reprogramming of CSCs to benignity.

SUMMARY

Cancers are complex dynamical systems. They remain the leading cause of disease-related pediatric mortality in North America. To overcome this burden, we must decipher the state-space attractor dynamics of gene expression patterns and protein oscillations orchestrated by cancer stemness networks. The review provides an overview of dynamical systems theory to steer cancer research in pattern science. While most of our current tools in network medicine rely on statistical correlation methods, causality inference remains primitively developed. As such, a survey of attractor reconstruction methods and machine algorithms for the detection of causal structures applicable in experimentally derived time series cancer datasets is presented. A toolbox of complex systems approaches are discussed for reconstructing the signaling state space of cancer networks, interpreting causal relationships in their time series gene expression patterns, and assisting clinical decision making in computational oncology. As a proof of concept, the applicability of some algorithms are demonstrated on pediatric brain cancer datasets and the requirement of their time series analysis is highlighted.

INTRODUCTION

Complex diseases, such as cancers, are dynamical systems. The lack of time series cancer transcriptomics and molecular profiling imposes a fundamental barrier in advancing oncology. Then, the question arises, what kind of patterns emerge in time series cancer datasets that cannot be inferred from the currently predominant static methods? There are universal patterns observed in the state space of dynamical systems referred to as attractors. Identifying these attractors in cancer signaling and cancer processes will be the primary intention of this review. Thus, a great emphasis is placed on the science of such patterns through the lens of complexity, networks, nonlinear dynamics, and chaos since the paradigm of complex systems remains relatively unfamiliar to cancer researchers. Given the breadth of these complex concepts, a glossary of essential terms and a summary table (Table 1) of the tools discussed per section are provided, as well.

Table 1. Summary table of algorithms and tools for time series cancer modeling

Section	Algorithm or technique	Description
2	Lyapunov exponents (λ_L)	A characteristic set of exponents used to characterize the rate of separation of trajectories in a dynamical system. If $\lambda_L>0$, there may be a chaotic attractor.
2	Frequency spectra	The frequency decomposition of a spatial signal. Various algorithms, such as fast-Fourier transforms, are applicable to convert the time series signal traces to a frequency spectrum. A broad frequency spectrum in the time trace of a signal may be an indicator of chaotic dynamics.
2	Fractal dimension	A statistical index of self-similarity and complexity. The box-counting algorithm, wavelet analysis-based methods, or multifractal analysis (when more than one fractal dimension exists) are used to compute the fractal dimension.
4	Master equation	A way of describing gene expression dynamics as the evolution of a probability function. The Fokker-Planck equations and Gillespie algorithms are approximations of the Master equation used in the stochastic modeling of cell fate transitions.
4	Boolean networks	Discrete models of biological networks. Most applicable in the study of gene expression network dynamics, wherein each gene can be on or off. However, continuous analogs exist.
5	Reaction-diffusion equations	Mathematical models used to describe pattern formation in chemical systems. Examples of applicability include tumor pattern formation and cancer stem cell differentiation. Chaotic attractors may emerge in these equations in a regime referred to as chemical turbulence.
6	Computational simulations	Simulating the differential equations characterizing gene expression or protein oscillation dynamics may provide a computational tool for identifying chaotic attractors. Simulations must be paired with experimental data to infer such complex patterns.
7	Network science	Various network visualization tools consisting of machine learning pipelines are discussed. These tools combine various algorithms to infer regulatory networks from single-cell datasets.
8	Convergent cross mapping (CCM)	A method of attractor reconstruction using Whitney's theorem and Takens's theorem (time-delay coordinate embedding).
9	Entropy	A powerful information theoretic in network science used as a measure of chaotic flows in dynamical systems.
9	Waddington landscape reconstruction	The general methods to reverse engineer the transcriptomic state space as an energy (epigenetic) landscape are discussed. scEpath was demonstrated as an example of such algorithms.
10	Deep learning neural networks	A type of multi-layered artificial neural networks. They are among the most powerful algorithms in machine learning capable of detecting complex patterns from empirical datasets.
10	Recurrent neural networks (RNNs)	A type of artificial neural networks used in time series analysis. Reservoir computing, a type of RNN, can be used to find the Lyapunov exponents of large-scale empirical datasets, from which chaotic attractors could be reconstructed.
11	Kolmogorov complexity, $K(s)$	$K(s)$ is the length of the shortest description of the data object in program space. Lossless compression algorithms and the block decomposition method (BDM) are available to approximate the $K(s)$ measure. BDM is the most robust estimate of a graph network's complexity available. Algorithmic information dynamics is the branch of computational science using $K(s)$ approximation algorithms to study complex networks.

The emerging paradigm of complex systems is essentially the study of patterns. Therefore, the term pattern science is interchangeably used with complexity science. Given the limited space allocated to this paper, the essentials of these concepts alone are introduced. Many network inference, clustering, and visualization algorithms useful in the dissection of cancer networks have been discussed in detail in Uthamacumaran.[1] However, the intuition for nonlinear dynamics and chaos were absent therein. This paper was written to serve that purpose, to steer the readers to develop an insight for pattern science and complex systems approaches (algorithms, simulations, and mathematical models) when investigating the state-space dynamics of cancer networks, given their time series datasets are made available.

Many of the algorithms discussed here not only sow insights into time series analysis of cancer processes but are in general, applicable to single-time frame (static) gene expression datasets. These algorithms are highlighted as tools for identifying personalized, targeted therapies in precision oncology. That is, comparing the single-cell gene expression profiles of healthy cells and contrasting them with those of the cancer cells of a patient, allows us to identify genes and protein networks that are abnormally expressed in cancer populations. The identified patterns serve as potential markers for (a combination of) target-specific therapeutic interventions upon further screening.

Pediatric cancers are used throughout the paper as a model-system to illustrate the algorithms and concepts, and direct scientists toward navigating these disease networks. By including

an implementation of some of the discussed algorithms on the patient-derived single-cell RNA sequencing (scRNA-seq) count matrices from Neftel et al.[2] and Richards et al.,[3] the clinical relevance of the presented tools, and the insights of the paper are strongly emphasized. Thus, the patient-centered section on cancer networks (section "Cancer networks") is to be treated as the heart of the paper. The molecular networks and patterns discussed therein will steer pattern science toward clinical oncology.

The "single-patient sample" results shown herein correspond to the scRNA-seq data of the pediatric glioblastoma multiforme (GBM) patient sample BT1160 from the study by Neftel et al.[2] (N = 335 cells, 16,488 genes). The "few pediatric GBM samples" results correspond to six patient samples with N = 846 cells analyzed. The "few adult GBM samples" correspond to three adult patient samples with N = 409 cells analyzed from the count matrix found in Neftel et al.[2] Finally, the adult glioblastoma stem cell (GSC) analysis correspond to sample BT127_L (N = 500 random cells) from the scRNA-seq counts obtained from Richards et al.[3] The source datasets (gene expression count matrices) for these data used to demonstrate some of the algorithms herein are available as .csv and .txt files in the Broad Single Cell Portal links provided in the data and code availability section at the end of this review. The source codes for all discussed algorithms are available in the links provided at the end of the paper and their corresponding citations. The various algorithms and approaches illustrated are still simplifications by themselves. However, when combined, they provide a powerful computational platform for the realistic modeling of cancer complexity from empirical data.

Allow me to encapsulate the tumor architecture of these patient samples in brief. The following terms will be used throughout the paper and must be clarified for non-biologists deciding to study cancer dynamics and complexity. In a nutshell, a cellular subpopulation with stem cell properties and neural stem cell marker expressions, including nestin and CD133, can be identified in brain tumors, referred to as cancer stem cells (CSCs). Self-renewal, high proliferation, and pluripotent potential are considered as the hallmarks of embryonic stem cells. CSCs express most of these stem cell properties unlike the well-differentiated cancer cells (see Glossary). CSCs are associated with cancer differentiation, resistance to therapies (chemo- or radiotherapy, immunotherapy, etc.), and recurrence.[4] Notably, the bulk of a GBM lesion consists of both CSCs referred to as GSCs (that are very rare) and the derived progeny of heterogeneously differentiated cancer cells (that form the essential of the tumor mass). However, many other cell types, such as ependymal cells, astrocytes, microglia, infiltrating immune cells, and healthy neurons, as well as the extracellular matrix and associated vasculature, form a complex interconnected system that supports the tumor, and nourishes and preserves the CSC properties.[4] We refer to this microenvironmental unit as the stem cell niche. Cancer stemness and phenotypic plasticity are maintained via both cell self-autonomous molecular networks and a dynamic interaction with their stemness niches. When their ecological relationships (i.e., predator-prey dynamics) are considered, this tumor microenvironment (niche) is referred to as an ecosystem. The unique microenvironment in the cancer stem niche sustains potency and self-renewal. Reconstructing the state space of cancer stemness networks will be a recurring central theme of the paper.

The overall intent of this communication is to direct the readers toward a basic intuition for the causal patterns that emerge in the state-space dynamics of cancer signaling networks. The methods to detect these patterns can be applied to any cancer process, if analyzed in time series. Some examples of such processes include cancer gene expression dynamics, the epigenetic landscape reconstruction of cancer (stem) cells, and the time-lapse imaging of intracellular protein-mediated pattern formation (i.e., reaction-diffusion systems). The primary focus of the paper would be to address the physical framework and algorithms to detect such causal patterns in cancer gene regulatory networks (GRNs) and infer their corresponding cell fate transitions on the epigenetic landscape. The epigenetic landscape is loosely defined as a potential energy projection of the state-space reconstructed from the single-cell transcriptomics of cancer cells. The epigenetic landscape reconstruction of cancer cells provides a robust tool to visualize cancer processes and decision making, at a single-cell resolution. Examples of these processes include the developmental landscape of CSCs undergoing differentiation, and the cell state transitions in cancer populations under therapy/drug-induced perturbations or in response to microenvironmental cues. With time series single-cell transcriptomics, these processes can be visualized as *dynamical systems* and the trajectories of cell states may be conceived as complex, causal patterns (attractors). The tools to classify, visualize, and experimentally detect these patterns from time series, if they exist, are hereby presented. However, it must be emphasized again that such approaches to time series modeling and attractor reconstruction are virtually absent in cancer research.

Cancers are complex systems with multi-scale processes.[1] In principle, attractors can emerge in any cancer process at any scale when treated as a dynamical system (i.e., time series quantification). However, two main categories of cancer processes will be discussed here in the context of attractor dynamics: (1) cell state transitions and (2) pattern formation. The first, cell state transitions, imposes that cell states themselves are attractors in the signaling state space of their molecular networks. The complex dynamics of the signaling pathways coordinating cell state transitions are best defined as attractors in state space. The time traces of these signals can be embedded on state space using Takens's theorem, and the various algorithms discussed herein can be applied to accurately reconstruct the underlying attractor. The universality of attractor dynamics implies that even the signaling of a single gene or single protein oscillation can form attractors in state space. The identified cellular signals responsible for the emergence of these attractors, such as stimulants or inhibitors, growth factors, transcription factors (TFs), cytokines, and other intracellular signals, would then serve as the control parameters of the attractor dynamics (these terms will be explained in the next section, for now the biological context is presented). An example would be the theoretical models of single TF oscillations mapped by Heltberg et al.,[5] exhibiting the self-organization of strange attractors (beyond a critical threshold). The key concept to understand here is the universality of attractors

in dynamical systems. That is, attractors can occur at any scales, from a single protein oscillation in a single cell, to a single cell's gene expression signaling or protein flow density dynamics in a bulk population of cancer cells. The trajectories of cell states on a Waddington epigenetic landscape are attractors, and so are the individual oscillators (genes and proteins) of the molecular networks driving them.

The second category, pattern formation, comprises chemical oscillations that govern the phenotypes of cancer cells. Examples include the differentiation of CSCs to mature (well-differentiated) cancer phenotypes. Such patterning processes can be well-described using reaction-diffusion systems as discussed later. The time-lapse imaging of chemical dynamics and protein oscillations during patterning processes provide empirical data that can be modeled by pairing with algorithms and simulations to identify attractor patterns. An example of this includes the chaotic attractors identified in the oscillations of the bacterial Min protein by Halatek and Frey,[6] which were shown to exhibit diffusion-mediated spatiotemporal chaos (i.e., chemical turbulence). The mammalian equivalent of Min proteins, Par proteins, are central role-players in the cell division and differentiation of cancer (stem) cells. Hence, if similar findings of chaotic attractors are observed in Par protein flows within cancer cells, it may be a key driver of tumor complexity and a control parameter to fine-tune cancer cell division dynamics.

Detecting chaos in cancer cellular systems implies finding the parameters that make the system behave "unpredictably." If chaos did not exist in cancer biology, it would be governed by stochastic processes, wherein their complex behaviors and patterns can only be described as probabilistic outcomes. However, if chaos is detected in these processes, short-term time series predictions are possible and there are causal structures (attractors) governing their behaviors. That is, chaos makes the time series of deterministic systems appear random but exhibit intrinsic patterns that govern the collective dynamics. The dynamics of the system are then bound to these causal structures that can be fine-tuned to control the chaotic system toward a more regular state (stable phenotype). Mathematical models and simulations of different cancer microenvironments have shown the emergence of chaotic attractors in cancer dynamics. However, their experimental validation remains untested.

It must be noted that, in the nonlinear sciences, one often distinguishes chaos from "order." This phrasing is extremely common in literature, but it is technically incorrect. For example, consider supersymmetric theory of stochastics (STS), a mathematical theory of stochastic differential equations (SDEs). STS provides a mathematical framework to explain the complex dynamics in systems exhibiting spontaneous dynamic long-range order (DLRO). Some examples of DLRO systems include self-organized criticality, collective dynamics, pattern formation, turbulence, and chaos.[7] It has recently been shown in the works of Ovchinnikov[7] on STS that the phase-transition from periodicity to chaos corresponds with the spontaneous breaking of (topological) de-Rham supersymmetry, and so, as the symmetry-broken phase, chaos is counter-intuitively the "ordered" phase of dynamical systems.[7] Moreover, a pioneer of complexity, Prigogine, would define chaos as a spatiotemporally complex form of order. As such, it is technically more accurate to distinguish chaos from "regularity" or "periodicity," rather than "order." Therefore, a chaotic cell state refers to an unstable phenotype which exhibits irregular (aperiodic and fluctuant) signaling dynamics, while a regular cell state refers to a stable phenotype with equilibrium-like dynamics, herein. Unlike a stochastic system, the irregular signaling of a chaotic system is bound to causally deterministic patterns in phase space, i.e., strange attractors.

Without time series data, cancer processes are limited to only one type of attractors: fixed points (equilibria). Fixed-point attractors are predictable and easily controllable. However, as will be presented, there are four major classes of attractor patterns observed in a three-dimensional dynamical system. Chaotic attractors are the most complex of these four classes. The emergence of chaotic attractors in cancer processes have been suggested as indicators of therapy resistance, cancer relapse, phenotypic plasticity, emergence of aggressive phenotypes, and metastatic invasion. Chaotic dynamics at the level of protein and gene oscillations may also confer dynamical heterogeneity, and increased survival (adaptation) in individual cell states, by allowing them to withstand extreme environmental conditions due to their large signaling fluctuations.[5] As such, it is further suggested here that chaotic attractors may be signature hallmarks of cancer stemness. For this reason, the first half of the paper strongly builds on the theoretical framework of nonlinear dynamics, chaos, and complexity to progress toward the algorithms needed to detect such patterns in cancer signaling.

CHAOS AND COMPLEXITY

Chaotic systems are ubiquitous in nature. The heat flows of coffee cups, protein-mediated pattern formation,[6] the dynamics of complex networks, the Hodgkin-Huxley model of neuronal oscillations,[8,9] the fluctuations of financial stock markets,[10] the patterns of fluid turbulence, the growth processes of biological populations,[11] ecological predator-prey dynamics,[12] aortic blood flows,[13] cardiac rhythms,[14] and the onset of complex dynamical diseases,[15,16] are some of the many illustrations of complex systems exhibiting chaotic behaviors. Despite its prevalence, the detection and long-term time series forecasting of chaotic systems remains an intractable problem and is widely considered the pinnacle of scientific progression. Tumors may be models of chaos exhibiting multifractal geometries and complex dynamics.[17,18] Of the numerous applications of causality inference in chaotic systems, precision oncology (medicine) may inarguably be the most benefited in times ahead of us. As such, this review intends to paint tumor ecosystems on the canvas of nonlinear dynamics, as complex systems fundamentally driven by chaotic processes.

Let us begin with a simple thought experiment. Imagine an abstract expressionist art piece, interconnected by the dripping and splashes of a myriad of colors. At glance, the heterogeneous patterns may seem stochastic due to its static appearance, i.e., apparent randomness. However, if one reconstructed every motion involved in the creative process, frame by frame, a symphony of dynamical expressions is interwoven. That is, in the temporal dimension, one can visualize patches of self-organized patterns, emergent structures, bifurcating multi-fractals, swirling

vortices, and cascading flows of fluid turbulence on the canvas cohering to an undivided whole. Let this painting be a metaphorical representation of a population of cancer cells. The static (single time frame) perspective is an analogy of our current approaches to studying cancer cybernetics, i.e., gene expression dynamics, cell signaling, and pattern formation. The lack of time series data-driven modeling of tumor ecosystems present a causal agnostic description of its complexity. What we are truly interested in is the temporal behavior of tumor ecosystems and their complex multi-scaled processes, which herein is discussed as that which is fundamentally coordinated by chaotic dynamics.

In systems science, one must draw comparisons between individual dynamics and the population dynamics of a nonlinear system. A key example would be to contrast a single cell versus a population of cells or rather the dynamics of a single gene versus that of gene networks. Self-organized patterns can spontaneously develop in nonlinear systems composed of positive and negative feedback loops maintained far from equilibrium. The statistical mechanics of the many nonlinearly interacting constituents in biological networks result in macroscopically irreducible systems. We refer to such emergent systems with multiscale structures (networks) in which the interactions change in time, as complex systems.[19,20] Complex systems generally consist of systems in which the number of particles or degrees of freedom approach a continuum. The physics of fluids, multicellular networks, social organizations, and financial markets are good examples. Due to nonlinear dynamics, complex systems exhibit collective behaviors, which cannot be inferred from the individuals. Cancers are complex systems, consisting of groups of adaptive malignant cells that self-organize in time and space, far from thermodynamic equilibrium.[1,21,22]

In traditional (analytical) physics, natural systems are dealt with using statistical mechanics to describe systems with large degrees of freedom, and dynamical systems theory to describe the changes in time of systems constrained to phase space. However, these tools by themselves are inadequate to describe the emergent scaling behaviors observed in complex systems.[23] In the 1980s, these two disciplines, nonlinear dynamics and nonequilibrium statistical mechanics, merged to the emerging paradigm of complex systems. Complex systems theory, also known as complexity science or simply pattern science, advocates the study of natural systems via the use of algorithms (machine intelligence) and computational physics.[19,24] Pattern science investigates the otherwise analytically intractable, emergent processes and behaviors of complex systems, which are computationally reproducible in program space. The pairing of simulations and computational physics with large-scale, empirical datasets is the modus operandi in pattern science. Pattern science paves the study of cancer ecosystems in a multidisciplinary space, which is referred to as computational oncology.

Some general features of complex systems include: emergence, self-organization, nonlinearity, nonequilibrium dynamics, criticality (i.e., poised between regularity and chaos), abrupt phase transitions, multi-nested, multi-scaled feedback loops, multi-layered (fractal) organization, computational irreducibility, unpredictability, NP-completeness (i.e., a measure of the resources exhausted to find a solution), non-locality (i.e., global features of a system cannot be reduced to local properties), undecidability, and non-analyticity (intractability).[19,20,24] Such notions are inherent to the choices made by the agents in complex systems and more generally, to certain classes of problems in computational complexity classified under the P versus NP umbrella.[24] Thus, complexity theory studies complex systems using algorithmic simulations. Furthermore, it exhibits a universality in algorithms (i.e., algorithms are as complicated as any complex system, and vice versa). The crowning glory of pattern science remains the vision of Laplace's demon; the time series forecasting of chaotic dynamical systems.

A dynamical system can be represented as a set of variables evolving in time. When the outcomes of the system do not change in direct proportion to a change in the inputs, we call it a nonlinear dynamical system. For instance, the cross-talks in physiological dynamics and signal transduction pathways in cellular systems are examples of biological nonlinearity.[16] Formally, a nonlinear dynamical system exhibiting sensitive dependence to its initial conditions and perturbations is a chaotic system. That is, moving or modifying small sections or pieces of a chaotic system can have unexpected, cascading effects. Even small fluctuations can amplify to exponentially large deviations in the system's behaviors. Chaotic systems can be continuous (e.g., the Kolmogorov energy spectrum of fluid turbulence) or discrete (e.g., the logistic map used to study population growth dynamics). The generic form of a continuous-time dynamical system is given by: $\frac{d\overline{X}}{dt} = \overline{f}(\overline{X}, t, \mu)$ where X is the state vector, t is time, μ is the set of control parameters and f is some arbitrary function. For example, the gene expression (count) matrix from an scRNA-seq experiment (as used in algorithms in the latter sections) is a state vector X. With time series analysis, the expression matrix evolves as a dynamical system. The generic form of a discrete-time system is given by a recurrence (difference) equation: $X_{k+1} = f(X_k(\mu))$, where k is the index of iteration for the state vector. When the control parameter μ exceeds a critical value, chaotic dynamics can emerge.[23]

In 1963, Edward Lorenz, Ellen Fetter, and Margaret Hamilton pioneered chaos theory by studying the predictability of hydrodynamic flows in simple mathematical models as an approximation to weather turbulence forecasting.[25,26] In 1969, Lorenz demonstrated the predictability of chaotic flows which preserve many scales of motion, are sensitively dependent to the errors in the small scales building up to large-scale anomalies in measurements.[27] These findings concluded two key insights in the study of chaotic systems: (1) even with reduced errors, we can only forecast the chaotic system's behavior up to certain Lyapunov times, and (2) that deterministic chaos is difficult to observationally distinguish from random systems.[27] However, there are certain statistical properties and topological measures we can extract to infer chaotic dynamics. The general workflow for chaos detection, is first to obtain time series measurements, convert the data into a geometric object via time-embedding, and assess the topological properties of this object/attractor (i.e., Lyapunov exponents, entropies, fractal dimensions, and multifractality).[1]

Generally, dynamical systems depend on certain control parameters, which, when changed, the stability of the fixed point can change, appear, or disappear, or give rise to a new object in its vicinity. An example would be to consider a limit cycle undergoing a Hopf bifurcation toward an unstable chaotic oscillator. However, in most complex systems, the control

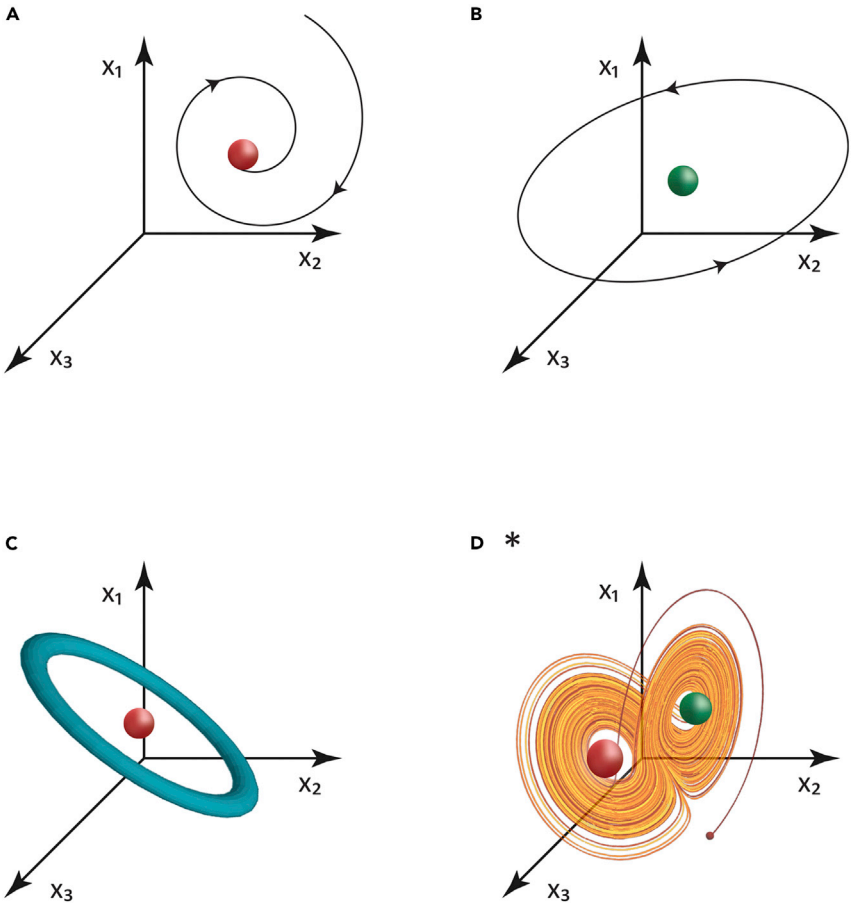

Figure 1. The anatomy of attractors

Attractors are universal causal patterns observed in the evolution of a dynamical system in state space. They represent the fundamental ways in which dynamical systems behave. The four major attractor patterns include: (A) trajectory convergence toward a fixed-point attractor in state space, (B) a limit cycle (periodic attractor), (C) a torus (quasi-periodic attractor), and (D) the birth of a chaotic attractor (strange attractor). Two unstable fixed points are shown for this strange attractor.

parameters and equations of motion may not be self-evident or difficult to infer from experimental time series measurements.[20,24] Despite their apparent randomness, chaotic systems have a deterministic structure (patterns) and an underlying mechanism generating the pattern. Although chaotic processes exhibit irregular and aperiodic oscillations, in principle their causal structures can be experimentally and computationally detected. We refer to these causal structures as attractors of a dynamical system.

Attractors are self-organized structures consisting of a set of points bounded in phase space to which all nearby trajectories of a dynamical system from a neighborhood (basin of attraction) will tend to with time. There are many sorts of attractors coexisting in the phase space of a complex dynamical system (Figure 1). Consider a general one-dimensional dynamical system, $\frac{dx}{dt} = \dot{x} = f(x, \mu)$, with some (set of) order parameter(s) μ. The equilibria of the system correspond to the eigenvalues when $\dot{x} = 0$. These equilibria correspond to the simplest of attractors, i.e., fixed-point attractors. Fixed points can be created or destroyed, or their stability can be changed by varying the order parameter(s). These qualitative changes are called bifurcations. When small changes in the parameter lead to large sudden changes in the state of the system, we call it a phase transition. Phase transitions are seen when nonlinear systems change from mono-stability to bistability and/or multistability (i.e., coexistence of attractors). When the number of dimensions increases from one to three or more, complex attractors may emerge. These dynamics are better visualized in state space, by letting $\dot{x} = -\frac{dU(x,\mu)}{dx}$, represent the gradient of a potential energy U describing the system. Such potential energy (attractor) landscapes can be used to study complex dynamical systems, such as the multidimensional epigenetic landscape of a population of cancer (stem) cells.[1]

As mentioned, the heterogeneity and phenotypic plasticity of tumors are vastly attributed to a dynamic subset of cells within the cancer population known as CSCs. CSCs possess the capacity to both self-renew and generate all cells within a tumor, and are believed to drive tumor recurrence, therapy resistance, phenotypic plasticity, and the emergence of aggressive phenotypes. Cancer cells (both stem and non-stem cells) transition amid a variety of functional and heterogeneous cell fates via the dynamical reconfiguration of their microenvironment-dependent multicellular, molecular networks.[1] Upon certain signaling network configurations, they differentiate toward specific cell fate commitments (phenotypes). One of the mechanisms most widely discussed in both CSC and non-stem cancer cell state transitions is the EMT (epithelial-mesenchymal transition)/MET (mesenchymal-epithelial transition) switch, which allows the interconversion of epithelial and mesenchymal cancer phenotypes.[28,29] The nonlinear, multi-nested feedback loops of many inflammatory signals, including nuclear factor κB (NF-κB), STAT-3, and embryonic developmental signals (morphogens), such as Wnt, TGF-β, Shh (sonic hedgehog), Notch, cytoskeletal networks, etc., orchestrate the EMT programs and CSC/cancer cell state transitions.[28,29] These cell state transitions are dynamical systems.

