Cancer: Strange Attractors and Complexity

Abicumaran Uthamacumaran
Concordia University
a_utham@live.concordia.ca

Cancer is a complex adaptive ecosystem and remains the lead cause of disease-related death in pediatric patients in North America. Machine learning, Network science, Fluid dynamics and Quantum Mechanics are hereby discussed as tools to advance oncology and cancer reprogramming. The fluid dynamics of cell-fate transitions, cancer pattern formation and invasion are reviewed in this paper. Cancer cell decision-making is investigated through dynamical systems and complexity theory. A fluid-dynamics grid-scheme based Deep Learning neural networks is proposed as the solution to identify strange attractors in time-series multi-omics (scRNA-Seq) data and time lapse imaging of cancer stem cell differentiation. Cancer is discussed within the context of two unsolved foundational problems in science: the three-dimensional Navier-Stokes equations global regularity, smoothness and existence and the P vs. NP problem.
PROPOSITION: A fluid-dynamics grid-scheme based Deep Learning neural networks is proposed as the solution to identify strange attractors in time-series scRNA-Seq data and time lapse imaging of cancer stem cell differentiation.

PEDIATRICS: THE FUTURE

Cancer is the leading cause of ‘disease’-related death among children in North America. According to the World Health Organization data, in 2015, near 397,000 children under 15 developed cancer globally and ~45% of those cases went undiagnosed (Oakes, 2019). Pediatric cancers are herein briefly discussed to describe a tumor and the multiple phenotypes of its cells as a dynamical system having multiple, chaotic attractors.

Leukemias, cancers of the bone marrow and blood account for ~30% of all pediatric cancers. Leukemia is myeloid or lymphocytic depending on which bone marrow cells the cancer originates. A leukemia is acute or chronic depending on whether most of the abnormal cells are immature (stem cell-like) or mature, respectively. Chronic myelogenous leukemia and Hodgkin’s disease are now virtually curable. Acute lymphoblastic leukemia (ALL) is the most prevalent pediatric cancer. The uncontrolled proliferation of progenitor lymphocytes occurs usually by translocations/fusions. The 5-year survival rate for children has greatly increased over time, presently >85%. However, increased survival does not imply improved living conditions. Among children, B-cell lineage ALL constitutes approximately 88% of cases while T-cell lineage represents the remainder. On the other hand, multi-omics screening show epimutations shape CLL (chronic lymphocytic leukemia) lineages and can aid in clinical interventions (Gaiti et al., 2019).

Many relapse ALL has NT5C3 and CREBBP (CREB-binding protein) gene mutations which are associated to chemoresistance. Poor prognosis for relapse (40%) occurs due to clonal heterogeneity. >20% of patients show ETV6-RUNX1 mutation (best survival) and is the most common fusion gene in pediatric ALL while MLL-gene rearrangements occur in 6% of patients (i.e. under 1 year of age). Meanwhile, BCR-Abl1 and like subtypes are difficult to treat (Boer and Boer, 2017). Children whose leukemia cells have a translocation between chromosomes 12 and 21 are more likely to be cured with current treatments. Whereas, those with a translocation between chromosomes 9 and 22 (the Philadelphia chromosome) or 4 and 11 tend to have a less favorable prognosis.

Different subtypes are observed in young childhood ALL (mostly ETV6-RUNX1 and hyperdiploid mutants) whereas the frequencies are shifted in adolescent and young adults (from ~60 to 20%). Single-cell CRISPR screens can identify many drug targets in these subtypes (Bhojwani et al., 2012). The whole genome sequencing of pediatric Wilms tumor, acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) identified sub-type specific clusters of driver events in many studies. There are fusion events and focal areas of gene deletion such as FLT3-ITD (FMS-like tyrosine kinase 3 -internal tandem duplication) with (wild-type 1) WT1 mutation or NUP98-NSD1 fusion with poor outcomes for AML. For instance, the FLT3-D835 mutants may confer chemoresistance to type II FLT3 inhibitors, while the least resistance was shown in Crenolanib treatment amidst relapse patients (Smith et al., 2015). Children with leukemia cells that have translocations between chromosomes 15 and 17 (seen in most cases of Acute promyelocytic leukemia) or between 8 and 21, or with an inversion (rearrangement) of chromosome 16 have higher probability of being cured. Children lacking a copy of chromosome 5 or 7 (monosomy) or chromosome 5 deletion have poorer prognosis. Dysregulation of T-cell mediated allorecognition is a driver of AML relapses post- hematopoietic cell transplantation (Health Statistics).
The evolution of intratumor heterogeneity consists of the interactions between hybrid (mixed) phenotypes, dormant states (quiescence and senescence) and the ability of one subtype to adaptively interconvert into another in relapse patients (i.e. cell fate reprogramming), making cancer an intractable problem. Even within a single patient, many adaptive subclones are found.