The statistical mechanics and dynamics of these complex networks at both the individual cell level and collective populations of CSCs can be assessed via quantitative single-cell measurements, including single-cell genomics and proteomics. The empirically reconstructed molecular networks can be projected in phase space to be visualized as an epigenetic landscape. The Waddington's epigenetic landscape is a metaphorical concept to represent cellular decision-making processes during development and differentiation.[30] The epigenetic landscape provides a visualization of developmental processes, such as

the differentiation of stem cells to their coordinated cell fate commitments.[31] The epigenetic landscapes also well apply to modeling the cell fate transitions of a cancer population under some given environmental conditions. For instance, one can map the differentiation of CSCs and dedifferentiation or phenotypic switching of the non-stem cancer cells under a therapy or drug-induced perturbation.

The epigenetic landscape exhibits the cell (often represented by a ball) rolling up hills and down valleys representing basins of attraction. The attractors may represent the distinct cell phenotypes, and the bifurcations toward these attractors represent the cell fate trajectories.[30] The critical points (extrema) of the potential energy surface $U(x, \mu)$ correspond to the equilibria of the dynamical system, where the energy minima denote stability (e.g., mature/differentiated phenotypes), and the maxima are unstable configurations (e.g., stem cell fates). There may be many hybrid and transient cell fates denoted by metastable attractors under epigenetic fluctuations of the landscape. Small perturbations and thermal noise can lead to drastic changes in the number and nature of these critical points, resulting in the emergence (destruction) of (un)stable attractors.[31] Current approaches naively assume that cell fates are static states represented by stable fixed-point attractors (i.e., cells are stuck in a valley of the energy landscape) (see "Control of chaos"). Fixed-point attractors represent stable phenotypes in equilibrium. However, fixed-point attractor can lose stability and bifurcate to complex attractors (Figure 2). Chaotic attractors, the most complex of attractors, may correspond to CSCs and/or aggressive cancer states.

There are several types of bifurcations of a fixed point to consider on the attractor landscape, such as saddle-node bifurcations, pitchfork bifurcations (in symmetric systems), and Hopf bifurcations, to name a few. Andronov-Hopf bifurcations occur at a critical point where the system's stability switches, and periodic solutions/limit cycles emerge. For example, they are observed in the birth of the Lorenz attractor and pattern formation in reaction-diffusion systems. In a system undergoing a Hopf bifurcation, when the control parameters μ are below the critical value needed for bifurcation, fixed points remain on the phase space. Above the critical value, symmetry breaking occurs.[32] Let us assume that we can represent some general dynamical system as an eigenvalue problem of the Jacobian matrix. If the Jacobian matrix of the system shows eigenvalues $\lambda < 0$, these fixed-point attractors are stable. When the parameter's critical value is exceeded, $\lambda > 0$ resulting in unstable point attractors (i.e., the real value of the eigenvalues will be positive).[32] The fixed point(s) is then no longer a steady-state attractor(s), but around it a new object can emerge such a limit cycle where the phase trajectory periodically oscillates around a fixed point. This phase transition from a stationary fixed-point (equilibrium) to a periodic oscillator (limit cycle) attractor is called a supercritical bifurcation. After the supercritical Andronov-Hopf bifurcation, the amplitude (A) of the limit cycle oscillator grows as:

$$A \approx \sqrt{\mu - \mu_c},$$

where the period T of the limit cycle is given by: $T = \frac{2\pi}{Im\lambda_{1,2}}$. The indices correspond to the two possible eigenvalues of the system at that bifurcation point and Im denotes their imaginary component.

The periodic oscillator can further become unstable and bifurcate into a quasi-periodic attractor or an aperiodic oscillator (i.e., chaotic attractor).[32] In quasi-periodicity, the trajectories are constrained to motion on the surface of torus with two separate frequencies. In a chaotic attractor, also known as a strange attractor, the bounded structure is characterized by a fractal dimension. If the system is hyperchaotic (i.e., has more than one positive maximal Lyapunov exponents), multifractal analysis is required (i.e., more than one fractal dimension exists). In such systems, multiple strange attractors can self-organize into multi-nested patterns (i.e., non-isolated). Strange attractors are geometric constructions exemplifying the two basic ingredients for chaotic dynamics, namely, the repeated actions of stretching and folding of a given phase space volume.[32] A classic example of bifurcations in chaos is the hierarchical fractal structures of eddies and vortices breaking down in the Kolmogorov-Richardson energy cascade of isotropic, homogeneous fluid turbulence.

Attractor dynamics allow us to predict the time series behavior of a chaotic dynamical system. However, this realization comes with a caveat, only fixed points can be found analytically and that too only in well-behaved low-dimensional systems. Analytically searching for periodic and chaotic attractors in a dynamical system's state space is related to Hilbert's 16th problem and is NP-hard. The quasi-periodic tori and strange attractors cannot be found analytically and require numerical methods, such as searching heuristics and computational algorithms. The second half of this manuscript will devote to the study of such algorithms. For the time being, let us consider the birth of chaos and strange attractors in a dynamical system.

It may be counter-intuitive to consider how a chaotic attractor exponentially diverges apart from its initial conditions and remains bound to an attracting set within phase space, simultaneously. Strange attractors consist of infinite curves bound to a finite space. The answer lies within the fractal geometry of the structure; the stretching, folding, and reinjection of the chaotic system's trajectory in phase space.[32,33] As a metaphor, think of the kneading of a pastry dough (phase space) into a flaky croissant (the strange attractor). The process of stretching and folding leads to a loss of information about the initial conditions.[32] Think of the set of initial conditions as a drop of food coloring (dye) applied to the pastry dough, the kneading process disperses the dye throughout the dough. These pastries in phase space exhibit some unique signatures, such as a positive characteristic Lyapunov exponent(s), and a fractal dimension that remains invariant under the system's time-evolution.[32] The iconic picture of a strange attractor is the Lorenz butterfly attractor, a causal pattern emerging as the solution of a set of ordinary differential equations describing the Rayleigh-Benard convection model.[25] The birth of a strange attractor commonly occurs via a sequence of period-doubling bifurcations. However, there are different routes to chaos one must consider. These include the following:

(1) Period-doubling bifurcations: starting from some limit cycle (a point in the Poincare section), a period-2 cycle develops as the control parameter increases and keeps repeating. This is also referred to as Feigenbaum chaos or the universal route to chaos, as seen in the logistic

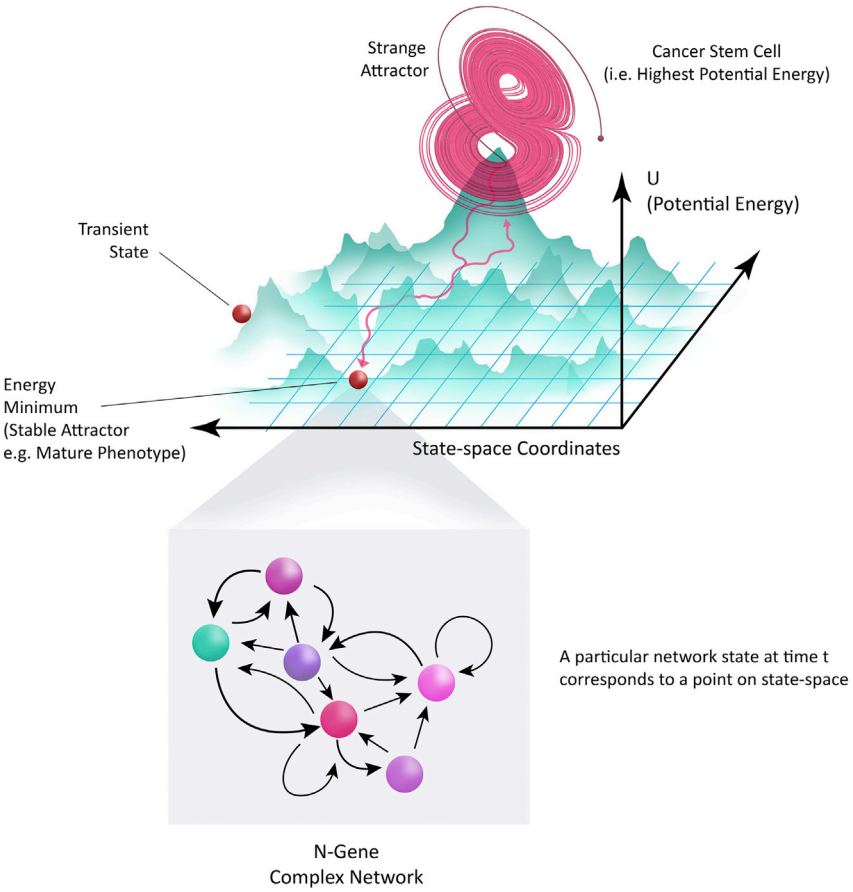

Figure 2. Cancer attractor landscape
The state-space projection of a cancer network is shown as an energy landscape. A particular network state at time *t* corresponds to a point on the state space. The *x*–*y* plane corresponds to state space, whereas the *z* axis denotes the potential energy *U*. The potential energy *U* denotes the relative stability of the cell fates (attractors) in state space. CSCs are populated at the highest peak on the epigenetic landscape, while stable differentiated (mature) phenotypes correspond to fixed-point attractors (energy minima). The interconvertibility (re-programmability) and plasticity of cancer states implies that mature cells may not be fixed-point attractors, as indicated by the red arrow going from the fixed-point attractor (red ball) back toward the strange attractor (CSC).

map equation above a critical threshold of the order parameter.

(2) Quasi-periodicity: the system begins with a limit cycle trajectory of a certain frequency and, as the control parameter changes, a spiral node fixed point turns into a spiral repellor and a stable limit cycle develops around it with a new frequency (i.e., Hopf bifurcation). The ratio of the two frequencies is an irrational number. The quasi-periodic oscillation can lead to chaos as the order parameter further increases. This is also known as the Ruelle-Takens route to chaos, seen in certain types of fluid turbulence.[33]

(3) Intermittency: irregular and aperiodic bursts of chaotic behaviors followed by intervals of regular behaviors. An example is the transitions between laminar and chaotic flows in fluid turbulence.

(4) Transient chaos: the long-term behavior is influenced by the interactions of unstable fixed points and/or unstable limit cycle attractors. For example, the rolling of a dice exhibits chaotic motion until it reaches its result (an equilibrium state).

It remains an open question which type of chaos applies to which dynamical systems and when. Li and Yorke first coined the term *chaos* for a deterministic dynamical system.[34] However, a formal mathematical definition was first given by Devaney.[35] He defined a map $F : X \rightarrow X$ to be chaotic if the following conditions are satisfied:

(1) F has sensitive dependence on X

(2) F is topologically transitive, i.e., there exists a point x^* such that a closure of an orbit O is given by $cl(O(x^*, F)) = X$

(3)The periodic points are dense in X

Conditions 2 and 3 were later shown to imply 1. A diffeomorphism or map F is said to be chaotic on an invariant set A provided that the following conditions are satisfied:

(a) the diffeomorphism F has sensitive dependence on initial conditions when restricted to an attractor A

b) F is topologically transitive

The set A is a chaotic attractor for a diffeomorphism/map F given that the set A is an attractor for F and that F is chaotic on A. The latter was defined by Alligood et al.,[36] where the Lyapunov exponent was established as a computational method to distinguish a chaotic attractor. They defined a chaotic attractor as an attractor such that:

(a) the Lebesgue measure of the basin of attraction of A, W(A), is positive

(b) there exists a point P where $P_0 \in A$, such that there exists at least one positive Lyapunov exponent $\lambda_L(P_0) > 0$ and that all the Lyapunov exponents of the dynamical system are non-zero. If more than one positive λ_L exists at various points, the system exhibits hyperchaos

The second condition also satisfies that the attractor is not a periodic orbit nor a set of fixed points. Distinguishing attractor types are analytically intractable, and computationally an NP-hard problem. There are various techniques to visualize chaotic dynamics and these algorithms should be incorporated into the study of time series cancer datasets when made available. To quantify sensitive dependence on initial conditions, consider two points separated by a very small distance in time, $s(t)$, in phase space, then:

$$s(t) \sim s_0 e^{\lambda_L t}$$

for a small time, t, where the λ_L is the Lyapunov exponent and not to be confused with the eigenvalue symbol used earlier. If $\lambda_L < 0$, there may be a fixed-point attractor, and if $\lambda_L > 0$, there may be a chaotic attractor.[32] If it is zero, there is a limit cycle (i.e., a stable periodic orbit). A negative largest Lyapunov exponent may also correspond to a limit cycle. As discussed, when a system's control parameter increases beyond its critical values, the frequencies of the oscillations multiply and bifurcate. As the system transitions to completely developed chaos, infinite frequencies are observed within the period-doubling bifurcations. The inverse of the maximal characteristic Lyapunov exponent(s) is the predictability horizon of the system, i.e., we can only predict a chaotic system to several multiples of its Lyapunov time.[32]

Fractal behaviors are statistical signatures of chaotic systems-they are the geometry of chaos.[37] We can characterize a strange attractor by its fractal (Lyapunov) dimension, a measure of statistical self-similarity and roughness of a complex structure. Various numerical methods, such as the box-counting algorithm, exist for the estimates of the Lyapunov dimension of attractors in dynamical systems. The most popular method to compute the attractor dimension remains the Kaplan-Yorke K dimension:

$$K = j + \frac{\sum_{i=1}^{j} \lambda_i}{|\lambda_{j+1}|},$$

where j is the index for which $\sum_{i=1}^{j} \lambda_i \geq 0$ and λ denotes the Lyapunov exponents.[38] Such measures allow the attractor patterns of chaotic systems to be described using phase space portraits. If the system exhibits hyperchaos, multifractal analysis is required which involves defining statistical measures, such as the Hölder exponents and Hurst exponents (for time series). Multifractal analysis will not be discussed herein, and readers are directed to Uthamacumaran[1] for further details of its cancer applications.

An alternate approach to studying chaos is to visualize the power spectral density (frequency space) of the system wherein the Fourier transform picks out the frequencies from the time series datasets. The frequency/power spectrum of chaotic oscillators, despite its complexity, demonstrates that systems exhibiting deterministic chaos often can be treated as a broadband with pronounced peaks at the natural frequency of the system.[33,39] The spectral properties of chaos reveal irregular harmonics and aperiodicity where a small amplitude (external) driving force can produce a large-amplitude response in the nonlinear oscillators, i.e., resonance.[33] There may be hidden spikes (sudden transitions) in the intermittent fluctuations. These resonances can accumulate toward cascading effects and dissipative dynamics. Therefore, a broadband spectrum in the frequency spectra (Fourier transform) of gene/protein signals' time traces can be an indicator of chaotic attractors. Multifractal analysis is required to further understand the complex dynamics and fluctuations of these time traces.

To illustrate an example, consider again the Kolmogorov energy spectrum for the three-dimensional isotropic turbulence of a homogeneous, incompressible fluid. The spectrum of fluid turbulence tells a story encapsulating the competition between the velocity and viscous forces encountered by a fluid, consisting of three distinct phases: (1) the stirring/creation of eddies and vortices, (2) the inertial phase denoting the transfer and transport of energy from large-scale to lower-scale vortex structures, and lastly (3) the dissipative phase embarking the destruction of vortices (i.e., viscosity dominant). The frequency (energy) spectrum of experimental fluid turbulence must be continuous, and broadband, in which only a small number of frequency modes produce the spectrum. However, the number of modes in the transition is an open question.[33,39] The period-doubling in frequency space shows that each time the period doubles, the fundamental frequency halves.[33,39] The detection of resonances in the frequency spectra remains a field of study in the time series forecasting of turbulent fluids.[33] In contrast, the phase-space dynamics of fluid turbulence exhibit a kaleidoscope of chaotic solutions—i.e., strange attractors.[33,37] In fact, the term strange attractors was first coined by Ruelle and Takens to describe the chaotic solutions seen in fluid turbulence models.[33] Fluid turbulence is an ideal model of complex systems to develop some intuition for chaotic dynamics and survey experimental techniques for its detection. These techniques can be adapted to the study of cancer processes when treated as dynamical systems.

There are a few additional caveats to consider with these approaches. The general methods for chaos detection (i.e., positive Lyapunov exponents and fractal dimension) defined above may not be directly applicable when equations-free or parameter-free systems are considered. There could be high levels of noise smearing the data and dimensionality constraints. Denoising algorithms are needed to pre-process the time series, followed by time-delay coordinate embedding algorithms, prior to their use. In the case of discrete-time models, there are Boolean derivatives and the spreading velocity of a perturbation (local damage spreading) available as analogs of the Lyapunov exponents.[40] There are also attractors that may be non-isolated and entangled with other attractors and, hence, ambiguous in their detection. Multistable systems are often enriched with the coexistence of these so-called hidden attractors. The basins of attraction of the hidden attractors do not intersect the (un)stable fixed points (if it exists) and nonlocally influence the dynamics of nearby attractors.[41] The entanglement of attractors in chaotic dynamics, and the infinitely many unstable periodic orbits/fixed points embedded in a single chaotic attractor, indicate that chaotic systems can be fine-tuned by controlling their complex dynamics. The study of steering a chaotic system by regulating its attractor dynamics is a powerful method in the control of chaos.

One can foresee that attractors draw parallels to oscillators, given that a dynamical system consists of periodic and/or aperiodic oscillations varying in time. A genetic oscillator is a simple model to describe how a gene's product, protein concentration,

oscillates. For instance, many essential genes of the circadian clock, embryonic development, tissue patterning, cell division, and metabolic networks can be treated as nonlinear coupled oscillators. As a mathematical trick, biophysicists often model genes and cells as nonlinear oscillators. Interactions between different coupled nonlinear genetic oscillators with a driving force can exhibit chaotic oscillations.[42] Many oscillatory genes, such as NF-κB, p53, and Wnt, involved in embryonic development, cell division, and morphogenesis are dysregulated in cancers acting as the precursors for tumor pattern formation.[43] These genes are required to be expressed at precise times and locations during developmental processes. Their aperiodicity would be implied in the emergence of pathological cell fates, such as cancer. Thence, if some of these developmental genes are identified as the master regulators of cancer stemness, the detection of periodic oscillators/limit cycles in the attractor space of gene networks could reduce the search space for candidate chaotic attractors (if they exist). However, as discussed, there are various routes to chaos to consider. Therefore, we require a more rigorous set of algorithms and experimental methods to capture strange attractor dynamics in cancer time series datasets.

Finally, the biological relevance of the various discussed attractors in cancer dynamics must be clarified. As stated, fixed-point attractors in cancer gene expression state space denote stable equilibrium states, such as well-differentiated, mature phenotypes. Thereby, in principle, if a patient's cancer cells are mapped as fixed-point attractors on the transcriptomic state space, they are likely easier to treat and a well-defined, relatively predictable molecular network drives the cancer state. Finding targeted therapies should be feasible from static (single-time frame) molecular profiling. However, even mature cancer cells exhibit phenotypic plasticity and can dedifferentiate. A good example is the EMT transition and its reverse-process observed in most cancers during metastatic invasion. Cell state transitions are dynamic processes. At the other end of the attractor spectrum, strange attractors may be associated with aggressive cancer phenotypes and complex adaptive behaviors, such as tumor relapse and therapy resistance.[44–47] As such processes are governed by cancer stemness, it is further suggested here that CSCs may be strange attractors governed by the chaotic dynamics of certain molecular networks. Unlike the currently predominant "static" picture of cancer dynamics (i.e., fixed-point attractors), strange attractors cannot be experimentally detected without time series analysis. A video of many time-slices is needed. If chaotic attractors are experimentally detected in cancer signals, the various algorithms as discussed above and those in the latter sections, can be implemented on experimental datasets paired with computational simulations, to infer the chaotic signals that may confer these adaptive cell states.[5]

Having laid the basic foundations of attractor dynamics, we can now explore its relevance in untangling cancer complexity. Assuming the interdisciplinary pattern scientists assessing this paper may not necessarily be experts in cancer biology, allow me to first introduce a brief background of pediatric cancers as a model system of tumor complexity. The complex molecular patterns governing pediatric cancer networks will provide the systems thinkers the necessary background to investigate the molecular networks driving cancer dynamics and pattern formation from a dynamical systems perspective. The gene and protein networks discussed in the following section can then be modeled using the various approaches discussed later, from differential equations to network state-space reconstruction algorithms, with time series data.

CANCER NETWORKS

Prior to addressing the mathematical framework of cancer GRNs, some context to the cancer stemness networks must be provided. Let us consider two cases of highly morbid pediatric brain tumors, glioblastoma multiforme (GBM) and medulloblastoma (MB). These lethal tumors are known for their therapy resistance, immunosuppression, migration/invasive dynamics, and intra/inter-tumoral heterogeneity, making them some of the most difficult cancers to treat.[4] Epigenetic alterations (that affect DNA transcription [i.e., gene expression] without modifying its nucleotide sequence) serve as characteristic molecular patterns distinguishing high-grade pediatric gliomas from adult-gliomas. Some examples include dysregulated multiprotein complexes/enzymes involved in direct DNA methylation, and histone H3.3 mutations, such as K27M (i.e., a mutation that changes the lysine [K] at position 27 to a methionine [M]) (prevalent in pediatric diffuse intrinsic pontine gliomas [DIPG]), K36M, and G34V/R, to name a few.[48,49] However, epigenetics is ambiguous and a poorly understood complex system. For instance, while the H3K27M mutant is a glioma phenotype associated with a lack of methylation, recent studies show mammalian target of rapamycin (mTOR) complexes promote the hypermethylation of H3K27 (in the form of trimethylated lysine 27 of histone H3 [i.e., H3K27me3]) that influence tumor progression and malignancy.[50] For simplicity, the epigenetic dysregulations, which serve as hallmarks of pediatric brain cancers, will not be discussed here. Rather it is implied that the dynamics of the discussed transcriptional network modules are inseparable from the aberrant changes in their epigenetic networks.

CSCs are the central drivers of tumor complexity and are herein predicted to be strange attractors of the tumor Waddington landscape. A central question then arises, can CSCs be reprogrammed to benign-like plastic states? This would be a fruit-bearing direction for clinical oncology. Regenerative biology was revived when it was shown that fibroblasts could be reprogrammed into embryonic stem cell-like fates—called induced pluripotent stem cells (iPSCs)—by introducing into them a minimal set of essential TFs: Oct4, Sox2, Klf4, and Myc (OSKM).[51] Changes of a very few TFs can transform a differentiated cell into a stem cell (i.e., transdifferentiation).[52] Altering the timing and concentration of specific TF expression programs steers the cell fate choices during stem cell differentiation.[52] GSCs are at the apex of an entropic hierarchy (i.e., maximal entropy state) in their epigenetic landscapes.[53] Maximal entropy state implies the maximum number of gene network configurations correspond to that cell state. Chaotic oscillations of cell states and intrinsic transcriptional fluctuations were suggested as the mechanism responsible for their heterogeneity, plasticity, and therapy resistance.[53] However, the term chaotic is mostly used in a qualitative manner in the description of GSC complexity.[53]

Recent advancements in high-throughput multi-omics sequencing have identified the essential TF networks and

genetic programs regulating CSCs. For example, Suvà et al.[54] identified a core set of neurodevelopmental TFs (*POU3F2*, *SOX2*, *SALL2*, and OLIG2) essential for the propagation and stemness of the GBM stem cells (GSCs). These TFs are part of the Yamanaka circuitry critical for the maintenance of self-renewal and pluripotency in embryonic stem cells. In principle, mature/differentiated GBM phenotypes could be reprogrammed into GSCs by the induction of these four TFs.[54] More recently, the YAP and TAZ transcriptional coactivators were shown to be master regulators of GSC stemness, irrespective of the GBM molecular subtypes.[55] Their expressions are highly intertwined with the Hippo signaling pathway and p53 expression. However, the cybernetics of these stemness networks remains poorly understood. Do the highly entropic CSCs at the peak of the epigenetic landscape roll down their valleys (energy minima) and get stuck in them, or can they climb uphill from the valleys and oscillate in between valleys as strange attractors? Time series analysis is required to address these fundamental questions.

While the complex dynamics of GBM cells remain ambiguous, distinct molecular patterns have been identified for the self-organization of GBM's transcriptional heterogeneity. Combining scRNA-seq and lineage tracing in glioblastoma models, an analysis of TCGA bulk specimens and scRNA-seq from both adult and pediatric GBMs were performed.[2] The study concluded that malignant cells in glioblastoma exist in four distinct cellular states influenced by the copy-number amplifications of the CDK4, EGFR, and PDGFRA loci and mutations in the NF1 locus.[2] These correspond to the Verhaak classification—the clinically relevant molecular subtypes of GBM, namely, proneural, classical, mesenchymal, and neural subtypes, respectively.[56]

The four distinct cell types also demonstrate microenvironment-dependent phenotypic plasticity and can undergo cell fate transitions under environmental fluctuations/perturbations.[57] Therefore, transcriptomic classifications of supposedly well-differentiated GBMs reveal that cancer cell fates exhibit a high degree of phenotypic plasticity, and targeted therapies are required for patient-specific molecular patterns. For instance, the mitochondrial GBM subtype, which displays the most favorable clinical outcome, relies exclusively on oxidative phosphorylation for energy production, whereas the glycolytic subtype relies on aerobic glycolysis (Warburg effect), amino acids, and lipid metabolism. While the mitochondrial subtypes are vulnerable to oxidative phosphorylation inhibitors, the glycolytic subtypes are unaffected.[58] Non-selective (non-target-specific) therapy regimens can allow mature, differentiated cancer cells of a patient's tumor population to revert to a highly plastic, CSC state.[57] These findings further blur the difference between a CSC and a well-differentiated cancer cell.[57] More importantly, they question the existence of a global pattern (attractor) governing the dynamics of GBM stemness, plasticity, and its differentiated heterogeneity.

In a recent study, scRNA-seq was performed on >69,000 GSCs cultured from the tumors of 26 adult patients.[3] The study found that GSCs mapped along a transcriptional gradient of two distinct cellular states resembling the normal neural development and inflammatory wound-healing response.[3] Various clustering algorithms were used to identify cell states (phenotypes) with similar gene expression patterns. Often these techniques infer maximally connected differentially expressed gene networks using graph-based clustering methods and correlation metrics. Cluster analysis, followed by measuring the Jaccard index and principal-component analysis (PCA), revealed an elevated expression of mesenchymal-related genes and enrichment of pathways associated to inflammation and immune cell activation, as well as NF-κB and STAT signaling in one cell state.[3] These mesenchymal subtypes may promote the pathways needed for GSCs to sustain growth and self-renewal in hypoxic tumor microenvironments. The other cell state, with lower PCA loading, consisted of cells associated with genes and signaling pathways related to gliogenesis and neural developmental markers, such as PTPRZ1, ASCL1, and SOX2, and those highlighted by the expression of oligodendrocytic (e.g., OLIG1, OLIG2), astrocytic (e.g., CLU, APOE, S100B), and neuronal (e.g., STMN3) neurodevelopmental lineage markers.[3] These molecular subtypes resemble the four major Verhaak classifications of GBM. Whether these two-cell states correspond to a single dynamical structure or strange attractors, such as the Lorenz system, can only be verified with time series analysis. The detection of a global attractor governing cancer stemness networks would provide a mechanism to fine-tune tumor behaviors and dynamics.

Similar phenotypic clustering by molecular profiling have been paved for MB.[59,60] Four major distinct molecular subtypes of MB, namely WNT, Shh, group 3 and group 4, exhibit distinctive transcriptional and epigenetic signatures that define clinically observed patient subsets.[59] The WNT and Shh subgroups are driven by mutations of the wingless and SSH signaling pathways, respectively. The emergence of the other two subgroups remains unclear. However, other genetic driver mutations and epigenetic changes have been repeatedly documented. Moreover, recent scRNA-seq studies have revealed that all subgroups mirror transcriptional programs observed in the developing brain, wherein group 3 MB resembled Nestin-positive stem cells, and group 4 MB resembled unipolar brush cells.[60]

Again, none of the above-discussed studies were performed in time series, a major drawback in cancer research. With time series analysis, a globally defined strange attractor(s) may exist in the state space (epigenetic landscape) of cancer gene expression dynamics. It is proposed herein that identifying such causal structures would pave reprogramming cancer (stem) cells toward benign-like, stable cell fates. The reprogramming of cancer (stem) cells to benignity provides a new avenue to investigate non-invasive, selective, and potentially universal cancer therapies. It also provides a form of empirically controllable, perturbation analysis to investigate cancers as dynamical systems and probe for causal patterns (attractors) in their gene expression state space, given time series data is available.[1] In support of these ideas, several studies have been reported indicating that certain genes act as epigenetic barriers preventing tumor reprogramming to benign stem cell fates.[61] For instance, it has been shown that CSCs can be reprogrammed *in vitro* to iPSC-like states or chemically directed toward specific cell lineage commitments.[62] In another study, human pancreatic cancer cells were reprogrammed to iPSCs *in vitro* using episomal vectors and demonstrated a lack of tumorigenicity post-reprogramming.[63] A more recent study further confirmed these findings by showing that inhibiting or expressing certain factor(s) can indeed reprogram cancer cells to premalignant epiblast-like states

(i.e., normal iPSC states). Kong et al.[64] demonstrated that suppression of the RAS-pathway can reprogram anaplastic thyroid tumors—aggressive, fast-growing lethal cancers—into transient iPSCs.

Sendai virus-mediated transduction of the OSKM factors enabled the generation of transient iPSCs only from tumor cells bearing a Ras mutation.[64] At the transcriptional and epigenetic level, the cancer-derived iPSCs clustered closely to human embryonic stem cells and lost many of their malignant markers. The genes upregulated in the cancer-derived pluripotent stem cells (SOX2, LIN28A, and SALL4) showed a decrease in DNA methylation surrounding their promoters, indicative of reprogrammed epigenetic landscapes and a loss of neoplastic potential.[64] While CD133+/CD44+ enriched CSCs may have been present in the starting tumor population, they were undetected post-reprogramming to epiblast-like states. Key molecular patterns regulated by the RAS-signaling cascade, such as NF-κB, were downregulated to levels seen in embryogenesis, as observed via the heatmaps constructed from gene set enrichment analysis. When a reversible pharmacological RAS-MEK inhibitor, PD0325901, was withdrawn, the reprogramming was blocked and the cancer-derived iPSCs converted back to their parental lineage-specific neoplastic cells.[64] The risks of these methods, such as the possibility of tumor regression/formation, teratoma formation, and other possibilities are ignored here, and only the ability to reprogram tumor plasticity is put into context.

These findings collectively demonstrate that the epigenetic landscape of cancer (stem) cells can be controlled and cancer cell fates can be selectively reprogrammed (fine-tuned). It is speculated that a different (set of) reprogramming factor(s) may be identified for converting each cancer subtype to healthy epiblast-like states (iPSCs). On the other hand, identifying a universal set of factors (a global attractor) capable of reprogramming any cancer phenotype to iPSC remains the holy grail of medicine. Such reprogramming experiments should be repeated on patient-derived cancer cells and isolated CSC subpopulations, followed by time series scRNA-seq analysis. These measurements should be further subjected to the algorithms for network state-space visualization and trajectory inference as discussed, followed by chaos detection and chaotic attractor reconstruction methods. These could serve as the initial conditions to identifying the minimal regulatory network driving cancer stemness. Herein, it is proposed that identifying these driver stemness networks is the equivalent of detecting strange attractors in the transcriptomic state space of cancer networks. If such attractors do exist and are shown to drive cancer stemness, it must be further investigated whether a universal strange attractor exists for all cancers.

Gene-editing methods (e.g., CRISPR technology) or pharmacological inhibitors/activators should be applied to perturbing the Suvà stemness clique and the essential gene network drivers of MB/GBM discussed above followed by a study of their gene expression in time series and detect the corresponding attractor dynamics. Reprogramming cancer (stem) cells to iPSCs and investigating their state-space attractor dynamics (energy landscapes) provides a powerful tool for searching the existence of global attractors governing cancer stemness/emergence. These techniques can also be applied to patient-specific tumors for identifying personalized cancer therapies and optimal treatment strategies.

In the later sections, various network inference pipelines and algorithms are discussed. Some pediatric cancer datasets from the above-mentioned works will be directly used to test-run some of these algorithms to illustrate their practical applications and critical importance in reconstructing patient-specific (personalized) cancer networks. A few selected algorithms are tested on the static (single-time frame) pediatric datasets to demonstrate how insightful complexity approaches are to clinical medicine and help illustrate the concepts discussed herein. However, time series RNA-seq datasets are severely lacking and therefore hindering true progress in clinical oncology.

GENE EXPRESSION DYNAMICS

In this section, I briefly review classical methods to model gene expression dynamics. As previously mentioned, treating GRNs as a network of oscillators coordinating developmental processes provide a deterministic approach to modeling gene expression. The attractor space of genetic oscillators provides a mathematical framework to visualize chaotic attractors and emergent behaviors. For instance, certain genetic oscillators exhibit synchronization during the collective cell migration (flocking) of cancer cells and their cytoskeletal fluids.[65] Oscillatory dynamics in gene expressions are necessary for stemness, a loss of which can be regarded as a loss of stem cell potential.[66] The cell fate transitions between stem cells and their differentiated counterparts have been discussed as regulated by a combination of noise fluctuations and oscillatory dynamics in gene expression levels.[66]

It is widely accepted that the cell fate transitions between the distinct functional phenotypes of a cancer population are governed by stochastic dynamics. Molecular observations can be explained by a Markov model in which cancer cells transition stochastically in between states and settle to equilibria (fixed-point attractors) of an epigenetic landscape. In stochastic models, the probability of state transitions are assigned to a population of cancer (stem) cells rather than describing the transitions using continuous, causal trajectories in state space.[67,68] Their ability to well-fit empirical cancer datasets must be appreciated.

In contrast, single-cell transcriptomics have revealed, in the past decade, that cell fate transition dynamics may be defined by smooth, continuous trajectories on the epigenetic landscape. Oscillator models and dynamical systems approaches are largely ignored in cancer cybernetics, as the wisdom of the crowd follows the central dogma of systems biology: gene expression dynamics is a stochastic process. In the stochastic modeling of continuous-time gene expression, three general approaches are widely considered: (1) the CME (chemical Master equation) is used to model the evolution of a probability distribution of chemical species across the system, (2) the Fokker-Planck equations (FPEs) provide a continuous approximation of molecular diffusion kinetics across a potential energy landscape, or (3) the Gillespie algorithm is used to simulate stochastic dynamics of individual trajectories. The latter two methods (2) and (3), are approximations of the CME approach (1).

CME is a relatively accurate framework for the stochastic modeling of gene expression dynamics. The CME provides a time-evolution equation for the probability of observing a gene

state and a certain number of gene products at a given time. It is a Markovian model in which the molecular numbers change in discrete (integer) amounts.[69] Continuous approximations of the CME include methods, such as the FPEs and partial integro-differential equations. These stochastic models of gene expression have proven to be quite insightful in the study of complex systems wherein FPEs are among the most widely used models to characterize cell fate transitions.[69] An alternate approach to solve the CME is the use of the Gillespie algorithm. However, the Gillespie algorithm is only computational tractable for small-scale systems with a few reactant species, and quickly becoming computational prohibitive as the number of species increases.

The simplest CME to consider is one that describes the gene expression without autoregulation or gene/protein bursting. For instance, consider the following set of ordinary differential equations:

$$\dot{m} = k_m - \gamma_m m$$

$$\dot{n} = k_p m - \gamma_p n$$

Where, m is the number of mRNA per cell and n is the number of proteins, γ is the degradation rate, and k_m *and* k_n are the kinetic rates (of synthesis) of the chemical species, respectively. A two-state model (often referred to as a constitutive expression model) describing promoter switching, transcription, and messenger RNA (mRNA) decay is the simplest stochastic model of mRNA fluctuations to consider. The model assumes that the gene is continuously turned on, producing mRNA at some constant rate, followed by mRNA decay or its dilution due to cell division.[70] If all these processes are approximated by effective first-order reactions, then the model is easy to solve and predicts a Poisson distribution of mRNA numbers in cells. Such a two-stage model of gene expression ignores that protein synthesis occurs in geometrically distributed bursts.[70] The CME for the two-stage model of gene expression assumes that the promoter is always active and has two stochastically fluctuating variables, the number of mRNA (m) and number of proteins (n). The CME is then given by:

$$\frac{\partial P_{m,n}}{\partial t} = k_0(P_{m-1,n} - P_{m,n}) + k_1(P_{m,n-1} - P_{m,n}) + d_0[\times (m+1)P_{m+1,n} - mP_{m,n}] + d_1[(n+1)P_{m,n+1} - nP_{m,n}]$$

Where k_0 is the probability per unit time of transcription, k_1 is the probability per unit time of translation, d_0 is the probability per unit time of degradation of an mRNA, and d_1 is the probability per unit time of degradation of a protein.[70] A Gillespie algorithm-based simulation of these deterministic ordinary differential equations (ODEs) will exhibit qualitatively different behaviors from the equations, such as the disregarded emergence of bursts. This demonstrates that the average of the many stochastic dynamics of molecules will eventually resemble the ODEs. We consider the behavior to be stochastic dynamics when $k_n \gg k_m$ or rather as the system increases in complexity, as many protein-protein interactions intertwine. The Gillespie algorithm is then a useful coping mechanism to overcome the complexity when studying the stochastic dynamics of large networks of chemical species' interactions.

Commonly, scientists *a priori* assume a Poisson distribution for the mean mRNA count per cell and mean protein synthesis per mRNA, then derive the Master equation for the distribution. The various parameters are further tweaked in the resultant differential equations, such as the synthesis and degradation rates of the products based on empirically quantified measurements.[71] A Poisson distribution is often sufficient to capture the technical noise in single-cell quantitative measurements, such as the RNA-seq counts.[71] The stochastic picture is repeatedly supported by findings in different cellular models, although not necessarily in pathological (cancer) networks. For instance, Stumpf et al.[72] demonstrated that stem cell differentiation may be best defined as a non-Markov stochastic process. Their findings collectively demonstrated a general increase in mean Shannon entropy as differentiation progressed from pluripotency to particular-cell fate commitments. When the data were fitted into a Hidden Markov process, the system was shown to obey a Poisson distribution. On such basis, Stumpf et al.[72] concluded that the increase in cell-cell variability observed indicated that cells do not synchronize their fate transitions and thereby stem cells progress in an uncoordinated, stochastic manner.

In contrast, there is a large body of experimental evidence showing that the distribution of molecule-number fluctuations in gene expression dynamics is non-Poisson.[73] If a gene can switch between an on and an off state, non-Poissonian mRNA fluctuations and bursting are observed. The model of mRNA bursting, commonly called the telegraph model, has thus been recently more widely adopted in the description of mRNA dynamics in eukaryotic cells.[73] In the Telegraph model, the CME is slightly altered to accommodate additional parameters, such as the marginal distribution of the nascent mRNA molecules (contains Gaussian-like terms). A major drawback in both coarse-grained models discussed here is that they do not take into consideration mRNA fluctuations during cell divisions/replication, an aberrant state of which is the signature hallmark of cancer.

Alternately, there are path-integral approaches involving the use of SDEs to study the nonequilibrium statistical mechanics of the Waddington landscape. The transitions between metastable, intermediate cell fates (i.e., mixed phenotypes or transient cell fates) are considered stochastic, and governed by the Diffusion matrix of the kinetic trajectories of involved chemical species in the transition. In principle, metastable attractor states may be nonequilibrium attractor states, such as strange attractors. A lack of dynamical systems approaches in the time series treatment of these systems remains a roadblock in deriving such conclusions, for the time being.

To simplify the epigenetic landscape's complexity, mean-field approximations are used to obtain steady-state distributions of gene expression.[74] A potential energy landscape is then reconstructed to visualize the attractors, such that given the steady-state probability distribution P, the potential energy is given by$U = -lnP$.[75,76] FPEs are used to describe the cell state transitions on the epigenetic energy landscapes. We shall not get into the finer details of these approaches herein and readers are directed to the citations. Note that for epigenetic landscapes exhibiting multistability, defining a potential U may be ambiguous. Furthermore, it is *a priori* assumed the solutions of the Waddington attractor landscape are steady state (i.e., solutions

negligibly change in time). Such nonequilibrium statistical approaches are only beginning to emerge in the study of cancer dynamics.

To illustrate, Rockne et al.[77] used a double-well state transition theory to map the state transition dynamics from healthy peripheral mononuclear blood cells in mice to acute myeloid leukemia (AML). They found that critical points on the transcriptomic state space could predict the cell fate trajectories of the disease. The two-dimensional state space was obtained from PCA dimensionality reduction on the time series bulk RNA-seq data.[77] The transcriptome was modeled as a particle undergoing Brownian motion (Langevin equation) in a double-well quasi-potential $U(x)$ with two stable states, representing the healthy and AML states. To calculate the mean expected behavior of the stochastic dynamics of a transcriptome-particle (cell), the evolution of the probability density function was obtained via the FPE.[77] However, the assumption that the healthy state and AML state critical points are stable phenotypes (i.e., valleys of a double-well potential) may have been the issue with this model. The cell states were *a priori* assumed to be stable, isolated fixed-point attractors of a conserved system. Simplifying the transcriptomic state-space trajectories as one-dimensional systems limits the phase space patterns to only fixed-point attractors and neglects oscillatory dynamics.[32] As discussed, cell phenotypes may correspond to attractors of various sorts, such as limit cycles (metastable fixed points) and chaotic attractors (unstable) on the multidimensional signaling state space. That is, in three-dimensions (R^3), trajectories can intertwine in state space and lead to chaos (i.e., strange attractors).

To avoid these assumptions and the rigorous use of differential equations and nonequilibrium statistical mechanics, system biologists have adopted a simpler approach to modeling gene expression dynamics: Kauffman's Boolean networks (BNs).[78] Most cancer network models in systems science have transitioned to the use of Boolean architectures since they are simple discrete-time models exhibiting complex behaviors.[79] BNs can model many types of biological networks, such as GRNs, metabolic networks, and protein-protein interactions (PPIs). The biological components (genes, proteins, or signaling molecules) are described by binary values of 0/1 and their interactions are represented by binary functions.[78,80] Let us consider GRNs as our biological network of interest. BNs contain a set of regulatory variables $x = \{x_1,\ldots,x_N\}$, $x_i \in B$, where each gene follows a Boolean logic, i.e., can be 0 representing the on state (expressed/activated) or 1 for the off state (repressed). The concentration levels of many regulatory processes in gene expression and their corresponding cell fate choices behave according to a Hill function.[81] For many values of Hill coefficients, the curve is sigmoidal and thus approximated as a dichotomous step function making BNs a representative model of gene expression fluctuations. BNs can be considered as a directed graph network, where each regulatory component is represented by a node of the graph. The directed edges between these nodes define the regulatory interactions and the dependencies are given by the Boolean functions. The state of a BN at a time t is given by the state vector $X(t)$, which at each discrete time t, a new network state is updated by applying the appropriate Boolean functions: There are three major updating schemes in how BN states transition:

(1) synchronous updates (all components take equivalent amount to change their value)
(2) asynchronous update (one random component updated at each time-step), wherein complex network topologies can emerge
(3) probabilistic update (alternative Boolean functions are used for each component), where the network can transition from synchronous to asynchronous states

GRNs and PPIs can exhibit scale-free topologies where the degree distribution of the regulatory interactions follows a power law distribution, i.e., most components are lowly connected while the rest form highly connected hubs.[82] Meanwhile, certain metabolic networks exhibit small-world graph topologies (i.e., the Watts-Strogatz network).[82,83] These network structures are highly robust and resilient in information flow. If we assume their network configuration to be fixed at some chosen time frame, their interactions can be defined as directed state graphs.[83] However, complex diseases, such as cancers, reveal dynamic non-trivial topologies with multi-nested feedback loops.

The long-term behavior of the network is a trajectory through the state graph. A periodic sequence of states forms an attractor on the state space, in which, once reached, the network configuration is stuck in unless some perturbation occurs. Therefore, attractors may correspond to cell phenotypes. There are various attractor types observed in BNs, namely, steady-state equilibria (fixed points), periodic cycles (can become chaotic with perturbations), and chaotic attractors (seen in asynchronous networks, but very rarely). The global behavior of BNs of N-genes (proteins) can be divided into three broad regimes: ordered (regular), chaotic, and critical.[80] Networks in the ordered (regular) regime, when placed into any random initial state in the state space, will quickly settle down to stable behaviors, such as a fixed-point attractor or limit cycle attractor with a small period T. In contrast, networks in the chaotic regime will exhibit apparent randomness in the phase space. Since the classical BN is discrete and finite, the network will eventually encounter a previously occupied state and repeat its trajectory (note: continuous BN models do exist). The period T of a limit cycle in the chaotic regime is very long and can be in the order of 2^N. Such an unstable limit cycle will exhibit aperiodicity and endless trajectory (permanent transient). The degree of chaos in the chaotic regime of BNs can be quantified based on the slope of the Derrida plot and spreading of information. Finally, the network in the critical regime is at the edge of chaos, where it can both generate novel structures and maintaining stability, simultaneously.[80] Edge of chaos is formally defined as a state of zero largest Lyapunov exponent.

When investigating BNs, one must consider the average connectivity per node k. It was found that when $k = 2$, the BNs are poised at the edge of chaos.[80] Above a critical k_c value (depends on various parameters), networks behave chaotically. Kauffman[80] proposed that most biological networks/cells are poised at criticality, allowing the dynamic transitions between regular and chaotic regimes. However, the implications of chaotic attractors in BNs remain poorly understood and without an interpretation in a biological context. No feedback mechanisms were strongly proposed to support criticality in gene networks or conceive their phase transition to chaos under symmetry

breaking. Self-organized criticality is only discussed qualitatively in analogy to Bak's sandpile model.[80] The EMT mechanism involved in CSC phenotype switching and cancer state transitions is often discussed as a process exhibiting critical dynamics (power law behaviors). However, their empirical investigation remains primitively developed. As stated above, STS is a mathematical framework to better understand the connection (transition) between chaos and criticality, and should be explored in cancer systems.

To envision the quantitative applications of BNs, consider a graphical representation constructed by the mapping of some experimentally derived gene expression matrix where each row represents a gene and each column a time point, under certain conditions. Transposing the matrix, mapping each column (condition) as a dot in a three-dimensional space where each axis is spanned by the number of genes expressed, creates a state space because each coordinate represents a state S of the transcriptome. With molecular profiling, a higher-dimensional state space consisting of epigenetic expressions can be reconstructed. A gene expression pattern reflects the dynamic configuration state of a GRN network and settles down to an attractor on phase space.[84,85] These states can be stable (e.g., a mature phenotype) or unstable (e.g., a transient state or stem cell phenotype)[86] (see Figure 2). Gene set functional enrichment analysis (GSEA) and clustering algorithms are then used to identify classes of genes or protein flows that are over-represented in a large set of genes or proteins associated to phenotypes, by navigating Gene Ontology databases.[87]

Huang et al.[88] used time series transcriptome data to experimentally verify the existence of robust attractors in the dynamics of hematopoietic progenitor cells. The trajectories of these stem cells toward neutrophil state differentiation, in multidimensional signaling state space, revealed a high-dimensional stable attractor denoting the distinct cellular phenotype. In contrast, cancer cells were discovered to be unstable attractors of their gene expression state space (epigenetic landscape).[84] Cancer cells were identified as states trapped in embryonic attractors of the epigenetic landscape. The EMT switch during metastatic progression was used as an argument for the emergence of embryonic attractor phenotypes.[89] Although they defined cancers as abnormal, unstable attractors, algorithms for the detection of chaotic attractors were not performed on time series CSC gene expression datasets in these analyses.

In principle, measuring the time series mRNA abundance and protein quantifications in the cancer stemness networks discussed above, exposed to various environmental conditions or drug/therapeutic perturbations, can provide an experimentally verifiable model for cancer signaling dynamics via these classical methods. The quantified data from kinetic and/or expression experiments can be fed into the CMEs and differential equations discussed in this section, projected in phase space, and searched for chaotic dynamics via computational algorithms/numerical methods. Alternatively, their gene expression matrix can be used to reconstruct BNs.

To contrast these two approaches, the CME offers a way of describing the evolution of a probability function. Thus, it is ideal for studying transition probabilities from one state to another. It is most applicable to describe discrete state transitions, although the equation is generally continuous. The FPE is a subset of the CME obtained by transformation/expansion, applied to a specific situation: Brownian motion (under drag forces and random forces). Thus, the CME is particularly relevant in the context of biochemical reaction networks where we have kinetic information about reaction rates (think of chemical reactions of reactants X and Y, where [X] + [Y] $\rightarrow$ [XY] at some rate k). From the CME, information about the evolution of the state space probabilities are derived, and we usually use truncation or moment closure to understand how the mean and the variances are changing over time. The classic BN is a discrete dynamical system in which there is only on or off (1 or 0) states and is relevant to expression rather than kinetics. There are limits to the type of attractors observed in classical BNs as defined above. In the "continuous" or semi-continuous forms of BNs, we get closer to the CME, but there is less integration of kinetics in BNs than in the CME. In BNs we are not looking at the evolution of probabilities, and rather observe the state space (unless some heuristics are applied to the whole network to infer probabilities).

While these classical approaches of gene expression modeling are powerful tools that have reshaped our understanding of network science, we are most interested in parameters-free, equations-free approaches to reconstruct cancer networks and their corresponding attractor dynamics (from time series datasets). Attractor dynamics and chaos detection are nearly absent in current practices of systems biology, which are adherent to stochastic modeling of gene expression. The current approaches naively hold many *a priori* assumptions by which the complex features of the dynamical system may be lost. Therefore, algorithms for parameter-free causality inference in GRN networks are essential to decipher cancer complexity.

PATTERN FORMATION

In this section, the essential physics of another category of cancer processes will be introduced: pattern formation. How do the protein products of the various gene expression patterns interact to form cancer structures? Protein-mediated pattern formation is another direction pattern scientists should take to detect chaotic attractor dynamics. Specific genetic programs and transcriptional networks are involved in synthesizing the proteins required for tumor pattern formation. The concentrations of these proteins are spatiotemporally distributed by the precise coordinated timing of multi-scale network processes, providing the blueprint for the self-organization of tumor structures. Turing coined the term morphogen to describe these protein flows orchestrating biological patterning.[90] Turing demonstrated the pattern formation (morphogenesis) of a wide variety of complex systems can be described using a set of nonlinear partial differential equations, i.e., reaction-diffusion equations.[90] The Turing model shows that even simple reaction-diffusion systems can generate complex structures, spatiotemporal oscillations, and chaos. Stationary patterns, also known as Turing patterns, are coherent stable structures acting as the underlying mechanism for cellular differentiation in healthy embryonic development.[91]

Many extensions of the reaction-diffusion models have been unfolded with various applications in systems biology. For instance, in the clock and wavefront model of somitogenesis, posteriorly moving molecular gradients sequentially slows the rate of clock (gene) oscillations, resulting in the patterning of

the developmental system.[92] In the French Flag Model of cellular differentiation, the effects of morphogen gradients in the positions of the various specialized cell types within a tissue are observed.[93] The concentrations below or above a critical threshold causes the activation of certain genes and hence, their chemical programs required for pattern formation.[93] A classic example, the stripes observed in the development of a fruit fly embryo, exhibits the precise positional localization of certain protein molecules alternating on the developmental plate of the embryo in high and low concentrations. Nonetheless, these reaction-diffusion systems are vastly disconnected, and it remains unclear how one model relates to another.