Central nervous system cancers account for more than one-third of cancer death in children. Pilocytic astrocytoma and medulloblastoma are amidst the most common primary brain tumors in children between 5 -14 years of age and are the most common spinal cord tumors in the pediatric population. Although the 5-year survival rates are increasing, the quality of life is poor and children suffer from devastating side effects including secondary tumors/relapse in future, growth inhibition, disabilities, heart failure, kidney disease, infertility, etc. (Collins et al., 2015). With aggressive forms of brain cancer, barely 10 percent of children survive three years after diagnosis. Another deadly pediatric-killer is Glioblastoma Multiforme (GBM), malignant high-grade gliomas (WHO grade IV astrocytoma) that are resistant to conventional therapies.

Exome sequencing of adult GBM identified 4 types of primary tumors: proneural, neural, classical and mesenchymal types. Aberrant gene expression of EGFR, NF1 and PDGFRA/IDH1 each define the Classical, Mesenchymal, and Proneural subtypes, respectively (Verhaak et al., 2010). The Proneural groups in the Verhaak classification consists of IDH (isocitrate dehydrogenase) mutants and can be low or high grade. The IDH1 mutant cluster of adult GBM favors mitochondrial oxidative phosphorylation. A more useful classification for gliomas in general, is the 2016 WHO classification: IDH mutant vs. IDH wild-type gliomas (Ceccarelli et al., 2016).

Response to aggressive therapy differs by subtype and patient, with the greatest relative benefit in the Classical subtype and no benefit in the Proneural subtype. However, all subgroups respond poorly to current treatments in neuro-oncology. However, the genomic landscape of cancer in pediatrics is very different from that of adults (Sweet-Cordero and Biegel, 2019). The lower genetic mutational burden and higher epigenetic alterations in pediatric cancer has been postulated to be due to a combination of the embryonic origin of cancer, dysregulation of developmental pathways (morphogens) and environmental carcinogens.

Mutations in IDH1 and less frequently in IDH2, occur in 80% of grade II and grade III astrocytomas and oligodendrogliomas. They are also found in high grade glioblastomas that have arisen over time from these lower grade gliomas (Staedtke et al., 2016; Azzarelli et al. 2018). In adults, the lack of IDH mutations, 1p19q co-deletion and/or presence of BRAF alterations are very helpful in sorting out the differential diagnosis. However, this does not apply in children, in whom such mutations are virtually absent. Epigenetic dysregulations are more useful in classifying pediatric brain cancers. IDH1 mutants is only about 6% of the cases in pediatric GBM (mostly in adolescents, early adult-transition) (Ceccarelli et al., 2016). Thus, features of pediatric and adult cancers are very distinct.

IDH mutations disrupt cellular metabolism causing the hypermethylation of histones and CpG islands (methylator phenotype). Interestingly, progression to higher grade disease is often accompanied by overall decrease in methylation, and hypermethylation of a small subset of CpG islands associated with developmental regulators, including FOX, SOX and TBX family genes (which may arrest cells into a permanently self-renewing state) (Bai et al., 2016). A common path of glioblastoma tumorigenesis is one or more of chromosome 7 gain, 9p loss, or 10 loss. TERT promoter/telomerase mutations often occurred
later for rapid growth and relapsed tumors indicating a critical attractor in cancer development (Korber et al., 2019). Survival promoting mutations commonly in TERT promoter is needed years pre-diagnosis of chromosomal alterations for detectable tumors. As well-established, the maintenance of telomeres (increased telomerase or ALT-alternative lengthening) is a hallmark of cancer. Recent studies showed 73% of cancers of 31 different types needed mutations in TERT promoter methylation (Barthel et al., 2017; Verhaak et al., 2019).

90% of GBM recurrences occur within 2 cm of original tumor resection site or radiation field. The post-operative residual tumor tend to be more aggressive post-non target specific therapies, especially if there is MGMT promoter methylation (confers chemoresistance) (Del Maestro, 2012). However, surgery remains one of the primary interventions for neuro-oncology. In 126 patients, those with more than 90% of tumour resection had a significantly better outcome, improving median survival from 225 to 519 days (Dea et al., 2012).