On the contrary to embryogenesis, nonequilibrium patterns and complex multi-scale (fractal) structures are universally observed in tumor pattern formation.[94] For instance, the migration/proliferation dichotomy (go-or-grow) hypothesis is an example of a reaction-diffusion-based model used in the description of glioma pattern formation.[95] The tumor density-dependent go-or-grow mechanisms can produce highly complex structures via Turing instabilities.[95] There is an inherent notion of disorder, aperiodicity, and instability in tumor patterning. Assuming the law of mass action and that the transport of these proteins is under Brownian motion (i.e., viscosity dominant), the variations of the involved chemical species' concentrations by these equations lead to spontaneous spatial patterns with a wavelength determined by the diffusion constants and reaction rates.[90] As discussed above, pediatric cancers, such as GBM and MB, are invasive diseases consisting of dynamical heterogeneous phenotypes interacting with an adaptive microenvironment.[2] The reaction-diffusion equations are accurate models to describe the dynamics of the various protein flows involved in cancer (stem) cells to give rise to their heterogeneous patterns. They provide a mathematical tool to visualize how the dynamics of cancer (stemness) networks give rise to cancer phenotypes, under a given set of environmental conditions. Therefore, a better understanding of the reaction-diffusion processes in the emergence and cell fate transitions of the distinct cancer phenotypes are fundamental to finding better therapy regimens.[96]

Consider a continuum description of the density of cells $u(r,t)$, and the density of some growth factor/protein or nutrient, $c(r,t)$ stimulating the growth diffusing to the tumor from far away, plated on an extracellular matrix. A tumor density-dependent dynamical switch of the glioma to invasive phenotypes is observed at critical concentrations of the growth factor, as governed by the reaction-diffusion equations:

$$\frac{\partial u}{\partial t} = \nabla\cdot(D_u\nabla_u) + \alpha u^2 c$$

$$\frac{\partial c}{\partial t} = \nabla\cdot(D_c\nabla_c) - \beta u^2 c.$$

Here, D_u and D_c are the diffusion coefficients of u and c, α is a proliferation exponent, and β is the coefficient of nutrient consumption. Reaction-diffusion models well-portray the growth-invasion dynamics of tumors and their dynamic pattern formation.[97] However, this simplified model transitions to the regime of nonequilibrium thermodynamics and nonlinear dynamics when we consider a many-body problem consisting of the interactions between tumors and their dynamic microenvironment.[22] For instance, chaotic patterns and strange attractor dynamics emerge when one considers the predator-prey dynamics of a three-body tumor ecosystem consisting of the competitive interactions among tumor cells, host cells, and immune cells, for resources.[47,98]

Fractals are the geometric signatures of chaos. Fractals are Nature's most stable structures exhibiting the optimization of space and spatially constrained resources. Think of trees, blood vessels, clouds, snowflakes, contours of geographic landscapes, etc. A system exhibiting fractal architectures may be robust to environmental changes. The fractal dimension is a statistical index of complexity. Tumor pattern formation exhibits complex, multifractal architectures.[18,99] The multi-fractality of cancer structures may imply their information and patterns are embedded across many length and time scales (i.e., self-similarity). Thus, a higher fractality in tumor structures may imply that the tumor is more complex, resilient (i.e., withstands environmental perturbations), aggressive, and difficult to treat.[18,99] The reaction-diffusion models underlying tumor growth-invasion dynamics have repeatedly demonstrated tumors can spontaneously self-organize into strange attractors and multifractal patterns in their state space. For example, strange attractors resembling the Lorenz attractor have been observed in reaction-diffusion models of cancers with nutrient-oxygen dependence.[100] These chaotic reaction-diffusion models were further validated on aggressive cancer models, such as malignant melanoma and mucinous ductal ectasia, which exhibited the emergence of strange attractors.[100,101] Experimentally validated mathematical models of tumor growth-invasion dynamics based on reaction-diffusion equations and predator-prey models (i.e., Lotka-Volterra-type equations) can exhibit chaotic attractors.[102] For example, the Itik and Banks model[44] consisted of three cell populations in a single tumor site component:

$T(t)$: *tumor cell number density at time t*

$H(t)$: *healthy cell number density at time t*

$E(t)$: *effector immune cell number density at time t*

Then, the system of ODEs describing the tumor growth dynamics were given by:

$$\frac{dT}{dt} = r_1 T\left(1 - \frac{T}{k_1}\right) - a_{12}TH - a_{13}TE$$

$$\frac{dH}{dt} = r_2 H\left(1 - \frac{H}{k_2}\right) - a_{21}TH$$

$$\frac{dE}{dt} = \frac{r_3 TE}{T + k_3} - a_{31}TE - d_3 E$$

Given the rate of change of the competing populations. Here, r_1 is the growth rate, k_1 is the maximum carrying capacity, a_{12} denotes the tumor cell population loss with competition with $H(t)$,

and a_{13} is the tumor cell population loss with competition with the effector cells $E(t)$. Here, $r_1 > r_2$, and the tumor cells also inactivate healthy cells at a rate of a_{21}. Experimental data can be fitted into these equations to model the predator-prey dynamics and obtain phase space portraits of their attractors. Itik and Banks[44] demonstrated that chaotic attractors with Shilnikov-like connections emerge when $a_{12} \geq 1$. The Lyapunov exponents were calculated using n-orthogonal tangent vectors satisfying initial conditions for the Jacobian matrix of the system, and the Gram-Schmidt orthogonalization followed by the Lyapunov K-dimension calculation. The K-dimension was computed to be about 2.03 in the model at selected parameters for chaotic onset, whereas the Lyapunov dimension of the Lorenz attractor is around 2.06. These chaotic reaction-diffusion models demonstrate the state-space dynamics of cancer cells exhibit strange attractors.

Itik and Banks[44] confirmed chaotic dynamics in cancer models by computing the Lyapunov exponents and the Lyapunov dimension, which was obtained to be fractal. The study strongly demonstrated that tumor-immune system interaction dynamics may result in the emergence of chaotic attractors in their phase space. As suggested by the authors, the existence of chaotic attractors may indicate tumor escape (immunosuppression) and the uncontrolled growth of cancer cells.[44] Thereby, a deeper understanding of the immune signaling pathways mediating tumor-immune interactions, such as cytokines, inflammatory pathways, etc., must be experimentally screened and probed as control parameters to study the state-space attractor dynamics of patient-specific tumors, using these Lotka-Volterra-type equations as a toy-model. As such, these equations and their phase space attractor dynamics are powerful tools to investigate the optimal dosage of immunotherapies and similar drug-based interventions in oncology. Simulations of these equations can be paired with experimental findings to confirm the detection of chaotic attractors. Time series molecular profiling may provide empirical data to be paired with the mathematical modeling.

In a similar study to Itik and Banks,[44] the competing populations of cancer cells, immune cells, and host (healthy) cells gave birth to chaotic attractors that were topologically equivalent to the Rössler attractor (i.e., a strange attractor).[45] Chaotic attractors were suggested to be indicative of fast-growing, aggressive tumors, therapy resistance, and metastatic potential. Therefore, controlling their unstable periodic orbits toward stable, non-chaotic orbits provides a screening method for identifying selective, targeted therapies.[45] Letellier et al.[45] further suggested that clinicians can use immunological parameters to assess tumor behavior during immunotherapy by studying their attractor dynamics using these equations-based modeling systems. Furthermore, the predator-prey dynamics and ecological relationships shared by the heterogeneous cancer cells and their microenvironment (i.e., immune cells, healthy cells, extracellular matrices, etc.) enables us to define the tumor's interaction with its microenvironment as a cancer ecosystem.

Many other mathematical models further confirmed the chaotic attractor model of tumor invasion-growth dynamics.[103,104] Numerical model extensions have demonstrated that periodic oscillations as well as chaotic attractors result in the complex dynamics of the tumor-immune cell predator-prey dynamics.[47,98] The emergence of these limit cycles and chaotic attractors were suggested as indicators of long-term tumor relapse.[47,98] For instance, a discrete map of a time-delayed cytotoxic T-lymphocyte response on tumor growth was mathematically modeled given the growth kinetics described by the logistic equation and the growth rate being the bifurcation parameter.[105] Increasing the control parameter beyond a critical threshold caused instability of the tumor state along with period-doubling bifurcations onto the emergence of strange attractors. These findings suggested metastatic progression and aggressiveness of tumor dynamics may correspond to not external parameters, but intrinsic tumor density-dependent chaotic dynamics. Nonlinear forecasting programs utilizing these equations were suggested to help monitor clinical tumor-immune dynamics under therapy progression.[105]

The emergence of chaotic patterns remains a poorly investigated branch of pattern science. While a significant importance has been put into how orderly patterns, such as stripes, swirls, and spots emerge in tissue patterns, the study of chaotic patterns in biosystems is in its infancy. Prigogine and Stengers,[106] pioneers of complexity science, demonstrated the spontaneous self-organization of highly ordered complex patterns known as dissipative structures in open reaction-diffusion systems. Their works established that reaction-diffusion systems far from equilibrium may exhibit spatiotemporal chaos. A recently emerging paradigm of pattern science is that spatiotemporal chaos can emerge in certain protein-mediated reaction-diffusion systems within cell populations.[107,108] The term chemical turbulence has been coined to describe these emergent chaotic patterns in cellular systems.[6] Chemical turbulence demonstrates that chaotic patterns, reminiscent of those seen at high Reynolds numbers in fluid dynamics, are observed in small-scale chemical reaction-diffusion systems. A broadband distribution in the power spectrum of the protein flows and a low spatial correlation length reminiscent of the Kolmogorov spectrum were observed.[6] Such complex systems are currently unaccounted for in tumor pattern formation. These findings on diffusion-mediated chemical turbulence should be extended to the study of cancer cells and CSCs and investigated for the emergence of chaotic attractors during their patterning processes (i.e., division and differentiation). Time-lapse videomicroscopy and similar imaging techniques can be used to capture potential chaotic dynamics of such proteins and reconstruct their attractor dynamics. The equations discussed in this section provide a theoretical framework to perform simulations and mathematical modeling, as done by Halatek and Frey,[6] and should be paired with the experimental data of protein dynamics during cancer (stem) cell differentiation and chemical perturbations/reprogramming.

CHEMICAL OSCILLATIONS AND SYNCHRONIZATION

The dance of various proteins, gene expressions, and cell signaling molecules at differing timescales (a)synchronize to this complex painting we call tumors. Tumor gene expression and pattern formation are complex processes fundamentally orchestrated by chemical oscillations. When these oscillations are aperiodic and irregular, chaotic dynamics are observed in cancer (stem) cells resulting in their adaptive emergent behaviors.[53] How can one understand aperiodicity, the signature of

chaos, without understanding oscillations? As such, the following are a few illustrations and theoretical models to build some insights for chemical oscillations in tumor cybernetics.

Winfree[109] and Kuramoto[110] pioneered the study of spontaneous synchronization of populations of biological oscillators. They demonstrated that each individual cell, organism, or biochemical process could be considered as an oscillator with its own distinctive frequency and phase. When considering a population of biological oscillators, a mean-field approximation can be considered where each oscillator responds to the average signal produced by all others.[109,111] In the case of a large population of weakly interacting oscillators, the mean-field distribution of natural frequencies may appear asynchronous and arrhythmic (i.e., random). As the coupling strength increases above a critical value of the order parameter, an abrupt phase transition often occurs where the population of biological oscillators enter a coherent (synchronized) state (i.e., phase locking).[111] These principles apply to the biochemical waves and oscillations observed within reaction-diffusion systems, whereby a lack of synchrony (i.e., aperiodic oscillations) may result in complex patterns and spatiotemporal chaos (i.e., chemical turbulence).[110,112] Kuramoto[110] established that reaction-diffusion systems share resemblance to the Navier-Stokes equations. The oscillatory units of the diffusion system can undergo synchronization when multiple periodic processes with different natural frequencies couple toward a collective frequency (phase). Mutual synchronization is a key mechanism for the self-organization of chemical waves into structured, coherent patterns. When synchronization is lost, chaotic patterns may emerge.

Cell division and morphogenesis is a highly coordinated process consisting of the precise timely oscillations and synchronizations of various cell-cycle regulators.[113,114] Cancers emerge from the dysregulated multicellular molecular networks of cell-cycle oscillations and aberrantly expressed cell-cycle proteins[115]. Interestingly, cell-cycle oscillations even in two-coupled mitotic oscillators with simple underlying kinetics can exhibit chaotic (strange) attractor dynamics.[116,117] Cell-cycle oscillations can be represented by simple first-order ODEs for simple gene circuits. A CME was derived to describe a mesoscopic model of tumor dynamics considering a cell-cycle master regulator, the p53-Mdm2 feedback loop system, in attempt to explain tumor apoptosis evasion.[118,119] Experimental studies of the tumor suppressor gene p53 and *Mdm2* oscillatory dynamics, in response to DNA damage, show damped oscillations of p53 concentration at the cell population level and undamped oscillation of p53 in single cells.[118]

Consider a simplified model where p53 activation is stimulated by some chemical species *Dm*, associated with the DNA damage level, and p53 stimulates the synthesis (positive feedback) of Mdm2, which in turn has a negative feedback (promotes the degradation) of p53.[118,119] Given A is the Mdm2 synthesis rate, B is the Mdm2 inhibition rate, and C is the rate constant for damage reparation by p53 dynamics, K is the p53 degradation constant, δ_1 is the p53 basal synthesis rate constant, and δ_2 is the rate constant associated to the DNA damage level, the transition probabilities of each species' concentration per unit time is given by the following ODEs, in the absence of external fluctuations:

$$\frac{d[p53]}{dt} = \delta_1 - K\left[p53\right]\left[Mdm2\right],$$

$$\frac{d[Mdm2]}{dt} = A[p53] - B[Mdm2][Dm],$$

$$\frac{d[Dm]}{dt} = \delta_2 - C[p53].$$

The model selected δ_1 and B as the control parameters dictating the dynamical system's bifurcations.

In this case, when $B<B_c$, p53 shows damped oscillations and when $B>B_c$, the cell apoptosis occurs as regulated by the p53 sustained oscillations (undamped oscillation, limit cycle).[119] The use of such simple kinetic equations provides basic insights into the oscillatory dynamics and their corresponding attractor spaces in cancer cell-cycle networks. These equations can be used as a model system to screen for the presence of chaotic dynamics in cancer cell-cycle oscillations. Although chaotic dynamics were not investigated in this study, the entropy production rates were shown as robust characteristics of the tumor complexity, measurable from the contour-based fractal dimensions of avascular tumor growths.[120] Increased entropy production rates were proportional to the growth rates (invasiveness) of tumors.[120] As will be discussed later, a positive (increasing) entropy production rate is a useful measure to infer chaotic systems (in combination with other methods).

The mutation/inactivation of p53 is a critical event in the development of cancer cells and CSCs. Indeed, p53 inhibition and hTERT activation followed by the aberrant acquisition of stem cell-associated morphogens, such as those conferred by EMT programs, Nodal, Shh, Notch, Pax, Par, TGF-β, and Wnt proteins, to name a few, determine the phenotypic plasticity in CSCs.[121] The oscillations of these morphogens provide the developmental cues for tumor structures and their regenerative potentia.l[121] Many of these morphogens regulating cancer stemness are also involved in the reaction-diffusion processes governing embryonic development as periodic oscillators. The time series measurements of these factors' oscillations in tumors fitted into the above-discussed equations, with appropriate state-space reconstruction algorithms, may provide a simple framework for detecting chaotic attractors in the oscillatory dynamics of such proteins in cancer (stem) cells.

Metabolic networks may also exhibit chaotic dynamics. Tumor ecosystems exhibit drastic phenotypic transitions under nutrients and growth factor oscillations. For example, under glucose starvation, glycolytic oscillations were observed in individual HeLa cervical cancer cells.[122] The intracellular oscillations and intercellular synchronization of such oscillatory dynamics were collectively observed under nutrient deprivation.[122] With FBS serum starvation, collective asynchronous and aperiodic oscillations with differing amplitudes and periods were observed forming heterogeneous behavioral patterns.[122] Aperiodicity, irregular, and asynchronous oscillations are hallmarks of chaotic dynamics. Further investigation is required in the plausibility/emergence of chaos in metabolic networks.

The key point to be recognized here is that periodic oscillators (limit cycles) can induce and transition toward chaotic behaviors

when stability is lost. Oscillators are potentially a rich source of chaotic dynamics. In fact, chaotic oscillations in dissipative systems were first observed when periodic self-oscillators were periodically perturbed.[123] The theoretical framework of chaotization of oscillations by periodic forcing was provided by Russian mathematicians like Afraimovich and Shilnikov, who indicated three possible bifurcation mechanisms by which this chaos can be generated, and these have been documented in turbulent fluid models, such as the Taylor-Couette flows.[123] For example, yeast extracts with periodic perturbation of varying glucose fluxes will begin to chaotically oscillate.[123]

While the investigation of chaotic oscillations at a single gene or single protein level remain under-explored in cancer research, a recent study in healthy systems may be of central interest to cancer systems. NF-κB is a central TF involved in complex processes, such as inflammation, tumor-immune homeostasis, and tumor microenvironment dynamics. The activation of the NF-κB pathway and its associated signaling cascades exert significant immunomodulatory effects involved in inflammation, cancer immortalization/telomerase overexpression, cellular proliferation, and cell survival, critical processes for cancer progression.[124] Oscillations in the nuclear cytoplasmic translocation of NF-κB regulate these complex dynamics. Jensen and Krishna,[125] using a mathematical model, demonstrated that when the amplitude of an externally applied stimulus, tumor necrosis factor (TNF), exceeded a critical threshold, the coexistence of multiple synchronized states and eventually chaotic dynamics of the nuclear NF-κB concentration (oscillation) are observed. Furthermore, it was suggested that this could be used as a way of externally controlling immune response, DNA repair, and apoptotic pathways.[125] In confirmation of these models, chaotic oscillations in gene expression and protein oscillation dynamics were confirmed in a recent set of studies by Heltberg et al.[5,126] The team demonstrated that NF-kB-dependent genes can be frequency controlled by controlling an externally applied TNF amplitude.[126]

Periodic inputs of a signal can cause the dependent genetic oscillations to undergo frequency-locking and phase coherence to the external signal. Thus, periodic stimulus with TNF was shown to synergistically enhance the NF-κB oscillations, which above a critical threshold at large enough amplitudes of TNF, overcame the noise/stochasticity of the gene expression landscape and resulted in the birth of chaotic attractors.[5,126] The NF-κB protein was shown to be most effective at activating downstream genes via signaling cascades and optimally tuning the immune cells when in a chaotic state. Strange attractors were shown to form by the cellular protein densities. Chaotic dynamics in protein/gene oscillations were also suggested to enhance the formation of protein complexes required for signal amplification.[5] These findings suggest chaotic dynamics may be used by cancer cells, master hijackers of the immune system, to regulate their complex signaling dynamics and cell fate decision making. Heltberg et al.[5,125,126] also demonstrated that chaotic dynamics of NF-κB in individual cell states may increase the dynamical heterogeneity of the populations and allow them to adapt to harsh, fluctuant environmental conditions. When cytotoxic drug perturbations were considered in the model, due to large fluctuations, chaotic cell states survived. Therefore, chaotic dynamics may provide cancer cell populations adaptive heterogeneity and therapy resistance.[5,125] The chaotic dynamics of these signaling pathways, if experimentally confirmed in cancer cells, would serve as critical targets for cancer-specific therapies and reprogramming cancer attractor (epigenetic) landscapes.

Using microfluidic cell cultures, Heltberg et al.[126] delivered periodic TNF simulation to fibroblasts and recorded the NF-κB nuclear localization by live-cell fluorescence imaging. CellProfiler and MATLAB peak analysis algorithms were used to track cells and quantify the NF-κB translocation, where the activation was quantified as mean nuclear fluorescence intensity normalized by mean cytoplasm intensity.[126] The transitions of phase locking in an oscillatory manner were observed even among nonequilibrium fluctuations. Simulations using differential equations further revealed that these oscillatory frequency jumps are chaotic attractors, whereas the driving TNF force oscillation increased, deterministic chaos was observed in state space.[5] The fourth-order Runge-Kutta and Gillespie algorithms were used in the deterministic and stochastic stimulations of the TF dynamics, respectively. Before the chaotic transition, period doubling, a signature of chaos, was observed in the NF-κB oscillation amplitudes in time. Even at the early onset of chaos, transient and unstable limit cycle behaviors were formed—confirming the birth of chaotic attractor as discussed in earlier sections. These findings and methods should be extended to tumor ecosystems, and their TF networks, such as the stemness networks driving pediatric GBM and MB discussed previously. Due to the critical relevance of these findings in chaos detection, below are the mathematical equations needed to simulate the chaotic attractors observed in this study.

Assume the NF-κB can bind to an enhancer or operator region and form complexes to bind the RNA polymerase with different affinity. The protein synthesis and mRNA level transcribed by the i-th genes (i.e., transcription and translation) is described by the following differential equations:

$$\begin{cases} \dot{m}_i = \gamma_i \dfrac{N^{h_i}}{k_i^{h_i} + N^{h_i}} - \delta_i m_i \\ \dot{p}_i = \beta_i m_i - \Delta_i p_i \end{cases},$$

where m_i is the concentration of mRNA of i-th-species, and p_i is the protein levels of i-th species, the affinity of the binding is denoted by the parameter k_i (determines the concentration of NF-κB that results in 50% maximal gene expression enhancement), and the corresponding Hill coefficient is h_i (a measure of cooperativity of TF at the gene). The first term in the equation is the Hill function of mRNA, used normally as a description of proteins produced by genes as governed by TFs. The sigmoidal curves produced by the Hill functions enable them to be ideal candidates for BN modeling as well. Here, γ_i describes the mRNA production per time, δ_i the decay of mRNA per time, β_i the protein production per time, and Δ_i the decay of protein per time. All these parameters can be experimentally quantified with time series cancer datasets.

It must be emphasized again that the study of chaos in cancer processes should, when possible, be paired with the analyses of computational models/simulations of such processes. In general, the gold standard for empirical detection of chaos is to

demonstrate that a particular parameter change produces chaos in a model as well as in the real system that the model is simulating. An example of this was the empirical works by Heltberg et al.[5] discussed above. The emerging paradigm of complex systems is essentially the marriage between computational algorithms/simulations and empirical analyses. On a tangential note, cancer dynamics can also be visualized from an evolutionary game theoretic perspective when studying the adaptive behaviors of tumor ecosystems to cancer therapies. Tumor heterogeneity and therapy resistance can be explained as adaptive game strategies in response to their environmental conditions.[127] For instance, the total chemotherapy delivery, usage, and time to recovery can be optimally controlled by solving a system-specific Hamilton-Jacobi-Bellman equation.[127] Such principles of game theory dynamics will not be discussed herein.

NETWORK SCIENCE AND PATTERN INFERENCE

From here onward, the central focus of the paper will be the inference and state-space reconstruction of cancer networks. To model dynamic networks from time series data, first we must know how to reverse engineer the underlying networks from real data. Let us suppose we do not necessarily adhere to a Boolean framework or the CMEs discussed so far to infer the dynamics of complex networks. How can we reconstruct parameter-free biological networks using experimentally derived datasets without such assumptions? This brief section is allocated to provide the mathematical framework of graph networks and a general recipe to reconstruct parameter-free GRNs and PPIs from experimental cancer datasets. Only then can algorithms to infer chaotic attractor dynamics on the networks' signaling state space be discussed in a meaningful manner. A multitude of algorithmic pipelines exist for reconstructing cancer signaling networks from molecular/gene expression datasets. These algorithms have been reviewed in Uthamacumaran,[1] and therefore will not be rediscussed in detail herein.

Network science is the interdisciplinary branch of complexity science in which the statistical mechanics and dynamics of complex, nonlinear networks are investigated. Networks are universal, versatile tools at the heart of pattern science. Networks are used to study complex diseases, such as cancer, and assist finding targeted therapies (personalized pharmacogenomics) in precision oncology.[128] Network medicine advocates the use of networks in clinical oncology and thus provides a bridge between pattern science and cancer research. The initial conditions of network theory stems from a discipline of mathematics, graph theory. Any biological network, regardless of its complexity, can be represented as a graph network. Graphs are abstract representations of the relationships between the various members of a system (network). A typical graph structure G (V, E) consists of the associations between various nodes (vertices) V indicated by the edges (lines) E connecting them. Graphs can be directed or undirected, while the nodes can have labels and features which determine which type of learning algorithms are most suitable to optimize their function or classification/prediction.[129]

For instance, Z scores assess the gene-gene correlations in cell clusters identified from the gene expression matrix of scRNA-seq datasets.[130] Z scores are preferred over Pearson or Spearman correlation metrics for single-cell datasets, while the latter may be better suited for bulk RNA-seq datasets. The cell clustering must be performed in a pre-processing step with log-normalized gene expression counts. The identified gene-gene pairwise correlations serve as the weighted edges of the inferred regulatory network, where the nodes represent the genes.[130] The same interpretation holds for other types of molecular networks (e.g., PPIs, metabolic networks). The bigSCale algorithm is available as a machine learning pipeline to automatize these processes and thereby serve as computational tools to infer cancer networks.[130]

A computer recognizes a graph as various matrices. For instance, a feature graph or an incidence matrix, of size $n \times m$, represents the statistical information about the relationships between two classes (i.e., vertex-edge), where n is the number of nodes and m is the number of edges. The information on the incidence matrix can be extracted to more useful matrices, such as the adjacency matrix, A_{ij}, a matrix wherein the components are binarized to either 0 or 1 depending on whether the nodes of the graph are adjacent (connected) or not. The connectivity (associations between the nodes) can also be weighted and updated.[83,129] The adjacency matrix of gene expression determines the network structure.

If the network is weighted, A_{ij}, can take positive values different from one representing the weight of the link (relationship). In general, directed, and undirected graph networks are represented by symmetric and unsymmetric matrices, respectively.[129] There are many other matrices used to assign the relationships in a complex network. For instance, a degree matrix is a diagonal matrix which contains information about the degree (the number of nodes connected to each edge of a graph network). Various network parameters, such as the degree, directionality, smoothness (i.e., Graph Laplacian), architecture, and connectivity (e.g., power law/scale-free, small-world, Erdos-Renyi, etc.), are used to characterize the graph network's topology.[83,131] Some measures to assess a network structure's statistical mechanics include the following: (1) clustering coefficients (to define the local cohesion of network structures, such as cliques, motifs, communities, etc.), (2) centrality (isolate hubs of information), and (3) degree distribution and degree-correlations.[83]

Functional assignment of genes is often performed using clustering or undirected graphical models, including Markov random fields, while pathway reconstruction is often based on directed probabilistic graphical models. Common network clustering algorithms (e.g., Bayesian clustering, K-means, spectral clustering, agglomerative hierarchical clustering) and pairwise interactions' information theoretic measures (e.g., entropy scores, mutual information) are then used to infer statistically correlated network structures.[82,132] For instance, cancer networks can be graph partitioned into smaller sub-networks via hierarchical clustering and dimensionality reduction techniques.[133] Several measures can be used to assess the similarity of two time series gene expression datasets, both model-free, such as Euclidean distance, correlation and lag-correlation, and model-based methods, such as the Kullback-Leibler distance. Spectral clustering algorithms remain among the most robust network inference/graph partitioning tools used to find community clusters and protein expression- and gene expression-based cancer subtypes within complex networks.[134,135]

The typical workflow for dissecting intra-tumoral heterogeneity from single-cell datasets is as follows: (1) filtering the gene expression matrix (i.e., discard low quality cells and zero counts), (2) data processing (i.e., normalization of gene expression, imputation, and feature-selection/identify differentially expressed genes via correlation metrics), (3) correlation analysis—additional correlation metrics, such as *Z* scores, or similarity metrics, such as the Jaccard index, are used to identify clusters of patterns, (4) visualization via dimensionality reduction (find a low-dimensional representation of the dataset while preserving essential structures), (5) clustering (algorithms to label cells with discrete states/types) and modularity (graph clustering/community structure detection), (6) identify differentially expressed patterns, and (7) lineage trajectory tracing (various pseudotime inference algorithms exist for differentiation and lineage trajectory mapping).[1,136]

The lineage trajectory tracing (pseudotime inference) techniques, traditionally applied to stem cell datasets, are extensively used in model systems that exhibit cell fate transitions, such as cancers. When time series datasets are absent (the status in cancer research), pseudotime trajectory analysis is used to infer the trajectories of cell fates from their static gene expression datasets. The pseudotime is a quantitative measure of single-cell progression through a biological process, such as differentiation/cell fate transition. Following trajectory analysis, the identified clusters and gene-gene or protein-protein correlations are used to reconstruct the inferred regulatory networks.[136] GSEA can further reveal clinically relevant transcriptomic markers and functional network modules. Network properties are then assessed to study the network structure and infer information flow (dynamics).

Some important, general measures/properties of complex networks include the following:

(1) Degree (connectivity): k, indicates how many links the node has to other nodes in a network. The connectivity can be undirected or directed.

(2) Degree distribution: $P(k)$, gives the probability that a selected node has exactly k-links. Unlike a random network, a power law degree distribution indicates that a few hubs hold numerous small nodes. The degree distribution is obtained by counting the number of nodes $N(k)$ with k-links and dividing by the total number of nodes N. Most biological networks are scale-free, and the degree distribution obeys a power law $P(k) \sim k^{-\gamma}$, where γ is the degree exponent.[82] The smaller the γ, the more important the role of the hubs in the network. When γ>3, the hubs become quite irrelevant and the network dynamics transitions toward randomness. When 2> γ> 3, there is a hierarchy of hubs, with the most connected hubs in contact with the large fraction of all nodes. If $\gamma = 2$, the largest hubs are in contact with the large fraction of all nodes. In general, universal properties of scale-free networks exist if γ<3, where the dispersion of the distribution $\sigma = \langle k^2 \rangle - \langle k^2 \rangle$ increases with the number of nodes, and higher the degree of robustness against accidental node failures or random perturbations.[82,83]

(3) Shortest path/mean path length: finding the shortest path of a complex dynamical network is NP-complete. Various heuristics exist to infer the mean path length.

(4) Clustering coefficient: $C_i = \frac{2n_i}{k(k-1)}$ denotes the clustering coefficient, where n_i is the number of links connecting the k_i neighbors of node i to each other.[82] The clustering coefficient gives the number of triangles that go through the node i. The average clustering coefficient $C(k)$ of all nodes with k-links, for many real networks is approximately k^{-1} indicating a hierarchical character.

(5) Modularity: most real-world, complex networks are heterogeneous, nonlinear, non-random, and exhibit modularity (i.e., cluster into distinct groups/modules with varyingly dense connectivity). Several algorithms have been proposed to compute partitions of networks into communities that score high on a graph clustering index called modularity. However, modularity optimization is an NP-hard problem, where our current approaches, including the most efficient graph clustering algorithms, are only heuristics.[137,138] A general benchmark to assess the modularity and the richness of communities within a complex network is given by the Newman-Girvan modularity measure, which states:

$$Q = \frac{1}{2m}\sum_{ij}[A_{ij} - P_{ij}]\delta(c_i, c_j),$$

where Q is the modularity index (expected fraction of edges forming a community), indices i and j correspond to the nodes, c represents their respective communities, and A_{ij} is the adjacency matrix of the graph partition of interest.[139] Consider the matrix formulation of modularity as follows. The actual number of edges between i and j is given by the adjacency matrix: $A_{ij} = \begin{cases} 1 \text{ if there is an edge } (i,j) \\ 0 \text{ otherwise} \end{cases}$ and P_{ij} is the expected number of edges between nodes i and j, then the modularity Q is the sum of $A_{ij} - P_{ij}$ over all pairs of vertices (i,j) falling into the same group. The simplest approach to partition the nodes is to consider a bipartition. This method is analogous to finding the ground state energy of a ferromagnetic random field, known as the Ising model.[140]

We define, $S_i = \begin{cases} +1 \text{ if vertex } i \text{ belongs to group } A \\ -1 \text{ if vertex } i \text{ belongs to group } B \end{cases}$, then $Q = \frac{1}{4m}[A_{ij} - P_{ij}]S_iS_j$. The optimization of this Ising spin model is NP-hard. Thus, to maximize Q, we can convert the discrete spin model into a continuous eigenvalue problem, such that we seek to maximize: $[A_{ij} - P_{ij}]S = \lambda S$, where λ is the Lagrange multiplier and the discrete constraint for the spin S is relaxed toward. $|S|^2 = n \epsilon Z$.

The Louvain community detection algorithm is often used as an efficient community structure searching heuristic, which functions based on two iterative steps: (1) find a local maximum and (2) compare neighbors to the local maximum by building a new network whose nodes are the communities, and iterate.[141] Due to the roughness of a network's modularity landscape, finding community structures is not the sole pattern used to assess complex network structure-function relationships. The dynamic curves of a network's associations are of central interest, especially when studying dynamical diseases.

The following are few examples of network inference and visualization tools available to pattern scientists investigating cancer

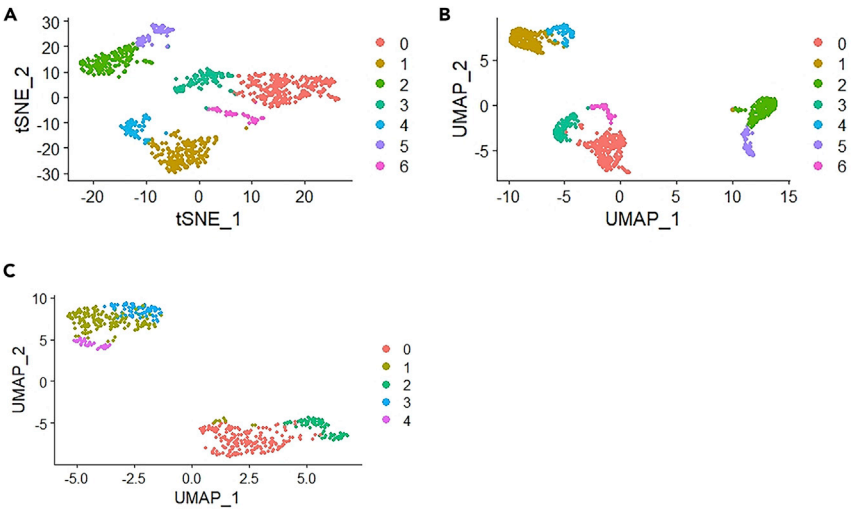

Figure 3. Results with Seurat pipeline: A framework for clustering, phenotyping, pseudotiming, and inferring GRNs from single-cell data

(A) TSNE plot in Seurat. The t-SNE pattern space of a few pediatric GBM samples generated by Seurat (n= 6 samples, with N = 846 cells analyzed; count matrix from Neftel et al.).[2]

(B) UMAP plot in Seurat. The UMAP Rplot of a few pediatric GBM samples; same data as in (A).

(C) UMAP plot of a GSC sample. UMAP plot of a GSC single sample shows two distinct cell states (sample BT127_L, *N* = 500 random cells in count matrix from Richards et al.).[3]

networks. Some results have been computed to help illustrate the essential concepts. Seurat is a machine learning pipeline used as a gold standard pre-processing and clustering tool for dissecting cancer gene expression datasets.[142] The identified clusters can be further sorted by other algorithmic packages into regulatory flow networks. Cells cluster together based on similarly expressed genes. Graph-based clustering algorithms are used by Seurat to assign potential phenotypic variations from differentially expressed genes. An example of a graph-based clustering algorithm is a *K*-nearest neighbor (KNN) graph, where the edges are drawn between cells with similar feature expression patterns, and the graph is partitioned into highly connected cliques or communities. Following on from this, the Louvain community detection algorithm is used for modularity optimization. Then, nonlinear dimensionality reduction methods, such as uniform manifold approximation and projection (UMAP) or t-distributed stochastic neighbor embedding (t-SNE), are used to visualize the sorted clusters. Positive and negative expression markers distinguishing the distinct phenotypic clusters can also be extracted to reconstruct the regulatory networks. In principle, once these cluster-specific markers are identified in cancer populations, chaos detection algorithms, such as Lyapunov exponents estimations, and entropy rates can be used to infer chaotic patterns in the individual signal's time series. Single-cell datasets can be highly noisy, and therefore more complex machine algorithms for chaos detection will be required, as discussed in the next sections. Regardless, Seurat-like tools are powerful pre-processing algorithms to visualize the cancer networks. To demonstrate the applicability of Seurat-like clustering and network visualization algorithms, allow me to illustrate Seurat's performance on the pediatric cancer datasets from Neftel et al.[2] and on the GSC datasets from Richards et al.[3]

Figure 3A shows the Seurat-generated Rplot of a few pediatric samples' GBM cells clustered into four global clusters in a t-SNE embedding. The tumor heterogeneity is well-visualized by the various colors indicating the six cell clusters with similarly expressed gene patterns. Figure 3B shows the same results for a few pediatric samples' GBM cells under a UMAP embedding space, where three distinct clusters are observed. Finally, Figure 3C shows two distinct cell clusters for adult GSC cells, single sample (BT127_L) from Richards et al.[3] dataset in UMAP embedding. The distinct clusters from the dimensionality-reduced static datasets show that different classes of drugs (targeted therapies) may be more effective in specific subtypes of a patient's cancer. The key insight to consider is that with time series, dynamical structures connecting the distinct clusters may emerge. If such attractor patterns are observed showing the transitions and potential interconvertibility of cell fates from one cluster to another, there are specific genes and molecular regulatory networks underlying these dynamical attractors. In principle, targeting such attractors, if detected, provides the most robust approach to cancer therapy. It may further pave road to reprogramming cancer (stem) cells toward benign states.

Figure 4 shows the visualization of selected markers in the cells projected on the UMAP dimensionality reduction space from Figure 1. As discussed, the selected markers compared in Figure 4 are three of the four essential TFs identified by Suvà et al.[54] for glioblastoma stemness. Figure 4A corresponds to the expression of these stemness factors in a few pediatric GBM samples, corresponding to Figure 3B. Figure 4B compares the stemness factors in a single pediatric GBM patient sample (BT1160) to demonstrate the applicability of these tools in assessing personalized, patient-specific targeted therapies. Finally, Figure 4C shows the markers' expression in adult GSC cells, corresponding to Figure 3C. The key remark established here is that these stemness factors are expressed in all pediatric GBM clusters, further blurring the distinction between GSCs and the "differentiated" cancer cells (GBMs). As mentioned previously, mature/differentiated GBM phenotypes could be reprogrammed into GSCs by the induction of these stemness TFs.[54] The dynamics of these stemness networks must be further investigated with time series analysis.

Figure 5 displays violin plots of the single cell data from Neftel et al.[2] for a few pediatric patient GBM samples (corresponding to Figure 3A). The sorted identity classes (on the *x* axis) indicate the distinct cell clusters identified by the average expression of the marker being potted. The compared gene markers are the four major molecular subtypes identified by Neftel et al.[2] discussed in "Cancer networks". From a static point of view they may be distinct cell clusters. The Seurat-guided clustering tool was used to generate these data. However, with time series analysis, dynamical attractors may join the clusters, as will be evident in

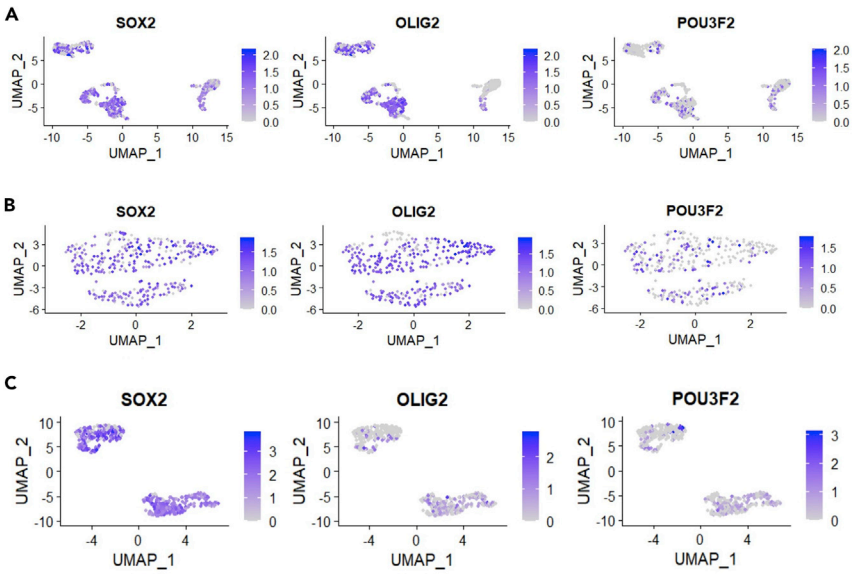

Figure 4. Feature plots of stemness marker expressions in GBM and GSC samples

(A) Selected stemness markers' expression in a few patients. Few pediatric GBM samples' UMAP pattern space of selected markers (GBM stemness factors). A higher expression is denoted by the blue color gradient.

(B) Selected stemness markers' expression in a single patient. Single pediatric GBM sample's UMAP pattern space of selected markers.

(C) Selected stemness markers in a single GSC sample. Single sample GSC feature plot of selected markers.

the single-cell energy path (scEpath) algorithm discussed in the following section.

Slingshot is a pre-processing, cell lineage classification, and trajectory inference algorithm in which diffusion maps, PCA, etc., are used as dimensionality reduction techniques followed by model-based clustering of cell populations.[143] The underlying GRNs and their cliques (maximally connected subnetworks) can be assessed from the distinct clusters mapped on the energy landscape using the above-defined network measures and correlation metrics.[143] To demonstrate, Figure 6A and Figure 6B show the pseudotime inference of cell states from the single pediatric GBM dataset (BT1160) under a PCA and diffusion map pattern spaces, respectively. As seen, the black curve indicating the inferred trajectory transitions from the distinct cell clusters depend on the type of dimensionality reduction technique used, indicating that neither by themselves provide a realistic description of the actual cell state trajectories. Three distinct clusters of similarly expressed gene expression patterns were identified in the sample from the generated expression heatmaps in Slingshot and in Seurat (data not shown).

As a final remark to this section, there are a few emerging empirical techniques to visualize trajectory inference/cell fate transitions in single-cell studies that pattern scientists should be aware of. The cell fate trajectories can be quasi-mapped as attractors in state space via these emerging methods and should be coupled to the algorithmic approaches discussed here. Nearly all currently used techniques for gene expression profiling in single cells do not directly capture transcriptional dynamics. Transcription dynamics can be quantitatively visualized directly *in vivo* using RNA imaging techniques or it can be inferred from (short-lived) protein reporters. For many genes, several RNAs are synthesized almost simultaneously, followed by a period of transcriptional quiescence, known as transcriptional bursting or intermittency.[144] La Manno et al.[145] demonstrated a method that enables the level and rate of change of gene expression dynamics to be estimated simultaneously for each gene in a single-cell dissection. Newly transcribed mRNA contains segments that are spliced during mRNA maturation. For a gene that is stably expressed, a small fraction of its mRNA is found in the unspliced form.

When a gene has been activated, for a brief time there will be a higher proportion of immature transcripts, and the converse holds true as well. Therefore, the ratio of unspliced to spliced mRNA can be used to infer instantaneous expression dynamics—referred to as the "RNA velocity" of each gene during cancer (stem) cell differentiation and cell fate decision making. The velocity field can be projected on a t-SNE plot's pattern space where the cell clusters labeled with specific differentiation markers can be visualized. More recently, Sci-fate was demonstrated as a new method using combinatorial cell indexing and 4-thiouridine labeling of newly synthesized mRNA to concurrently profile the single-cell transcriptome.[146] Lung adenocarcinoma cells were treated by a drug for perturbation analysis in gene expression dynamics, for various time intervals up to 10 h. RNA velocity analysis revealed to be a powerful tool to capture the main cell fate trajectory types in the lung cancer profiles.[146]

In principle, the previously discussed general chaos detection algorithms, such as Lyapunov exponents estimation and fractal dimension analysis, can be applied on the time series of the identified markers in these RNA velocity methods.[145] Despite that, there are two caveats: (1) as the cell count increases, the pattern space can become difficult to follow and, more fundamentally, (2) the projection of scRNA-seq data into low-dimensional latent spaces may remove many of the essential features of the complex dynamics. For instance, to obtain the UMAP or t-SNE plot, certain hyper-parameters must be tuned, and changing these parameters even slightly can drastically alter the clustering and results/interpretation (i.e., hyperparameter tuning is NP-hard). Regardless, suppose chaos is indeed detected in these patterns, either at the level of cell fate trajectories or at the level of their individual gene/protein networks. How does one determine its underlying causal structure, the strange attractor (if it exists)?

ATTRACTOR RECONSTRUCTION METHODS

By now we have cultivated an intuition for the patterns and causal structures (attractors) complex systems can produce. The analytical detection of limit cycles, tori, and strange attractors within the state-space reconstruction of dynamical systems remains an intractable, NP-hard problem. Furthermore,

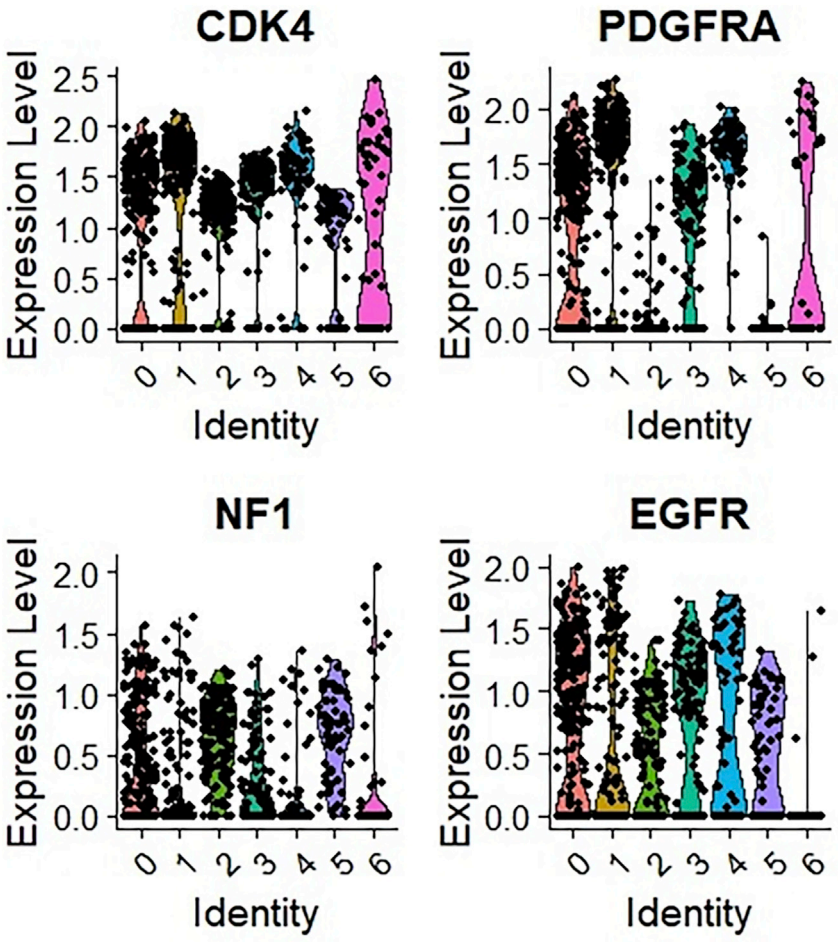

Figure 5. GBM cell clustering by Seurat
Rplot of a few pediatric GBM samples' expression levels of selected markers in the distinct identified cell clusters are shown. The markers correspond to the four GBM molecular subtypes, as discussed.

correlation does not imply causation. This remains a fundamental issue in our current statistical approaches to studying cancer signaling dynamics. The statistical patterns detected by contemporary methods provide mostly a static picture of gene expression, but how does one infer dynamics? If there are chaotic signatures in an experimentally acquired time series dataset, how will they be distinguished from noise? To reconcile these problems, the following sections are allocated to pertinent algorithms in causal discovery that should be incorporated into the network inference algorithms and Waddington landscape reconstruction algorithms discussed herein. The following algorithms can be merged with many of the freely available network reconstruction algorithms discussed not only in this paper, but those reviewed in Uthamacumaran[1] as well.

As mentioned, cancer signals, such as time series gene expression and protein oscillatory dynamics, may reveal rich chaotic structures, such as strange attractors and fractal patterns. The classical approaches to detect such causal structures include: Lyapunov exponents, fractal dimension, entropy, time-delay embedding algorithms, and empirical dynamic modeling (e.g., convergent cross-mapping [CCM]).[147] Lyapunov exponents and fractal dimensions are useful algorithms to verify whether a detected attractor in time series is chaotic. However, entropy deserves special attention and, hence, its discussion has been allocated to the next section on control of chaos. Although formally, a system's degree of chaos is quantified by the magnitude of its largest Lyapunov exponents, they are computationally difficult to estimate from finite, noisy time series measurements. Another fruitful direction in the detection of chaotic systems is the use of spectral frequency decomposition (e.g., Fourier transform, wavelet analysis). Still, these methods too require some other machine algorithms or attractor reconstruction techniques as follow-ups for causality inference in parameter-free, complex systems (note: the term parameter-free is used here to also denote systems for which the governing equations of motions and parameters are unknown).

Time-delay coordinate embedding is the state-of-the-art approach for phase space reconstruction and causal pattern discovery. Nevertheless, they may have dimensionality limits to consider. Let us assume we can empirically measure two variables in a dynamical system (e.g., time point and normalized gene expression level or protein oscillation). Takens's theorem gives us a one-to-one mapping between the original manifold and the reconstructed shadow manifolds via time-lags of the two compared variables.[148] These methods allow us to validate whether two time series variables belong to the same dynamical system (i.e., causally related) (see Figure 7). The points that are nearby on the shadow manifold of one variable will also correspond to those on the other's manifold wherein their nearest neighbors are denser and closer with longer time series (Figure 7). Cross-mapping estimates will further increase the precision of causal structures.[148]

CCM is a causal network inference technique that reconstructs mechanistic couplings in variables when one variable can encapsulate information about a dynamical attractor observed in other patches of the state space.[148] CCM is most applicable for time series with recurrent or oscillatory behaviors, as in the case of cancer cybernetics. CCM relies on two theorems central to the time-delay coordinate embedding of attractors: Whitney's theorem and Takens's theorem. CCM allows one to embed chaotic signals from cancer time series on state space and reconstruct their underlying attractor dynamics. The implications of CCM were recently demonstrated in a study that consisted of patients with markers for cyclic thrombocytopenia, in which multiple cells and proteins undergo abnormal oscillations[149]. An R-software implementation of CCM algorithms is available under the rEDM package (see data and code

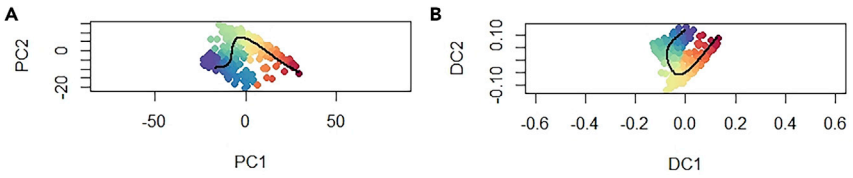

Figure 6. Pseudotime trajectory inference in GBM sample

(A) PCA on a single GBM patient. The PCA plot of a single pediatric GBM sample's pseudotime trajectory inference by the Slingshot algorithm.

(B) Diffusion map on a single GBM patient. Diffusion map plot of a single pediatric GBM sample's pseudotime ordering by the Slingshot algorithm.

availability). The rEDM package can be applied in the embedding of any chaotic time series cancer signal.

An empirical dynamical toolbox was developed for inferring networks of biomarker interactions in complex diseases in pattern space. Krieger et al.[149] combined CCM, transfer entropy, and dynamical mode decomposition (DMD), three non-parametric causal inference techniques, to study attractor dynamics in these complex disease networks (e.g., hemato-immune networks). Transfer entropy is an information theoretic network inference tool that draws on mutual information and seeks to quantify the amount of entropy that is shared between two causally linked time series. DMD attempts to quantify a smaller number of behaviors (modes) of a system by decomposing the time series, which, as an ensemble, fit well to complex exponential functions.[149] Empirical dynamics serve as a blueprint with broader applicability in understanding and treating complex disorders, such as cancers, as dynamical systems. These algorithms for attractor reconstruction should be paired with the network inference and epigenetic landscape reconstruction tools for single-cell datasets discussed herein.

CONTROL OF CHAOS

Entropy, to most systems thinkers, rings a bell to the second law of thermodynamics denoting the tendency of natural processes to evolve toward the increased disorder of the universe. In statistical mechanics and information theory, entropy is used as a measure of complexity. Entropy is a quantitative measure of the uncertainty (predictability) of an outcome. The larger the entropy of a system, the more complex the system tends to be, and more information is required for its description. As mentioned, tumors that exhibit a greater value of entropy production rate have an increased invasive ability, indicating entropy is a quantitative index of the metastatic potential of tumors.[120] However, in dynamical systems, entropy is a robust measure of chaotic dynamics. It emerges in the algorithms used for chaos detection and control. For instance, Toker et al.[150] demonstrated that the chaos decision tree algorithm performs with very high accuracy across a wide variety of complex systems, even in the presence of relatively high levels of measurement noise. The chaos decision tree algorithm approximates the degree of chaos by calculating the permutation entropy of the inputted signal, after it has been de-noised and corrected for possible over-sampling.[150] It has been tested on a model of the transcription of the NF-κB protein complex, adding further algorithmic support to the findings by Heltberg et al.,[5] by comparing the permutation entropy of the original time series to the permutation entropies of random surrogates of that time series.

There are many variants of entropy applicable in the study of complex dynamical systems. For instance, the Kolmogorov entropy, h_K, is related to the Lyapunov characteristic exponents λ_i, as given by:

$$h_K = \int \sum_{\lambda_i > 0} \lambda_i d\mu.$$

It is an information metric to characterize the irregularity in the oscillations of dynamical systems. A positivity in topological entropy is another measure of chaos related to the exponential growth rate of the number of UPOs (unstable periodic orbits) embedded within a chaotic attractor.[151] A set of UPOs can also be thought of as the skeletal framework for chaotic dynamics.[152] Recall that an unstable periodic orbit (limit cycle) can bifurcate into a strange attractor at the onset of chaos. Furthermore, many dynamical averages, such as the Lyapunov exponents, the fractal dimensions, and the entropy, can by efficiently expressed in terms of a sum over the UPOs. The UPOs can be further fine-grained toward the detection of periodic unstable points (PUPs). In principle, detecting these UPOs and PUPs in cancer datasets would reduce the search space for chaotic attractors. The science of navigating the state space for such causal structures and unstable orbits, to control chaotic systems is defined as cybernetics.

Cybernetics is the control theory of the regulatory processes and communication within complex systems.[153] Cybernetic approaches can be designed to stabilize by means of small perturbations one of the many UPOs in a chaotic attractor of a feedback loop system, to switch it toward a more stable attractor. For instance, synchronization of certain chemical oscillations during tumor pattern formation would reduce tumor heterogeneity (i.e., synchronization of chaos). A cybernetician seeks to control the UPOs embedded in a chaotic attractor, to reprogram the dynamical system's chaotic behaviors toward stability, regularity, and predictability. Recall that, due to sensitive dependence on initial conditions, a tiny perturbation (control) can lead to a drastic change in the chaotic system's behavior. While this definition is the generating mechanism of chaos, it is also its weak point used to control chaos. Here we discuss some general cybernetic methods for the detection and control of chaos in dynamical systems. This is easy to conceive in biological systems when target genes or proteins are identified in regulatory networks, which can be controlled by various techniques, such as pharmacological stimulants/inhibitors or gene-editing techniques (CRISPR-Cas9, viral transfection, etc.). Even so, what if we could use certain algorithms to know how much of a stimulant/inhibitor is needed to fine-tune the chemical oscillations of a system? Is there a mathematical framework to control chaos (if it exists) given a specific master regulatory cancer network? Maximal entropy provides

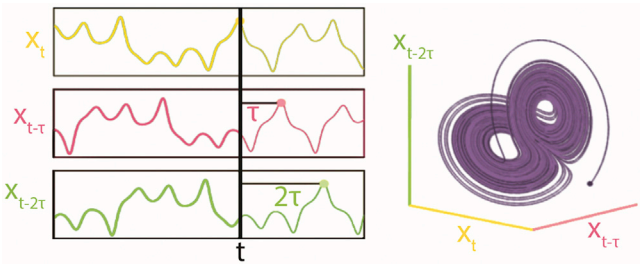

Figure 7. Takens's theorem

The time-delay coordinate embedding of a single variable *X* of a signal's time trace is shown. The time delay is represented by the tau (τ) parameter and *t* corresponds to time. The attractor reconstruction of the embedded signal is shown as a Lorenz attractor.

an answer, at least in part, for short-term time series forecasting and control of such complex systems.

A control chaos algorithm for the detection of PUP's was developed by Ott, Grebogi, and Yorke: the OGY method.[154] In control theory, chaos is defined as a superposition of a very large number of unstable periodic motions.[155] In the OGY algorithm, small forced, periodic kicks are applied to the system to maintain it near the desired unstable periodic orbit of a chaotic attractor. The most studied example is the transformation of periodic motion into chaotic motion, and vice versa, by an external harmonic excitation (periodic kicks) knowing the oscillation amplitudes and frequency spectra of the dynamical system (i.e., feedforward control). Pierson and Moss[156] showed that chaotic attractors can be detected reliably by using these methods on the van der Pol oscillator, known to exhibit chaotic behaviors above a critical threshold. An example of such a method was earlier seen in the findings by Heltberg et al.,[5] in controlling the attractor dynamics of the NF-κB TF kinetics via periodic forcing of TNF amplitudes. In OGY control, we exploit the understanding of the chaotic attractor's unstable manifold structure and the denseness of the UPOs on the attractor, for controllability of the chaos. For instance, the OGY algorithm has been demonstrated in controlling the oscillations of cardiac and brain tissues.[155] Another approach is the Pyragas method based on a time-delay feedback to generate or suppress chaos, as seen in the findings of Khajanchi and colleagues,[47,98] where a time-delayed model of chaotic attractors was demonstrated in tumor growth invasion. Furthermore, works on the stabilization/spatiotemporal control of chaos in reaction-diffusion systems have been documented. For example, Matsumoto and Tsuda[157] demonstrated the ability to suppress spatiotemporal chaos in the Belousov-Zhabotinsky reaction by adding a white noise disturbance. Topological entropy and entropy rates are at the heart of all these methods in chaos detection.

Given entropy is a measure of control in chaos, entropy metrics and entropy rates can be used to reconstruct complex signaling dynamics from empirical datasets. Several single-cell data-processing pipelines have demonstrated that entropy serves as an estimate of chaos detection and, thereby, strange attractors inference if time series datasets are made available. Some entropy-based cancer networks and trajectory inference algorithms are reviewed in Uthamacumaran,[1] all of which can be applied to time series analysis. Examples include the scEpath algorithm discussed above (energy distribution based on maximum entropy), Markov chain entropy (MCE), and single-cell entropy rate (SCENT) algorithms.[158,159] These algorithms use the signaling entropy rate or maximum entropy, as statistical measures to infer cell fate transitions from sCRNA-seq datasets. The entropy measures can be used to project the corresponding Waddington energy landscape (i.e., multidimensional transcriptomic state space) from single-cell datasets and to further infer the master regulatory networks driving the distinct cell clusters on the landscape.[158,159]

Entropy was demonstrated as a robust tool for characterizing the height of the attractors (cell fates), a measure of the cell potency on the landscape.[159] When using the MCE or SCENT algorithms to analyze RNA-seq data, the gene expression matrix (often in the form of a text file or csv file) is read into the algorithm's software space. From it, a matrix object is generated, with rows labeling genes and columns labeling cell barcodes. The data matrix should be normalized i.e., typically via a log-transformation of the form $\log2(\text{counts.m} * \text{sf} + 1.1)$, where counts are the raw expression counts and where sf is a cell-specific scaling factor. The +1.1 pseudocount ensures that empty cells (0s) are avoided in the data matrix if one subsequently uses the single-cell entropy method (zero counts are typical issues in RNA-seq experiments). Hence, the +1.1 will guarantee upon log-transformation that all values of the matrix are >0. A quicker and simpler CCAT (correlation of connectome and transcriptome) method is available in the SCENT-R packages, where a pseudocount of +1 can be used as usual.[160] The CCAT is essentially a Pearson correlation coefficient, which can take on values between -1 and 1, with increasing values indicating higher potency. The CCAT approximation measure is simpler, faster, and proposed to be equally accurate as the MCE measure in estimating single-cell differentiation and/or cell fate transition from gene expression datasets.[160] The entropy rate values can be clustered into various phenotypic cell states on an energy landscape. This is accomplished with the "InferPotencyStates" function, which fits a mixture of Gaussian distributions to log-transformed entropy rate values, or *Z* score-transformed CCAT values. The phenotypic cluster estimate obtained using these entropy measures can be integrated with diffusion maps or other dimensionality reduction techniques to assign lineage trajectories.[160]

These entropy-based inference tools are integrated in a similar user-friendly MatLab interface in the scEpath algorithm.[158] scEpath is a computational algorithm for mapping the three-dimensional energy landscape of cells and visualize their transcriptional dynamics. The algorithm allows us to estimate the transition probabilities between cell states from patient-derived scRNA-seq datasets.[158] The lineage trajectories and pseudotemporal ordering can be visualized from single-cell gene expression data using the scEpath algorithm.[158] The scEpath algorithm reconstructs a three-dimensional energy landscape for the single-cell gene expression dataset by incorporating nearest neighbor correlations based on maximum entropy. In addition, scEpath performs downstream analyses, including unsupervised clustering of differentially expressed marker genes or TFs, cell lineage hierarchy clustering by minimally directed spanning tree, and pseudotime trajectory inference.

To visualize the energy Waddington-like landscape, scEpath performs PCA analysis on the energy matrix $E = (E_{ij})$ and fits a surface using piecewise linear interpolation over the first two

PCA components and energy of each cell. The energy of each attractor (cell state), E_j, on the landscape is computed according to the following function:

$$E_j(y) = \sum_{i=1}^{n} E_{ij}(y) = -\sum_{i=1}^{n} y_{ij} ln \frac{y_{ij}}{\sum_{k \in N_i} y_{kj}},$$

where y_{ij} represents the normalized gene expression level (between 0 and 1) of gene I and cell j, and $N(i)$ is the neighborhood of node I in the network. Each gene is assigned a local energy state E_{ij}. The local energy is then normalized for each cell. The normalized single-cell energy is projected on the PCA pattern space to obtain the Waddington landscape reconstruction. To identify the distinct cell states (attractors), structural clustering is performed using an unsupervised framework called single-cell interpretation via multikernel learning.[158] A probabilistic directed graph network is reconstructed by the scEpath algorithm to infer cell state transitions on the Waddington landscape.

Although such Waddington landscape reconstruction algorithms exist, they remain vastly unutilized in cancer gene expression datasets, both static (single-time frame) and time series (the absence of which itself is the grand problem). To demonstrate their applicability, the scEpath-generated epigenetic landscapes of the various GBM patient samples are shown in Figures 8A–8D. The vertical axis, scEpath, denotes the energy potential of the epigenetic landscape, i.e., a measure of the cell fate's differentiation potential, whereas the PCA components 1 and 2 indicate the reduced coordinates of the transcriptome (i.e., state-space analog). The cell states from a few GBM patients' energy landscape shows distinct cell fates, the colors indicating tumor heterogeneity in Figure 8A, corresponding to the same data visualized in Figure 3A. The higher the cell states are on the energy landscape, the higher their potency (differentiation potential). Due to the lack of time series gene expression analysis, the cell states are stuck in their valleys (i.e., fixed-point attractors). Nonetheless, in time series, certain cell fates may fluctuate within or in between distinct valleys periodically (i.e., limit cycles) or exhibit aperiodic trajectories on the energy landscape with a fractal dimension (i.e., strange attractors). Figure 8B shows the Waddington landscape of a single GBM patient corresponding to the results in Figures 4B and 6A–6C (slingshot results). As seen in Figures 8B and 8A dynamical structure can be inferred even without time series by the energy distribution of cell states. The landscape topography explains the two different pseudotime trajectory curves seen in Figures 6A and 6B. These two trajectory inference curves can be superimposed on the trajectory inference seen on Figure 8B. The energy landscape of a few adult GBM patient samples is shown in Figure 8C. There seems to be three or four distinct cell states, which can be interconnected by a lineage bifurcation from the cell clusters at the top of the landscape to the two local minima.

Finally, recall that Richards et al.[3] found that the transcriptional states of their 28 adult GSC samples were divided into two distinct states. When the scRNA-seq counts of a single adult GSC sample were analyzed using the Seurat algorithm, as shown in Figure 3C, their findings are confirmed. However, their data looked at the dimensionally reduced clustering of the cells. In Figure 8D, when the same data count matrix for a single GSC sample is visualized on the energy landscape in three-dimensions using the scEpath algorithm, a dynamical structure connecting the two-states is inferred, indicating that they may not be two distinct states but rather a dynamical attractor may intertwine them. To truly visualize these cells as dynamical systems and test the presence of dynamical attractors in their energy landscape, the dynamical structures inferred from these results shown must be verified with time series datasets. When observing these energy landscapes, think of how will they look if they were played as a video instead of being observed as a single picture.

All discussed network inference and Waddington landscape reconstruction algorithms must be applied to time series cancer datasets to infer causal structures and experimentally detect the presence of chaotic attractors. It would be of significant clinical interest to perturb the cancer biopsy of a patient with a target-specific therapy identified by the network clustering algorithms and obtain a time series transcriptomics of the patient's cancer cell population under the therapy response. Therapies can be of any form, ranging from target-specific drugs to immunotherapy regimens. The epigenetic landscape can be reconstructed for these therapy-perturbed samples, and their corresponding attractor dynamics can be visualized corresponding to their molecular network dynamics. If chaotic attractors are observed, the genes/proteins corresponding to the emergence of these patterns can be identified using these network approaches. The identified targets responsible for the emergence of chaotic attractors are predicted to confer cancer stemness, tumor relapse, and therapy resistance.

CAUSALITY INFERENCE VIA MACHINE INTELLIGENCE

Perhaps the most celebrated cybernetic method designed to study the time series of complex systems are Artificial Neural Networks (ANNs).[161] Single-cell datasets are highly noisy and sparse. Thus, the above-discussed network visualization and epigenetic landscape reconstruction algorithms should be paired with a more robust machine intelligence capable of deciphering patterns in noisy time series datasets. ANNs may provide the solution. They resemble multipartite graphs where the graph is separated into different sets of nodes, and the nodes in each set can share edges between different sets but not within itself.[161,162] Each set of nodes corresponds to a layer in the network architecture, which can be arranged in a multitude of ways. The data propagation can be feedforward or backward through the network. If the ANN learns to recognize patterns as the information is iterated through its multi-layers, we call it deep learning.[163]

The design of the ANN mimics the brain's neuronal information processing, in a simplified (linearized) manner. To illustrate, consider the linear equation: $y = mx + b$, as the neuron structure of a simple ANN, such as a perceptron. Here, the output y is the sum of the bias (b) and the input data (x) times the weight (m). The activation function, such as tan$h$, (leaky) ReLu, sigmoid, softmax, etc., then decides if a given neuron's output y is activated to propagate the information toward the function.[164] The coupling of a collection of such neurons forms a layer, and a bunch of stacked layers forms the ANN allowing the neural network to pass the information toward a value that is closer to the target (label) data on which the network is trained on. The training is done via adjusting the weights with a loss function, such as the minimization of the error using gradient descent learning, nearest neighbors optimization, or backpropagation algorithms.[162,164]

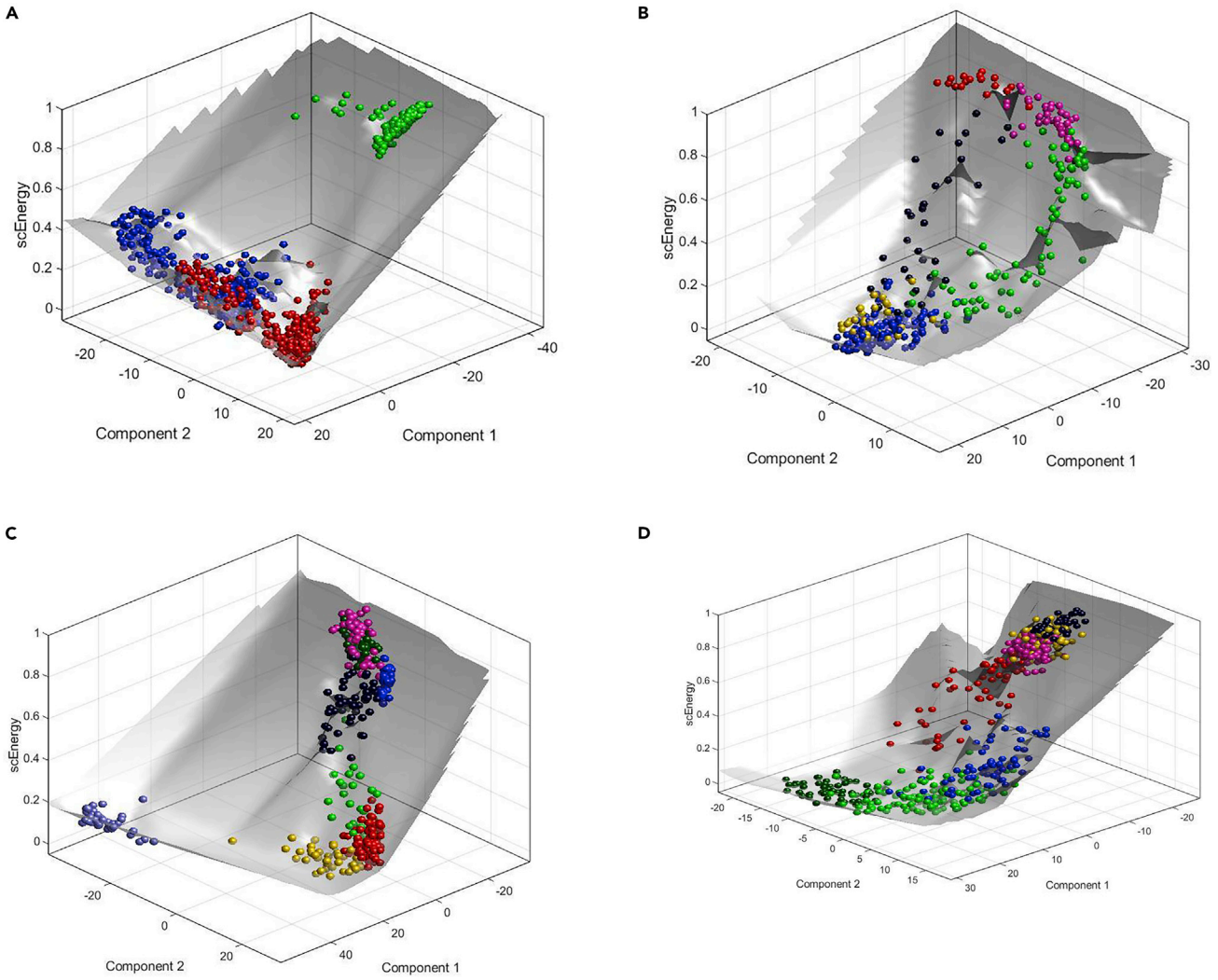

Figure 8. Waddington landscape reconstruction of GBM and GSC samples

(A) Energy landscape of a few pediatric GBM patients. scEpath energy landscape for a few pediatric GBM patient samples.
(B) Energy landscape of a single pediatric GBM patient. scEpath energy landscape for a single pediatric GBM sample.
(C) Energy landscape of few adult GBM patients. scEpath Waddington landscape for a few adult GBM samples.
(D) Energy landscape of a single GSC sample. scEpath epigenetic landscape for a single GSC sample.

In addition to adjustable parameters, there are hyperparameters that specify the global properties of the ML system, such as the learning rate, regularization, duration of memory, strength of nonlinearity, and layers of neuronal architecture. The amount of weight adjustment to minimize the loss function is defined as the optimization function, and various hyperparameter tunings are required to settle to the optimal solution, which is by default an NP-hard problem requiring user-defined trial-and-error tuning.[165] Neural networks may exhibit many of the general problems faced by ML, such as overfitting, bias-variance trade-off, training generalization in pattern recognition, etc. The problem of overfitting is overcome with performance predictors, such as cross-validation and hyperparameter tuning, in addition to reproducible, large sets of data needed for the training.

Machine algorithms and in specific, neural networks, are emerging as state-of-the-art approaches in time series forecasting and classification of complex systems, including gene expression dynamics.[166] Once again, none have been applied in cancer research due to the lack of cancer time series datasets. In principle, the neural networks discussed below can and should be utilized to study attractor dynamics from time series cancer datasets. Common machine algorithms for computing attractors are stochastic searching algorithms, those which start with a randomly selected initial state and finish with exhaustive search of the state space of a network, such as Monte Carlo algorithms and tree-based search algorithms. The computational complexity of these methods can grow exponentially with respect to the number and length of the attractors. Despite that, recent advancements in searching algorithms in combination with deep learning architectures can potentially overcome these barriers. For instance, one of the most concrete examples of adaptive emergent behaviors in artificial intelligence (AI) is Google DeepMind's AlphaGo algorithm (and more recently the AlphaFold algorithm). AlphaGo can optimize the possibilities of a vast search space

effortlessly during the Go game. By use of a Monte Carlo tree search algorithm it finds its moves based on knowledge previously acquired by a deep learning ANN.[167] These progresses in machine intelligence demonstrate pairing deep learning algorithms with search tree algorithms may provide ways to find context-specific causal dependencies in complex (cancer) datasets.

Recently, deep learning architectures have shown applications in inferring gene expression patterns. The following studies are surveyed here to stimulate the use of such neural network algorithms in the investigation of time series cancer datasets. Some examples include the use of convolutional neural networks (CNNs)[168] and the Decode algorithm.[169] To decode differential gene expression, the Decode algorithm predicts personalized (patient tissue-specific) gene-regulatory interaction modules. Co-expression analysis of transcriptomic data is achieved via XGBoost-based hyperparameter tuning and training of the deep learning network.[169] The AttentiveChrome algorithm shows consistent improvement over its predecessor DeepChrome's predictive performance in assessing positional histone information in gene expression analysis by coupling a CNN to hierarchical long short-term memory (LSTM) modules.[170]

Another deep learning algorithm, D-GEX, was shown to outperform classical ML classifiers, such as linear regression and KNN clustering algorithms in gene expression analysis. Starting with a smaller input set of genes, it can predict the output of a higher number of target genes. For upscaling gene expression, D-GEX uses a multi-task multi-layer feedforward neural network with a nonlinear activation function (tan*h*) and hyperparameter tuning.[171] Schmauch et al.[172] used deep learning algorithms to reconstruct RNA-seq gene expression profiles from whole-slide tumor sections demonstrating a potent application of digital pathology for gene expression inference. HE2RNA, the deep learning algorithm, was devised for the prediction of specific molecular phenotypes containing microsatellite instabilities and the transcriptomic profiles from the corresponding high-definition whole-slide images.[172] In essence, these algorithms can be adapted to time series tumor imaging techniques.

Deep learning architectures using capsule networks called scCapsNet were shown to reconstruct single-cell gene expression modules via identifying hierarchical relationships in cellular sub-types from single-cell transcriptome analysis.[173] The scCapsNet algorithm used two-dimensional PCA on the internal weight parameters in the feature extraction layer and two-dimensional t-SNE visualization for the embedding representation of all the genes and clustered modules.[173] Nevertheless, the implementation of capsule networks depends critically upon the availability of large, high-quality datasets. The problem of zero counts or low RNA captures in the technique also remain as technical roadblocks. In addition, in the scCapsNet's mathematical framework, the leakage transcription rate and degradation rates of different proteins were considered identical. A random search for chaotic states in the network was found to be a rare event (low in frequency). The Hill equations used in this deep learning algorithm further iterate how both theoretical and experimental/computational methods must be merged to decipher complex systems. The set of Hill equations used in scCapsNet are as follows:

$$\frac{dp_i}{dt} = f_i(p) - p_i,\ where\ f_i(p) = \begin{cases} A_i(p) \\ R_i(p) \\ A_i(p)R_i(p) \end{cases}$$

$$(p) = (p_1, \ldots, p_N)$$

$$A_i(p) = \frac{Act_i^h}{Act_i^h + K^h},\ R_i(p) = \frac{K^h}{Rep_i^h + K^h}$$

$$Act_i = \sum_{j=1}^{N} \alpha_{ij} p_j,\ Rep_i = \sum_{j=1}^{N} \beta_{ij} p_j,\ (i,j = 1,2,\ldots N).$$

In this simplified gene expression model, $p_i \in [0, 1]$ denotes the expression level of gene *I*, the adjacency matricesα_{ij} and β_{ij} determine the network structure of the system defined such that $0<\alpha_{ij}<1$ if gene *j* activates gene *i*, and $0<\beta_{ij}<1$ if gene *j* inhibits the *i*-th gene, and both equal zero when there is no regulation of the *i*-th gene by the *j*-th gene. Act and Rep correspond to the sum of active/repressive transcriptional factors to node *i*. The regulatory interactions are defined by Hill functions with the Hill (cooperativity) coefficient *h* and activation coefficient *K*. This model demonstrates that, in principle, one can adopt any of the earlier discussed mathematical models for gene expression inference or pattern formation mapping into a deep learning framework.

As discussed, classical chaos detection tools, such as Lyapunov exponent estimates, fractal dimension, and entropy rates, can confirm the presence of chaotic attractors from time series datasets only if such attractors are detected. Classical chaos detection methods, such as Lyapunov exponents and fractal dimensions, can be applied on the time-embedded signal to infer chaotic attractors. Furthermore, these general techniques may only apply to low-dimensional systems, such as analyzing the time series of a single gene, a well-defined network of very few genes, a single TF's oscillation, or the time trace of a small network of proteins. How do we infer Lyapunov exponents from the time series of a large-scale, multi-dimensional system (think of thousands of gene expressions or protein oscillations in thousands of cells at once)? RNNs and deep learning networks are emerging as robust solutions to this problem. Recurrent neural networks (RNNs) are recently emerging as the state-of-the-art machine algorithms for the spatiotemporal prediction/forecasting of chaotic dynamics and attractor reconstruction in complex time series. For example, reservoir computing (RC), a type of RNN has recently demonstrated applicability in the prediction of multi-dimensional time series of spatio-temporally chaotic systems, such as the forecasting of the Kuramoto-Sivashinsky equation up to a few multiples of the Lyapunov time.[174,175]

Pathak et al.[174,175] exploited the reservoir dynamics to find the Lyapunov exponents of high-dimensional dynamical systems, from which chaotic attractors could be reconstructed. RC is an implementation of ANN, where the reservoir is a network of *D* neuron-like units. Each node of the reservoir *i* has multiple inputs and outputs, and a scalar state denoted by $r_i(t)$. The weighted connections between the nodes are represented by an adjacency matrix A: D x D. The input is coupled to the reservoir through a fixed randomly generated input matrix. The training phase is given by: $r(t + \Delta t) = \tanh(Ar(t) + W_{in}u(t))$, where *W* is

the weights, and $u(t)$ is the training dataset, following which we find the output weight matrix that minimizes the loss function: $\sum_{t=-T}^{0}[norm(W_{out}r(t)-u(t))]^2$.[175] Such methods remain to be explored in cancer time series datasets and are suggested here as possibly the most powerful algorithm discussed in this paper prospective of capturing cancer attractor dynamics.

While autoregression models are better suited for short-term time series forecasting in low-dimensional systems, deep learning algorithms are the state-of-the-art approaches for time series gene expression forecasting in higher-dimensional systems.[166] Furthermore, neural network algorithms used in turbulence forecasting may be of great interest in the application to cancer dynamics prediction. If the AI can predict causal patterns, such as strange attractors in turbulent flows, then it should (in principle) within cancer datasets as well. A general AI-based approach to forecast turbulence-like chaotic systems is to perform singular value decomposition followed by dimensionality reduction and dynamic mode decomposition (SVD/PCA/DMD) on the time series dataset. RNN architectures, such as RC or LSTM, can then be used to forecast the evolution of the coefficient of the modes.[176] Another approach would be to use CNNs, for which backpropagation through time (BPTT) algorithms are used instead of RC computing.[176,177] Regularization procedures utilizing BPTT were also shown to be more effective in chaotic attractor reconstruction.[176,177] CNN-based methods may be better suited for the classification of time series gene expression dynamics, while RNN algorithms, such as RC, are best for time series forecasting/attractor reconstruction.[166] RC is computationally cheaper to train in the case of full-state information acquisition whereas, gated architectures are better for reduced order states/observables.

On a final note, with regard to AI, in 2019, Google confirmed its quantum processors had reached quantum supremacy.[178] Quantum supremacy is the alleged claim that quantum computers can perform certain computational tasks or solve certain problems exponentially faster on a quantum processor than on a classical processor in a feasible amount of time.[179] Quantum computing (QC) allows a vast search space for solving combinatorially complex optimization and decision problems, via exploiting the fundamental properties of quantum mechanics, such as superposition and the entanglement of information (i.e., allows parallel computation). While a lower time complexity than classical algorithms is implied, whether there are any computability advantages remains highly debated and anticipated. The prospects of QC and quantum machine learning in biological networks/attractor reconstruction can only be verified in time.

KOLMOGOROV COMPLEXITY

Current approaches to the inference of cancer networks have computational complexity-imposed limitations. For instance, the hyperparameter optimization of machine learning algorithms and modularity optimization in community detection are *NP-hard problems*, requiring user-defined brute-force searching and heuristics. Furthermore, these methods may reveal some of the statistical properties of complex networks, but not necessarily their dynamics. Information theoretic measures, such as the maximal Shannon entropy and entropy production rates, are powerful indicators of chaotic dynamics in time series analysis when combined with other approaches, such as Lyapunov exponents and energy landscapes. Entropy allows us to visualize the state-space dynamics in terms of information flows in phase space. However, Shannon's *entropy is not a robust measure of a graph network's complexity*. It fails to capture the algorithmic content of a complex system. There remains a powerful tool in computational complexity theory, vastly unexplored in the causality inference of cancer networks. The *Kolmogorov complexity*, also known as algorithmic complexity, is a descriptive measure of a network/data object's complexity.[180] K-complexity is analogous to the Occam's razor, wherein the simpler or more compressible a data object, the shorter the program length p required to compute it or capture its information content when run onto a universal Turing machine.[180,181]

Assume there is a certain dataset (e.g., a gene expression adjacency matrix) that can be represented as a binary encoded text or string for some universal Turing machine. The length of the shortest program in a defined language that can generate this string (data object) as its output, denoted as s, is defined as the Kolmogorov complexity $K(s)$.[182] That is, $K(s)$ is the length of the shortest description of the data object (network or system) in program space. Despite that, the Kolmogorov complexity of the system is in principle, incomputable (can only be estimated).[182] Then, the expected probability $P(s)$ that a random program p running on the Turing machine produces the string s upon halting is given by the Solomonoff-Levin algorithmic probability measure: $P(s)=\sum_{p:T(p)=s}\frac{1}{2^{|p|}}$. The algorithmic probability is a mathematical method of assigning a prior probability to a given observation.[183] The coding theorem method (CTM), then establishes an equality between the algorithmic probability $P(s)$ and the Kolmogorov complexity $K(s)$, formally expressed as:

$$P(s)=2^{-K(s)}+c,$$

where c is some constant or, equivalently, $K(s)\approx -log_2P(s)$. The CTM states that an object of a short computer program is more likely to be generated at random than an object generated by a longer computer program (i.e., complex systems).[184]

Due to the incomputability of $K(s)$, computational algorithms are required for indirectly approximating the K-complexity, such as lossless compression algorithms (such as the Lempel-Ziv-Welch algorithm), the Block Decomposition Method (BDM), Minimal Information Loss Sparsification, and Maximal Algorithmic Randomness Preferential Attachment.[182,185] The lossless compression algorithms are similar to Shannon entropy (a poor estimate of K-complexity). Their applications in estimating graph complexity remain limited. BDM is the most robust of all techniques for estimating the $K(s)$ complexity of a graph network (see BDM code in the data and code availability section). K-Complexity approximation algorithms are emerging in the detection of causal structures in complex networks within algorithmic information dynamics (AID). AID is a branch of algorithmic information theory emerging as an AI platform for causality inference in complex datasets.[182] AID is a promising tool for studying the cybernetics and attractor landscapes of complex disease networks using perturbation analysis in software space.[182,185] A higher K-complexity implies non-randomness (if not lower than the bits of information encompassing the system). As such, complexity measures allow

us to distinguish chaos from randomness. It is proposed as an alternative approach for detecting strange attractors and inferring chaotic solutions in cancer datasets, in addition to the ANN-based approaches discussed above. The algorithms for approximating the Kolmogorov complexity $K(s)$ should be incorporated into the machine learning pipelines and algorithms for single-cell gene expression dissection.

CONCLUSIONS

To conclude, a blueprint of chaotic dynamics and their patterns in the state space of time series cancer signaling is provided. Cancers are complex cybernetic systems which may exhibit strange attractors in their gene expression signaling and pattern formation (i.e., reaction-diffusion systems). The description of complex systems, such as cancer stemness networks requires a fundamental understanding of nonlinear dynamics, fractals, and chaos. Chaos, despite appearing to be random-like, may serve as a robust biomarker for tumor complexity and is bound to well-defined pattern structures in state space: strange attractors. Chaotic dynamics are efficient methods to control the gene expression of lowly activated downstream genes and signaling cascades (cross-talks) required for complex cellular processes involved in homeostasis, development, phenotypic plasticity, and survival. Chaotic attractors may be the hallmarks of therapy resistance, tumor recurrence, and cancer stemness. Although mathematical cancer models have demonstrated that the emergence of chaotic attractors may be indicators of aggressive (adaptive) cancer states, their detection from empirical datasets remains primitively explored. As such, it is suggested herein that the study of chaotic dynamics and the detection of strange attractors in the signaling state space of cancer stemness networks, at the cellular level and at the individual gene/protein levels, via the pairing of empirical and computational approaches, may provide a fruitful direction for precision oncology. Time series molecular datasets (gene expression profiling and protein quantification) is severely lacking in cancer research. Therefore, the generation of time series scRNA-seq and bulk RNA-seq from cancer patients are strongly encouraged.

The mathematical models pertaining to the complex dynamical processes governing cancers have been revealed in this paper to serve as a primer for the empirical investigation of chaotic dynamics in cancer processes. Various algorithms for the detection of chaotic attractors in time series cancer datasets have been surveyed. The general road map to detecting a chaotic attractor (if it exists) in cancer signaling dynamics is as follows. First the time traces of the signal of interest, such as the gene expression or protein oscillation, is acquired. The time series is then subjected to a time-delay coordinate embedding to be visualized in a three-dimensional space. Following, the fractal dimension and positive Lyapunov exponent(s) can be computed on the time-embedding to verify if the detected attractor (if it exists) is chaotic. Empirical dynamic modeling (e.g., CCM), Waddington landscape reconstruction, deep learning architectures, RNNs, and AID are suggested as powerful tools for complex attractor reconstruction, which remain vastly unexplored in deciphering cancer time series datasets. The various algorithms and tools discussed in each section should be combined and contrasted for realistic cancer modeling.

The investigation of cancers as dynamical systems may provide innovative targeted therapies for relapsed/refractory cancers in the emerging frontier of computational oncology. Finally, in principle, if strange attractors are empirically verified in the state-space dynamics of cancer regulatory networks, and chemical pattern formation, the control of chaos may pave the reprogramming of cancers (stem) cells to benign/healthy cell fates (Box 1). More precisely, the detection of globally existent strange attractors in the state space of cancer stemness networks are suggested as the directions pattern scientists should take to reprogram cancer (stem) cells to benignity.

Data and code availability

Source codes for scRNA-seq datasets

Columns are cell barcodes, row names are genes, in all expression (count) matrices.

(1) https://singlecell.broadinstitute.org/single_cell/study/SCP393/single-cell-rna-seq-of-adult-and-pediatric-glioblastoma#study-summary.[2] Expression Matrix (Expression Matrix [log2(TPM/10 + 1)] (Smartseq2): IDHwtGBM.processed.SS2.logTPM.txt.gz

(2) https://singlecell.broadinstitute.org/single_cell/study/SCP503.[3]

Count matrix. Richards_NatureCancer_GSC_scRNAseq_-counts.csv.gz.

The corresponding metadata files can be used to subset the count matrix to cells from a single sample. A subsetted raw count matrix was used in the analysis, GSC sample (BT127_L, 500 random cells). All cells in the gene expression matrix are malignant. All samples are of good technical quality and have been quality checked (see the Methods sections in corresponding citations). Refer to Figures 1 and 2 in Richards et al.,[3] to determine which GSC sample is of your interest based on cluster number, gradient position, etc.

Source codes for demonstrated algorithms

https://satijalab.org/seurat/v3.2/pbmc3k_tutorial.html (Stuart et al.[142]) [Seurat Algorithm];

http://www.bioconductor.org/packages/release/bioc/html/slingshot.html (Street et al.[143]) [Slingshot algorithm]

https://github.com/sqjin/scEpath (Jin et al.[158]) [scEpath algorithm]

https://mran.microsoft.com/snapshot/2018-06-22/web/packages/rEDM/vignettes/rEDM-tutorial.html (Sugihara et al.[148]) (rEDM package and tutorial for Convergent Cross Mapping)

http://complexitycalculator.com/(Zenil et al.[182]) (BDM calculator).

Figure codes

The Lorenz attractor on Figures 1 and 2 were rendered from the Wolfram Demonstrations Project.[186] Thanks to Rik Bhattacharja for redesigning these two figures, the graphical abstract, and Figure 7. Figure 7 was rendered from the rEDM package and tutorial website given above.[148] All other (data) figures corresponding to the Seurat, Slingshot, and scEpath algorithms were generated by the author.

Box 1. Glossary

Attractors: as the word implies, they are regions or shapes in phase-space to which the trajectories of a dynamical system are attracted to (pulled towards) for a wide range of initial conditions. Attractors are universal patterns characteristic of a dynamical system's evolution in state space. They may correspond to cell states on an epigenetic landscape or patterns at the single-molecular dynamics, such as the oscillation of a single gene or transcription factor (note: repellors are the opposite analog of attractors, where the dynamical trajectories (flow) is in the opposite direction of an attractor).

Cancers: cancers are diseases caused by the uncontrolled cell division of malignant cells. The malignant cells emerge from genetic mutations/translocations and result in aberrant (dysregulated) gene expression patterns. These patterns revert the tumor microenvironment partially or completely to an embryonic-like developmental state (dedifferentiation). Emergent hallmarks of cancers include uncontrolled cell division, telomerase overexpression/re-lengthening, immune system hijacking, apoptosis/death evasion, rewired (epi)metabolomes, and metastatic invasion.[21]

Cancer stem cells: stem cells are immature cells that exhibit unlimited replicative potential (self-renewal) and the ability to develop (differentiate) into many other cell types (i.e., potency). Embryonic stem cells have the highest potency and can differentiate into any somatic (adult) cell type, by the combined expression of a relatively few transcription factors. Similarly, a subset of cancer populations, expressing specific stemness-related transcription factors and surface markers have stem cell properties, analogous to an embryonic stem cell. These cells are known as cancer stem cells (CSCs). The interconvertibility of non-stem (mature) cancer cells and CSCs, make cancers a complex dynamical system.

Cell fate-transitions: also referred to as state transitions or phenotypic switching. In developmental biology and cancer processes, this refers to the differentiation of a cell from one phenotype to another. Stochastic models naively assume cell state transitions can be modelled by Fokker-Planck equations. This review outlines that these are complex dynamical systems requiring the study of attractors dynamics.

Chaos: chaos is aperiodic, irregular long-term behavior in a system that exhibits sensitive dependence on initial conditions and exponential trajectory separation. Chaotic dynamics exhibit causal structures, namely, dense fractal attractors bound in state space (i.e., strange attractors).

Complex systems: complex systems are irreducible dynamical systems composed of many interconnected parts, the collective behaviors, and properties of which are more than the sum of its local interactions (i.e., nonlinear).

Complexity science: complexity science, also known as complex systems theory or sometimes complexity theory, is the emerging interdisciplinary study of complex systems through computational physics. It evolved from dynamical systems theory and now comprises of the study of patterns in complex systems through the lens of nonlinear dynamics, chaos, fractals, nonequilibrium systems, networks, and algorithms. As such, herein it is also referred to as pattern science. Complexity science merges large-scale empirical data on complex systems with computational algorithms and mathematical modelling.

Ecosystem: the ecological relationships (e.g., predator-prey dynamics) among the various cell types and environmental factors composing the tumor microenvironment. It resembles a stem cell niche, however, and includes the non-stem (mature, differentiated) cancer cells, healthy cells, and the abiotic signaling factors (e.g., molecules, proteins, secretory vesicles, circulating tumor DNA, epigenetic modulators, etc.).

Fractal: a geometric structure or pattern observed in chaotic systems in which each part exhibits a statistical self-similarity to the whole across many length scales (or scale-invariance). Some examples include snowflakes, trees, coastlines, tumor contours, and fluid turbulence. The latter two are considered multifractal systems. The fractal dimension is a non-integer dimension to characterize fractals. When a single fractal dimension is insufficient to describe the fractal geometry of a system/structure, it is considered a multifractal.

Glioblastoma multiforme: a type of glioma. Gliomas, deadly lethal diseases responsible for cancer-related pediatric mortality, is used as a model system to describe cancer networks as dynamical systems herein. Three clinically distinguished glioma types exist, namely, oligodendroglioma (isocitrate dehydrogenase [IDH]-mutated and 1p/19q co-deleted), IDH-mutated astrocytoma, and IDH wild-type astrocytoma, which include the most malignant glioma-type, glioblastoma (GBM).[4]

Pseudotime: a measure of how a cell transitions between different states in some cancer-related process. Examples include the trajectory of a cell on an epigenetic landscape during differentiation, in response to environmental stimuli, a drug perturbation or decision-making processes. Only with time series datasets can real attractor dynamics and cell state trajectories be visualized.

State space: also referred to as phase -space or attractor space is the space in which all possible states of a dynamical system are represented. The axes correspond to the specified variables of the dynamical system, and the state of the system can be represented by a state vector X.

Strange attractors: also known as chaotic attractors, are attractors of certain nonlinear, aperiodic systems which exhibit sensitivity to initial conditions (i.e., chaotic systems). Trajectories on a strange attractor appear to be random-like in time series. However, they have a defined causal structure to which the trajectories of the dynamical system are bound to in state space. The strange attractor has a fractal dimension. In hyperchaotic systems, strange attractors may exhibit multifractality.

Trajectory inference: also known as pseudo-temporal ordering is a computational algorithm used in single-cell transcriptomic data processing pipelines to infer the patterns of a dynamic cellular process (e.g., differentiation).

(Continued on next page)

Box 1. *Continued*

Waddington's landscape: Waddington's epigenetic landscape is an abstract metaphor frequently used to represent the relationship between gene-regulatory network dynamics and their corresponding cell fates during developmental processes. Here, the epigenetic landscape is used as a state-space projection of cancer-related processes in which cell fate transitions can occur. As such, the terms attractor landscape and energy landscape are also interchangeably used. Most often, the epigenetic landscape is the transcriptomic state-space reconstruction of gene expression datasets, with some energy function (quasi-potential) as its vertical axis defining the landscape topography.

ACKNOWLEDGMENTS

My sincere thanks to Dr. Mario D'Amico of Concordia University, Dr. Phil Gold, Dr. Rolando Del Maestro, Dr. Goffredo Arena, and Dr. Mohamed Abdouh of McGill University, and to Dr. Morgan Craig of Université de Montreal, for being my initial conditions. This article is dedicated to the innate systems thinkers [children] at the Glen Hospital, B7- Oncology (Montreal, QC) and their paintings, which have served as the greatest teachings of chaos, complexity, and turbulence.

REFERENCES

1. Uthamacumaran, A. (2020). A review of complex systems approaches to cancer networks. Complex Syst. *29*, 779–835, https://doi.org/10.25088/ComplexSystems.29.4.779.

2. Neftel, C., Laffy, J., Filbin, M.G., Hara, T., Shore, M.E., Rahme, G.J., Richman, A.R., Silverbush, D., Shaw, M.L., Hebert, C.M., et al. (2019). An integrative model of cellular states, plasticity, and genetics for glioblastoma. Cell *178*, 835–849.e21, https://doi.org/10.1016/j.cell.2019.06.024.

3. Richards, L.M., Whitley, O.K.N., MacLeod, G., Cavalli, F.M.G., Coutinho, F.J., Jaramillo, J.E., Svergun, N., Riverin, M., Croucher, D.C., Kushida, M., et al. (2021). Gradient of developmental and injury response transcriptional states defines functional vulnerabilities underpinning glioblastoma heterogeneity. Nat. Cancer *2*, 157–173, https://doi.org/10.1038/s43018-020-00154-9.

4. Jung, E., Alfonso, J., Osswald, M., Monyer, H., Wick, W., and Winkler, F. (2019). Emerging intersections between neuroscience and glioma biology. Nat. Neurosci. *22*, 1951–1960, https://doi.org/10.1038/s41593-019-0540-y.

5. Heltberg, M.L., Krishna, S., and Jensen, M.H. (2019). On chaotic dynamics in transcription factors and the associated effects in differential gene regulation. Nat. Commun. *10*, 71, https://doi.org/10.1038/s41467-018-07932-1.

6. Halatek, J., and Frey, E. (2018). Rethinking pattern formation in reaction-diffusion systems. Nat. Phys. *14*, 507–514, https://doi.org/10.1038/s41567-017-0040-5.

7. Ovchinnikov, I.V. (2016). Introduction to supersymmetric theory of stochastics. Entropy *18*, 108, https://doi.org/10.3390/e18040108.

8. Aihara, K. (2008). Chaos in neurons. Scholarpedia *3*, 1786, https://doi.org/10.4249/scholarpedia.1786.

9. Rasmussen, R., Jensen, M.H., and Heltberg, M.L. (2017). Chaotic dynamics mediate brain state transitions, driven by changes in extracellular ion concentrations. Cell Syst. *5*, 591–603.e4, https://doi.org/10.1016/j.cels.2017.11.011.

10. Mandelbrot, B., and Hudson, R.L. (2004). The Misbehavior of Markets: A Fractal View of Financial Turbulence (Basic Books).

11. Allen, J.C., Schaffer, W.M., and Rosko, D. (1993). Chaos reduces species extinction by amplifying local population noise. Nature *364*, 229–232, https://doi.org/10.1038/364229a0.

12. Shih, H.-Y., Hsieh, T.-L., and Goldenfeld, N. (2016). Ecological collapse and the emergence of travelling waves at the onset of shear turbulence. Nat. Phys. *12*, 245–248, https://doi.org/10.1038/nphys3548.

13. Stein, P.D., and Sabbah, H.N. (1976). Turbulent blood flow in the ascending aorta of humans with normal and diseased aortic valves. Circ. Res. *39*, 58–65, https://doi.org/10.1161/01.RES.39.1.58.

14. Goldberger, A.L., Rigney, D.R., and West, B.J. (1990). Chaos and fractals in human physiology. Sci. Am. *262*, 42–49, https://doi.org/10.1038/scientificamerican0290-42.

15. Philippe, P. (1993). Chaos, population biology, and epidemiology: some research implications. Hum. Biol. *65*, 525–546.

16. Goldberger, A.L. (2006). Giles F. Filley Lecture. Complex systems. Proc. Am. Thorac. Soc. *3*, 467–471, https://doi.org/10.1513/pats.200603-028MS.

17. Sedivy, R., and Mader, R.M. (1997). Fractals, chaos, and cancer: do they coincide? Cancer Invest. *15*, 601–607.

18. Coffey, S.D. (1998). Self organization, complexity, and chaos: the new biology for medicine. Nat. Med *4*, 882–885, https://doi.org/10.3109/07357909709047603.

19. Thurner, S., Klimek, P., and Hanel, R. (2018). Introduction to the Theory of Complex Systems (USA: Oxford University Press).

20. Ladyman, J., and Wiesner, K. (2020). What Is a Complex System? (New Haven: Yale University Press).

21. Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674, https://doi.org/10.1016/j.cell.2011.02.013.

22. Fang, X., Kruse, K., Lu, T., and Wang, J. (2019). Nonequilibrium physics in biology. Rev. Mod. Phys. *91*, 045004, https://doi.org/10.1103/RevModPhys.91.045004.

23. Shalizi, C.R. (2006). In Methods and Techniques of Complex Systems Science: An Overview, Complex Systems Science in Biomedicine, T.S. Deisboeck and J.Y. Kresh, eds. (Springer), pp. 33–114.

24. Wolfram, S. (1988). In "Complex Systems Theory," in Emerging Syntheses in Science: Proceedings of the Founding Workshops of the Santa Fe Institute, Santa Fe, New Mexico, D. Pines, ed. (Addison-Wesley), pp. 183–189.

25. Lorenz, E.N. (1963). Deterministic nonperiodic flow. J. Atmos. Sci. *20*, 130–141, https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

26. Sokol, J. (2019). Hidden Heroines of Chaos, Quanta Magazine. https://www.quantamagazine.org/hidden-heroines-of-chaos-ellen-fetter-and-margaret-hamilton-20190520/.

27. Lorenz, E.N. (1969). The predictability of a flow which possesses many scales of motion *21*, 289–307, https://doi.org/10.1111/j.2153-3490.1969.tb00444.x.

28. Sehl, M.E., Shimada, M., Landeros, A., Lange, K., and Wicha, M.S. (2015). Modeling of cancer stem cell state transitions predicts therapeutic response. PloS one *10*, e0135797, https://doi.org/10.1371/journal.pone.0135797.

29. Batlle, E., and Clevers, H. (2017). Cancer stem cells revisited. Nat. Med. *23*, 1124–1134, https://doi.org/10.1038/nm.4409.

30. Waddington, C.H. (1957). The Strategy of the Genes; a Discussion of Some Aspects of Theoretical Biology (London: Allen and Unwin).

31. Li, C., and Wang, J. (2014). Quantifying the underlying landscape and paths of cancer. J. R. Soc. Interf. *11*, 20140774, https://doi.org/10.1098/rsif.2014.0774.

32. Strogatz, S.H. (2015). Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering (Boulder, CO: Westview Press).

33. Ruelle, D. (1995). Turbulence, strange attractors and chaos (Singapore. . World Scientific Series on Nonlinear Science Series A, *16*.

34. Li, T., and Yorke, J.A. (1975). Period three implies chaos. Amer. Math. Monthly. *82*, 985–992, https://doi.org/10.2307/2318254.

35. Devaney, R. (1989). An Introduction to Chaotic Dynamical Systems (Addison-Wesley Publ. Co.).

36. Alligood, K., Sauer, T., and Yorke, J. (1997). Chaos: An Introduction to Dynamical Systems (Springer-Verlag).

37. Mandelbrot, B. (1982). The Fractal Geometry of Nature (W.H. Freeman and Company).

38. Frederickson, P., Kaplan, J., Yorke, E., and Yorke, J. (1983). The Lyapunov dimension of strange attractors. J. Diff. Eqs. *49*, 185–207, https://doi.org/10.1016/0022-0396(83)90011-6.

39. Feigenbaum, M.J. (1983). Universal behavior in nonlinear systems. Physica D: Nonlinear Phenomena *7*, 16–39, https://doi.org/10.1016/0167-2789(83)90112-4.

40. Bohr, T., Jensen, M.H., Paladin, G., and Vulpiani, A. (1998). Dynamical Systems Approach to Turbulence (Cambridge University Press).

41. Dudkowski, D., Jafari, S., Kapitaniak, T., Kuznetsov, N.V., Leonov, G.A., and Prasad, A. (2016). Hidden attractors in dynamical systems. Phys. Rep. *637*, 1–50, https://doi.org/10.1016/j.physrep.2016.05.002.

42. Balanov, A., Janson, N., Postnov, D., and Sosnovsteva, O. (2009). Synchronization: From Simple to Complex (Springer-Verlag), pp. 192–197.

43. Mengel, B., Hunziker, A., Pedersen, L., Trusina, A., Jensen, M.H., and Krishna, S. (2010). Modeling oscillatory control in NF-κB, p53 and Wnt signaling. Curr. Opin. Genet. Dev. *20*, 656–664, https://doi.org/10.1016/j.gde.2010.08.008.

44. Itik, M., and Banks, S.P. (2010). Chaos in a three-dimensional cancer model. Int. J. Bifurcation Chaos *20*, 71–79, https://doi.org/10.1142/S0218127410025417.

45. Letellier, C., Denis, F., and Aguirre, L.A. (2013). What can be learned from a chaotic cancer model? J. Theor. Biol. *322*, 7–16, https://doi.org/10.1016/j.jtbi.2013.01.003.

46. Zenil, H., Schmidt, A., and Tegnér, J. (2017). Causality, information and biological computation: an algorithmic software approach to life, disease and the immune system. In In: From Matter to Life: Information to Causality, Part III: Chapter 11, S.I. Walker, P.C.W. Davies, and G. Ellis, eds. (Cambridge University Press), pp. 244–280.

47. Khajanchi, S., Perc, M., and Ghosh, D. (2018). The influence of time delay in a chaotic cancer model. Chaos *28*, 103101, https://doi.org/10.1063/1.5052496.

48. Lulla, R.R., Saratsis, A.M., and Hashizume, R. (2016). Mutations in chromatin machinery and pediatric high-grade glioma. Sci. Adv. *2*, e1501354, https://doi.org/10.1126/sciadv.1501354.

49. Bechet, D., Gielen, G.G., Korshunov, A., Pfister, S.M., Rousso, C., Faury, D., Fiset, P.O., Benlimane, N., Lewis, P.W., Lu, C., et al. (2014). Specific detection of methionine 27 mutation in histone 3 variants (H3K27M) in fixed tissue from high-grade astrocytomas. Acta neuropathologica *128*, 733–741, https://doi.org/10.1007/s00401-014-1337-4.

50. Harachi, M., Masui, K., Honda, H., Muragaki, Y., Kawamata, T., Cavenee, W.K., Mischel, P.S., and Shibata, N. (2020). Dual regulation of histone methylation by mTOR complexes controls glioblastoma tumor cell growth via EZH2 and SAM. Mol. Cancer Res. *18*, 1142–1152, https://doi.org/10.1158/1541-7786.MCR-20-0024.

51. Yamanaka, S. (2012). Induced pluripotent stem cells: past, present, and future. Cell Stem Cell *10*, 678–684, https://doi.org/10.1016/j.stem.2012.05.005.

52. Graf, T., and Enver, T. (2009). Forcing cells to change lineages. Nature *462*, 587–594, https://doi.org/10.1038/nature08533.

53. Prager, B.C., Bhargava, S., Mahadev, V., Hubert, C.G., and Rich, J.N. (2020). Glioblastoma stem cells: driving resilience through chaos. Trends Cancer *6*, 223–235, https://doi.org/10.1016/j.trecan.2020.01.009.

54. Suvà, M.L., Rheinbay, E., Gillespie, S.M., Patel, A.P., Wakimoto, H., Rabkin, S.D., Riggi, N., Chi, A.S., Cahill, D.P., Nahed, B.V., et al. (2014). Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. Cell *157*, 580–594, https://doi.org/10.1016/j.cell.2014.02.030.

55. Castellan, M., Guarnieri, A., Fujimura, A., Zanconato, F., Battilana, G., Panciera, T., Sladitschek, H.L., Contessotto, P., Citron, A., Grilli, A., et al. (2020). Single-cell analyses reveal YAP/TAZ as regulators of stemness and cell plasticity in glioblastoma. Nat. Cancer *2*, 174–188, https://doi.org/10.1038/s43018-020-00150-z.

56. Behnan, J., Finocchiaro, G., and Hanna, G. (2019). The landscape of the mesenchymal signature in brain tumours. Brain *142*, 847–866, https://doi.org/10.1093/brain/awz044.

57. Xiong, S., Feng, Y., and Cheng, L. (2019). Cellular reprogramming as a therapeutic target in cancer. Trends Cell Biol *29*, 623–634, https://doi.org/10.1016/j.tcb.2019.05.001.

58. Garofano, L., Migliozzi, S., Take Oh., Y., D'Angelo, F., Najac, R.D., Ko, A., Frangaj, B., Caruso, F.P., Yu, K., Yuan, J., et al. (2021). Pathway-based classification of glioblastoma uncovers a mitochondrial subtype with therapeutic vulnerabilities. Nat. Cancer *2*, 141–156, https://doi.org/10.1038/s43018-020-00159-4.

59. Northcott, P.A., Buchhalter, I., Morrissy, A.S., Hovestadt, V., Weischenfeldt, J., Ehrenberger, T., Gröbner, S., Segura-Wang, M., Zichner, T., Rudneva, V.A., et al. (2017). The whole-genome landscape of medulloblastoma subtypes. Nature *547*, 311–317, https://doi.org/10.1038/nature22973.

60. Vladoiu, M.C., El-Hamamy, I., Donovan, L.K., Farooq, H., Holgado, B.L., Sundaravadanam, Y., Ramaswamy, V., Hendrikse, L.D., Kumar, S., Mack, S.C., et al. (2019). Childhood cerebellar tumours mirror conserved fetal transcriptional programs. Nature *572*, 67–73, https://doi.org/10.1038/s41586-019-1158-7.

61. Bang, J.S., Choi, N.Y., Lee, M., Ko, K., Park, Y.S., and Ko, K. (2019). Reprogramming of cancer cells into induced pluripotent stem cells questioned. Int. J. Stem Cell *12*, 430–439, https://doi.org/10.15283/ijsc19067.

62. Carpentieri, A., Cozzoli, E., Scimeca, M., Bonanno, E., Sardanelli, A.M., and Gambacurta, A. (2015). Differentiation of human neuroblastoma cells toward the osteogenic lineage by mTOR inhibitor. Cell Death Dis. *6*, e1974, https://doi.org/10.1038/cddis.2015.244.

63. Khoshchehreh, R., Totonchi, M., Ramirez, J.C., Torres, R., Baharvand, H., Aicher, A., Ebrahimi, M., and Heeschen, C. (2019). Epigenetic reprogramming of primary pancreatic cancer cells counteracts their in vivo tumourigenicity. Oncogene *38*, 6226–6239, https://doi.org/10.1038/s41388-019-0871-x.

64. Kong, Y., Gimple, R.C., McVicar, R.N., Hodges, A.P., Yin, J., Liu, Y., Zhan, W., and Snyder, E.Y. (2020). Reprogram enablement" as an assay for identifying early oncogenic pathways by their ability to allow neoplastic cells to reacquire an epiblast state. Stem Cel. Rep. *15*, 761–775, https://doi.org/10.1016/j.stemcr.2020.07.016.

65. Uriu, K., and Morelli, L.G. (2014). Collective cell movement promotes synchronization of coupled genetic oscillators. Biophysical J. *107*, 514–526, https://doi.org/10.1016/j.bpj.2014.06.011.

66. Furusawa, C., and Kaneko, K. (2012). A dynamical-systems view of stem cell biology. Science *338*, 215–217, https://doi.org/10.1126/science.1224311.

67. Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C., and Lander, E.S. (2011). Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell *146*, 633–644, https://doi.org/10.1016/j.cell.2011.07.026.

68. Moris, N., Pina, C., and Arias, A.M. (2016). Transition states and cell fate decisions in epigenetic landscapes. Nat. Rev. Genet. *17*, 693–703, https://doi.org/10.1038/nrg.2016.98.

69. Holehouse, J., Cao, Z., and Grima, R. (2020). Stochastic modeling of autoregulatory genetic feedback loops: a review and comparative study. Biophys. J. *118*, 1517–1525, https://doi.org/10.1016/j.bpj.2020.02.016.

70. Shahrezaei, V., and Swain, P.S. (2008). Analytical distributions for stochastic gene expression. PNAS *105*, 17256–17261, https://doi.org/10.1073/pnas.0803850105.

71. Wang, J., Huang, M., Torre, E., Dueck, H., Shaffer, S., Murray, J., Raj, A., Li, M., and Zhang, N.R. (2018). Gene expression distribution deconvolution in single-cell RNA sequencing. PNAS *115*, E6437–E6446.

72. Stumpf, P.S., Smith, R., Lenz, M., Schuppert, A., Müller, F.J., Babtie, A., Chan, T.E., Stumpf, M., Please, C.P., Howison, S.D., et al. (2017). Stem cell differentiation as a non-Markov stochastic process. Cell Syst. *5*, 268–282.e7, https://doi.org/10.1016/j.cels.2017.08.009.

73. Cao, Z., and Grima, R. (2020). Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells. PNAS *117*, 4682–4692, https://doi.org/10.1073/pnas.1910888117.

74. Cao, Z., and Grima, R. (2018). Linear mapping approximation of gene regulatory networks with stochastic dynamics. Nat. Commun. *9*, 3305, https://doi.org/10.1038/s41467-018-05822-0.

75. Wang, J., Zhang, K., Xu, L., and Wang, E. (2011). Quantifying the Waddington landscape and biological paths for development and differentiation. Proc. Natl. Acad. Sci. U S A. *108*, 8257–8262, https://doi.org/10.1073/pnas.1017017108.

76. Wang, J. (2015). Landscape and flux theory of non-equilibrium dynamical systems with application to biology. Adv. Phys. *64*, 1–137, https://doi.org/10.1080/00018732.2015.1037068.

77. Rockne, R.C., Branciamore, S., Qi, J., Frankhouser, D.E., O'Meally, D., Hua, W.K., Cook, G., Carnahan, E., Zhang, L., Marom, A., et al. (2020). State-transition analysis of time-sequential gene expression identifies critical points that predict development of acute myeloid leukemia. Cancer Res. *80*, 3157–3169, https://doi.org/10.1158/0008-5472.CAN-20-0354.

78. Kauffman, S. (1969). Homeostasis and differentiation in random genetic control networks. Nature *224*, 177–178, https://doi.org/10.1038/224177a0.

79. Schwab, J.D., Kühlwein, S.D., Ikonomi, N., Kühl, M., and Kestler, H.A. (2020). Concepts in Boolean network modeling: what do they all mean? Comput. Struct. Biotechnol. J. *18*, 571–582, https://doi.org/10.1016/j.csbj.2020.03.001.

80. Kauffman, S.A. (1993). The Origins of Order (Oxford University Press).

81. Macarthur, B.D., Ma'ayan, A., and Lemischka, I.R. (2009). Systems biology of stem cell fate and cellular reprogramming. Nat. Rev. Mol. Cel. Biol. *10*, 672–681, https://doi.org/10.1038/nrm2766.

82. Barabási, A.-L., and Oltvai, Z.N. (2014). Network biology: understanding the cell's functional organization. Nat. Rev. Genet. *5*, 101–113, https://doi.org/10.1038/nrg1272.

83. Newman, M.E.J. (2003). The structure and function of complex networks. SIAM Rev. *45*, 167–256, https://doi.org/10.1137/S003614450342480.

84. Huang, S., Ernberg, I., and Kauffman, S. (2009). Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. Semin. Cell Dev.l Biol. *20*, 869–876, https://doi.org/10.1016/j.semcdb.2009.07.003.

85. Huang, S., and Kauffman, S. (2013). How to escape the cancer attractor: rationale and limitations of multi-target drugs. Semin. Cancer Biol. *23*, 270–278, https://doi.org/10.1016/j.semcancer.2013.06.003.

86. Huang, S. (2006). Multistability and multicellularity: cell fates as high-dimensional attractors of gene regulatory networks. In Computational Systems Biology (Academic Press), pp. 293–326, https://doi.org/10.1016/B978-012088786-6/50033-2.

87. Kim, Y., Choi, S., Shin, D., and Cho, K.H. (2017). Quantitative evaluation and reversion analysis of the attractor landscapes of an intracellular regulatory network for colorectal cancer. BMC Syst. Biol. *11*, 45, https://doi.org/10.1186/s12918-017-0424-2.

88. Huang, S., Eichler, G., Bar-Yam, Y., and Ingber, D.E. (2005). Cell fates as high-dimensional attractor states of a complex gene regulatory network. Phys. Rev. Lett. *94*, 128701, https://doi.org/10.1103/PhysRevLett.94.128701.

89. Huang, S., and Ingber, D.E. (2006). A non-genetic basis for cancer progression and metastasis: self-organizing attractors in cell regulatory networks. Breast Dis. *26*, 27–54, https://doi.org/10.3233/bd-2007-26104.

90. Turing, A.M. (1952). The chemical basis of morphogenesis. Philos. Trans. R. Soc. B: Biol. Sci. *237*, 37–72, https://doi.org/10.1098/rstb.1952.0012.

91. Serna, H., Munuzuri, A.P., and Barragan, D. (2017). Thermodynamic and morphological characterization of Turing patterns in non-isothermal reaction-diffusion systems. Phys. Chem. Chem. Phys. *19*, 14401–14411, https://doi.org/10.1039/C7CP00543A.

92. Cooke, J., and Zeeman, E.C. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. J. Theor. Biol. *58*, 455–476, https://doi.org/10.1016/s0022-5193(76)80131-2.

93. Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. J. Theor. Biol. *25*, 1–47, https://doi.org/10.1016/s0022-5193(69)80016-0.

94. Zheng, Q., Shen, J., and Wang, Z. (2018). Pattern dynamics of the reaction-diffusion immune system. PLoS ONE *13*, e0190176, https://doi.org/10.1371/journal.pone.0190176.

95. Pham, K., Chauviere, A., Hatzikirou, H., Li, X., Byrne, H.M., Cristini, V., and Lowengrub, J. (2012). Density-dependent quiescence in glioma invasion: instability in a simple reaction-diffusion model for the migration/proliferation dichotomy. J. Biol. Dyn. *6*, 54–71, https://doi.org/10.1080/17513758.2011.590610.

96. Gatenby, R.A., and Gawlinki, A. (1996). A reaction-diffusion model of cancer invasion. Cancer Res. *56*, 5745–5753.

97. Khain, E., and Sander, L.M. (2006). Dynamics and pattern formation in invasive tumor growth. Phys. Rev. Lett. *96*, 188103, https://doi.org/10.1103/PhysRevLett.96.188103.

98. Khajanchi, S. (2019). Chaotic dynamics of a delayed tumor-immune interaction model. Int. J. Biomathematics. *13*, 2050009, https://doi.org/10.1142/S1793524520500096.

99. Baish, J.W., and Jain, R.K. (1998). Cancer, angiogenesis and fractals. Nat. Med. *4*, 984, https://doi.org/10.1038/1952.

100. Ivancevic, T.T., Bottema, M.J., and Jain, L.C. (2008). A theoretical model of chaotic attractor in tumor growth and metastasis. arXiv.

101. Ivancevic, T.T. (2011). Chaotic Attractor in Tumor Growth. Wolfram Demonstrations Project. http://demonstrations.wolfram.com/ChaoticAttractorInTumorGrowth/.

102. Vano, J.A., Wildenberg, J.C., Anderson, M.B., Noel, J.K., and Sprott, J.C. (2006). Chaos in low-dimensional Lotka-Volterra models of competition. Nonlinearity *19*, 2391–2404, https://doi.org/10.1088/0951-7715/19/10/006.

103. Abernethy, S., and Gooding, R.J. (2018). The importance of chaotic attractors in modelling tumour growth. Phys. A *507*, 268–277, https://doi.org/10.1016/j.physa.2018.05.093.

104. Bi, P., Ruan, S., and Zhang, X. (2014). Periodic and chaotic oscillations in a tumor and immune system interaction model with three delays. Chaos *24*, 023101, https://doi.org/10.1063/1.4870363.

105. Voitikova, M.V. (1997). Strange attractors in inmunology of tumor growth. arXiv.

106. Prigogine, I., and Stengers, I. (1984). Order Out of Chaos: Man's New Dialogue with Nature (Bantam Books).

107. Denk, J., Kretschmer, S., Halatek, J., Hartl, C., Schwille, P., and Frey, E. (2018). MinE conformational switching confers robustness on self-organized Min protein patterns. PNAS *115*, 4553–4558, https://doi.org/10.1073/pnas.1719801115.

108. Glock, P., Brauns, F., Halatek, J., Frey, E., and Schwille, P. (2019). Design of biochemical pattern forming systems from minimal motifs. eLife *8*, e48646, https://doi.org/10.7554/eLife.48646.

109. Winfree, A.T. (1967). Biological rhythms and the behavior of populations of coupled oscillators. J. Theor. Biol. *16*, 15–42, https://doi.org/10.1016/0022-5193(67)90051-3.

110. Kuramoto, Y. (2003). Chemical Oscillations, Waves, and Turbulence, 1st edition, Dover., ed. (Springer-Verlag).

111. Winfree, A.T. (1986). The Timing of Biological Clocks (Scientific American Library).

112. Cross, M., and Greenside, H. (2009). Pattern Formation and Dynamics in Nonequilibrium Systems (Cambridge University Press).

113. Wang, R., and Chen, L. (2005). Synchronizing genetic oscillators by signaling molecules. J. Biol. Rhythms *20*, 257–269, https://doi.org/10.1177/0748730405275653.

114. Uriu, K. (2016). Genetic oscillators in development. Dev. Growth Differ *58*, 16–30, https://doi.org/10.1111/dgd.12262.

115. Otto, T., and Sicinski, P. (2017). Cell cycle proteins as promising targets in cancer therapy. Nat. Rev. Cancer *17*, 93–115, https://doi.org/10.1038/nrc.2016.138.

116. Klevecz, R.R. (1992). A precise circadian clock from chaotic cell cycle oscillations. In Ultradian Rhythms in Life Processes, D. Lloyd and E.L. Rossi, eds. (Springer), pp. 41–69.

117. Lloyd, D. (1997). Chaos and ultradian rhythms. Biol. Rhythm Res. *28*, 134–143, https://doi.org/10.1076/brhm.28.1.134.12982.

118. Heltberg, M.L., et al. (2019). Inferring leading interactions in the p53/Mdm2/Mdmx circuit through live-cell imaging and modeling. Cell Syst *9*, 548–558.e5, https://doi.org/10.1016/j.cels.2019.10.010.

119. Izquierdo-Kulich, E., and Nieto-Villar, J.M. (2013). Morphogenesis and complexity of the tumor patterns. In Without Bounds: A Scientific Canvas of Nonlinearity and Complex Dynamics. Understanding Complex Systems, R. Rubio, et al., eds. (Springer), pp. 657–691.

120. Izquierdo-Kulich, E., Allonso-Becerra, E., and Nieto-Villar, J.M. (2011). Entropy production rate for avascular tumor growth. J. Mod. Phys. *2*, 615–620, https://doi.org/10.4236/jmp.2011.226071.

121. Plaks, V., Kong, N., and Werb, Z. (2015). The cancer stem cell niche: how essential is the niche in regulating stemness of tumor cells? Cell Stem Cell *16*, 225–238, https://doi.org/10.1016/j.stem.2015.02.015.

122. Amemiya, T., Shibata, K., Itoh, Y., Itoh, K., Watanabe, M., and Yamaguchi, T. (2017). Primordial oscillations in life: direct observation of glycolytic oscillations in individual HeLa cervical cancer cells. Chaos *27*, 104602, https://doi.org/10.1063/1.4986865.

123. Janson, N. (2012). Non-linear dynamics of Biological systems. Contemp. Phys. *53*, 137–168, https://doi.org/10.1080/00107514.2011.644441.

124. Taniguchi, K., and Karin, M. (2018). NF-κB, inflammation, immunity and cancer: coming of age. Nat. Rev. Immunol. *18*, 309–324, https://doi.org/10.1038/nri.2017.142.

125. Jensen, M.H., and Krishna, S. (2012). Inducing phase-locking and chaos in cellular oscillations by modulating the driving stimuli. FEBS Lett. *586*, 1664–1668, https://doi.org/10.1016/j.febslet.2012.04.044.

126. Heltberg, M., Kellogg, R.A., Krishna, S., Tay, S., and Jensen, M.H. (2016). Noise induces hopping between NF-κB entrainment modes. Cell Syst. *3*, 532–539.e3, https://doi.org/10.1016/j.cels.2016.11.014.

127. Gluzman, M., Scott, J.G., and Vladimirsky, A. (2020). Optimizing adaptive cancer therapy: dynamic programming and evolutionary game theory. Proc. R. Soc. B *287*, 20192454, https://doi.org/10.1098/rspb.2019.2454.

128. Barabási, A.-L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. Nat. Rev. Genet. *12*, 56–68, https://doi.org/10.1038/nrg2918.

129. Barabási, A.-L., and Posfai, M. (2016). Network Science (Cambridge University Press).

130. Iacono, G., Massoni-Badosa, R., and Heyn, H. (2019). Single-cell transcriptomics unveils gene regulatory network plasticity. Genome Biol. *20*, 110, https://doi.org/10.1186/s13059-019-1713-4.

131. Harush, U., and Barzel, B. (2017). Dynamic patterns of information flow in complex networks. Nat. Commun. *8*, 2181, https://doi.org/10.1038/s41467-017-01916-3.

132. Galvão, V., and Miranda, J.G.V. (2008). A computational model for cancer growth by using complex networks. Phys. A Stat. Mech. Its Appl. *387*, 5279–5286, https://doi.org/10.1016/j.physa.2008.05.010.

133. Choo, S.-M., and Cho, K.-H. (2016). An efficient algorithm for identifying primary phenotype attractors of a large-scale Boolean network. BMC Syst. Biol. *10*, 95, https://doi.org/10.1186/s12918-016-0338-4.

134. Tang, F., Wang, C., Su, J., and Wang, Y. (2019). Spectral clustering-based community detection using graph distance and node attributes. Comput. Stat. *35*, 69–94, https://doi.org/10.1007/s00180-019-00909-8.

135. Liu, E., Zhang, Z.Z., Cheng, X., Liu, X., and Cheng, L. (2020). SCNrank: spectral clustering for network-based ranking to reveal potential drug targets and its application in pancreatic ductal adenocarcinoma. BMC Med. Genomics *13*, 50, https://doi.org/10.1186/s12920-020-0681-6.

136. Choi, Y.H., and Kim, J.K. (2019). Dissecting cellular heterogeneity using single-cell RNA sequencing. Mol. Cells *42*, 189–199, https://doi.org/10.14348/molcells.2019.2446.

137. Brandes, U., Delling, D., Gaertler, M., Goerke, R., Hoefer, M., Nikoloski, Z., and Wagner, D. (2006). Maximizing modularity is hard. arXiv.

138. Arenas, A., Duch, J., Fernandez, A., and Gomez, S. (2007). Size reduction of complex networks preserving modularity. New J. Phys. *9*, 176, https://doi.org/10.1088/1367-2630/9/6/176.

139. Newman, M.E.J., and Girvan, M. (2004). Finding and evaluating community structure in networks. Phys. Rev. E. *69*, 026113, https://doi.org/10.1103/PhysRevE.69.026113.

140. Son, S.-W., Jeong, H., and Noh, J.D. (2006). Random field Ising model and community structure in complex networks. Eur. Phys. J. B. *50*, 431–437, https://doi.org/10.1140/epjb/e2006-00155-4.

141. Blondel, V.D., Guillaume, J.-P., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. J. Stat. Mech. P10008, https://doi.org/10.1088/1742-5468/2008/10/P10008.

142. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21, https://doi.org/10.1016/j.cell.2019.05.031.

143. Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC genomics *19*, 477, https://doi.org/10.1186/s12864-018-4772-0.

144. Brouwer, I., and Lenstra, T.L. (2019). Visualizing transcription: key to understanding gene expression dynamics. Curr. Opin. Chem. Biol. *51*, 122–129, https://doi.org/10.1016/j.cbpa.2019.05.031.

145. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498, https://doi.org/10.1038/s41586-018-0414-6.

146. Cao, J., Zhou, W., Steemers, F., Trapnell, C., and Shendure, J. (2020). Sci-fate characterizes the dynamics of gene expression in single cells. Nat. Biotechnol. *38*, 980–988, https://doi.org/10.1038/s41587-020-0480-9.

147. Sauer, T.D. (2006). Attractor reconstruction. Scholarpedia *1*, 1727, https://doi.org/10.4249/scholarpedia.1727.

148. Sugihara, G., May, R., Ye, H., Hsieh, C.H., Deyle, E., Fogarty, M., and Munch, S. (2012). Detecting causality in complex ecosystems. Science *338*, 496–500, https://doi.org/10.1126/science.1227079.

149. Krieger, M.S., Moreau, J.M., Zhang, H., Chien, M., Zehnder, J.L., and Craig, M. (2020). A blueprint for identifying phenotypes and drug targets in complex disorders with empirical dynamics. Patterns *1*, 100138, https://doi.org/10.1016/j.patter.2020.100138.

150. Toker, D., Sommer, F.T., and D'Esposito, M. (2020). A simple method for detecting chaos in nature. Commun. Biol. *3*, 11, https://doi.org/10.1038/s42003-019-0715-9.

151. So, P. (2007). Unstable periodic orbits. Scholarpedia *2*, 1353, https://doi.org/10.4249/scholarpedia.1353.

152. Boccaletti, S., Grebogi, C., Lai, Y.-C., Mancini, H., and Maza, D. (2000). The control of chaos: theory and applications. Phys. Rep. *329*, 103–197, https://doi.org/10.1016/S0370-1573(99)00096-4.

153. Wiener, N. (1961). Cybernetics: Or Control and Communication in the Animal and the Machine, 2nd version (MIT Press).

154. Ott, E., Grebogi, C., and Yorke, J. (1990). Controlling chaos. Phys. Rev. Lett. *64*, 1196–1199, https://doi.org/10.1103/PhysRevLett.64.1196.

155. Ditto, W.L. (1996). Applications of chaos in biology and medicine. AIP Conf. Proc. *376*, 175, https://doi.org/10.1063/1.51060.

156. Pierson, D., and Moss, F. (1995). Detecting periodic unstable points in noisy chaotic and limit cycle attractors with applications to biology. Phys. Rev. Lett. *75*, 2124–2127, https://doi.org/10.1103/PhysRevLett.75.2124.

157. Matsumoto, K., and Tsyda, I. (1983). Noise-induced order. J. Stat. Phys. *31*, 87–106, https://doi.org/10.1007/BF01010923.

158. Jin, S., MacLean, A.L., Peng, T., and Nie, Q. (2018). scEpath: energy landscape-based inference of transition probabilities and cellular trajectories from single-cell transcriptomic data. Bioinformatics *34*, 2077–2086, https://doi.org/10.1093/bioinformatics/bty058.

159. Shi, J., Teschendorff, A.E., Chen, W., Chen, L., and Li, T. (2020). Quantifying Waddington's epigenetic landscape: a comparison of single-cell potency measures. Brief. Bioinform. *21*, 248–261, https://doi.org/10.1093/bib/bby093.

160. Teschendorff, A.E., Maity, A.K., Hu, X., Weiyan, C., and Lechner, M. (2020). Ultra-fast scalable estimation of single-cell differentiation potency from scRNA-seq data. Bioinformatics, btaa987, https://doi.org/10.1093/bioinformatics/btaa987.

161. Hopfield, J.J., and Tank, D.W. (1985). Neural computation of decisions in optimization problems. Biol. Cybernetics *52*, 141–152, https://doi.org/10.1007/BF00339943.

162. Bishop, C.M. (1995). Neural Networks for Pattern Recognition (Oxford University Press).

163. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature *521*, 436–444, https://doi.org/10.1038/nature14539.

164. Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep Learning (MIT Press).

165. Feurer, M., and Hutter, F. (2019). Hyperparameter optimization. In Automated Machine Learning. The Springer Series on Challenges in Machine Learning, F. Hutter, L. Kotthoff, and J. Vanschoren, eds. (Springer). https://doi.org/10.1007/978-3-030-05318-5_1.

166. Tripto, N.I., Kabir, M., Bayzid, M.S., and Rahman, A. (2020). Evaluation of classification and forecasting methods on time series gene expression data. PLoS One *15*, e0241686, https://doi.org/10.1371/journal.pone.0241686.

167. Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature *529*, 484–489, https://doi.org/10.1038/nature16961.

168. Yuan, Y., and Bar-Joseph, Z. (2019). Deep learning for inferring gene relationships from single-cell expression data. Proc. Natl. Acad. Sci. U S A *116*, 27151–27158, https://doi.org/10.1073/pnas.1911536116.

169. Tasaki, S., Gaiteri, C., Mostafavi, S., and Wang, Y. (2020). Deep learning decodes the principles of differential gene expression. Nat. Machine intelligence *2*, 376–386, https://doi.org/10.1038/s42256-020-0201-6.

170. Singh, R., Lanchantin, J., Sekhon, A., and Qi, Y. (2017). Attend and predict: understanding gene regulation by selective attention on chromatin. Adv. Neural Inf. Process. Syst. *30*, 6785–6795.

171. Chen, Y., Li, Y., Narayan, R., Subramanian, A., and Xie, X. (2016). Gene expression inference with deep learning. Bioinformatics *32*, 1832–1839, https://doi.org/10.1093/bioinformatics/btw074.

172. Schmauch, B., Romagnoni, A., Pronier, E., Saillard, C., Maillé, P., Calderaro, J., Kamoun, A., Sefta, M., Toldo, S., Zaslavskiy, M., et al. (2020). A deep learning model to predict RNA-seq expression of tumours from whole slide images. Nat. Commun. *11*, 3877, https://doi.org/10.1038/s41467-020-17678-4.

173. Wang, L., Nie, R., Yu, Z., Xin, R., Zheng, C., Zhang, Z., Zhang, J., and Cai, J. (2020). An interpretable deep-learning architecture of capsule networks for identifying cell-type gene expression programs from single-cell RNA-sequencing data. Nat. Machine Intelligence. *2*, 693–703, https://doi.org/10.1038/s42256-020-00244-4.

174. Pathak, J., Lu, Z., Hunt, B.R., Girvan, M., and Ott, E. (2017). Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data. Chaos *27*, 121102, https://doi.org/10.1063/1.5010300.

175. Pathak, J., Hunt, B., Girvan, M., Lu, Z., and Ott, E. (2018). Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach. Phys. Rev. Lett. *120*, 024102, https://doi.org/10.1103/PhysRevLett.120.024102.

176. Vlachas, P.R., Byeon, W., Wan, Z.Y., Sapsis, T.P., and Koumoutsakos, P. (2018). Data-driven forecasting of high-dimensional chaotic systems with long short-term memory networks. Proc. R. Soc. A: Math. Phys. Eng. Sci. *474*, 2213, https://doi.org/10.1098/rspa.2017.0844.

177. Vlachas, P.R., Pathak, J., Hunt, B.R., Sapsis, T.P., Girvan, M., Ott, E., and Koumoutsakos, P. (2020). Backpropagation algorithms and reservoir computing in recurrent neural networks for the forecasting of complex spatiotemporal dynamics. Neural Networks *126*, 191–217, https://doi.org/10.1016/j.neunet.2020.02.016.

178. Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J.C., Barends, R., Biswas, R., Boixo, S., Brandao, F., Buell, D.A., et al. (2019). Quantum supremacy using a programmable superconducting processor. Nature *574*, 505–510, https://doi.org/10.1038/s41586-019-1666-5.

179. Bravyi, S., Gosset, D., Konig, R., and Tomamichel, M. (2020). Quantum advantage with noisy shallow circuits. Nat. Phys. *16*, 1040–1045, https://doi.org/10.1038/s41567-020-0948-z.

180. Sipser, M. (1997). Introduction to the Theory of Computation (PWS Publishing Co.).

181. Kolmogorov, A.N. (1968). Three approaches to the quantitative definition of information. Int. J. Comp. Math. *2*, 157–168, https://doi.org/10.1080/00207166808803030.

182. Zenil, H., Kiani, N.A., Marabita, F., Deng, Y., Elias, S., Schmidt, A., Ball, G., and Tegnér, J. (2019). An algorithmic information calculus for causal discovery and reprogramming systems. iScience *19*, 1160–1172, https://doi.org/10.1016/j.isci.2019.07.043.

183. Solomonoff, R.J. (1997). The discovery of algorithmic probability. J. Comput. Syst. Sci. *55*, 73–88, https://doi.org/10.1006/jcss.1997.1500.

184. Zenil, H., Kiani, N.A., Abrahao, F.S., and Tegner, J.N. (2020). Algorithmic information dynamics. Scholarpedia *15*, 53143, https://doi.org/10.4249/scholarpedia.53143.

185. Zenil, H., Kiani, N.A., and Tegnér, J. (2018). A review of graph and network complexity from an algorithmic information perspective. Entropy *20*, 551, https://doi.org/10.3390/e20080551.

186. Morris, R. (2011). Lorenz Attractor. Wolfram Demonstrations Project http://demonstrations.wolfram.com/LorenzAttractor/.