Surgical intervention of high-grade gliomas such as pediatric GBM sometimes may result in the rupture of the inner membranes attributing to the spreading of GBM (rare cases). GBM patients usually die within 15 months of diagnosis demonstrating their high aggressiveness, heterogeneity and morbidity without metastatic invasion (Klughammer et al., 2018). The cancer lies very close to the blood-brain barrier and utilize GSC (glial stem cells), stromal cells and ECM (extracellular matrix-niche) through EMT (epithelial-mesenchymal transition) pathways to increase their cancer stemness, tumor propagation and cancer-promoting signal transduction. Moreover, the inflammatory niche of brain cancers varies (e.g. DIPG generally shows a lower inflammatory milieu yet highly aggressive).

As mentioned, epigenetic dysregulation is a hallmark of pediatric cancers, resulting in translocations and fusion oncoproteins. Epimutations are more common in pediatric cancers than in adults (Filbin and Monje, 2019). The best-studied example of this is the histone-3 (H3)-K27M mutation that occurs in diffuse midline gliomas, such as diffuse intrinsic pontine glioma (DIPG) and thalamic and spinal cord gliomas of childhood (Filbin et al., 2018). For e.g., exome sequencing showed 7 hypermethylated samples in 19 patients, with TMZ (temozolomide) treatment+ angiogenesis inhibitor (e.g. bevacizumab) (Cancer Genome Atlas, 2008). The most prominent difference between pediatric and adult cancers is the frequency of hotspot mutations in genes encoding histone H3 variants (0.2% in adults vs 50.3% in DIPG pediatric). Epigenetic alterations often include complexes involved in chromatin remodelling such as SWI/SNF, polycomb repressor complexes and oncohistone mutations in GBM. Pediatric and young adult GBMs show frequent alterations in H3.3-ATRX-DAXX (Salloum et al., 2017). For e.g., diffuse intrapontine gliomas have frequent histone H3F3A mutations. Whereas, CREBBP, SETD2 etc. play key roles in the altered epigenetic hallmarks of ALL. Hence, DNA methylation is highly aberrant in pediatric cancers, especially brain tumors, indicating a target for deep-learning pattern recognition (Det et al., 2018).

In Diffuse intrinsic pontine glioma (DIPG), the enzyme ACVR1 mutation cooperates with histone H3 mutation. H3.1 K27M mutations occur in 20% of the brain tumor developments. The majority of DIPG patients carry a H3.3 Lys27Met (K27M) missense mutation, whereas a minority exhibit a H3.3 G34R/V mutation (Mackay et al., 2017). K27M H3.3 tumours are found in 70% of DIPG and non-brainstem midline paediatric gliomas. H3 mutations are also a major cancer stem cell attractor in pre-leukemic hematopoietic stem cells (Boileau et al., 2019).
Meanwhile, treatment with immunotherapies such as PD-1 inhibitors (e.g. nivolumab or pembrolizumab) revealed a significant enrichment of PTEN mutations associated with immunosuppressive signatures in non-responders, and an enrichment of MAPK pathway alterations (PTPN11, BRAF) in responders (Zhao et al., 2019). However, there is low clinical efficiency of immune checkpoint inhibitors in GBM, although nivolumab-ipilimumab combination is considered for recurrent GBM (relapse). Most importantly, whether the most advanced of our current therapy regimes including immunotherapies target nanoscale dynamics (i.e. cancer exosomes) is undetermined. Rather, the inverse is well-established (i.e. immunotherapy resistance is conferred by exosomes and NETs).

Functional genomics platforms and in vitro genome editing allows for efficient immunotherapies and multi-targeted CAR (Chimeric Antigen Receptor)- T cells. Arrayed CRISPR-Cas9 screens need to address immunomodulators, tumor microenvironiments and in vivo effects of the cancer. CRISPR screens are effective tools to study chemoresistance in patients. CRISPR screens identified stemness regulators in patient-derived glioblastoma stem cells and sensitivity to chemotherapy including members of Sox (Sox9, Socs3, USP8, DOT1L and protein ufmylation), while TMZ-resistance genes were found to be Fanconi Anemia/ICL R, MCM8/MCM9, MLH1, mismatch repair genes, etc. (Macleod et al., 2019).

Conventionally, alkylated TMZ (temozolomide) and radiation are usually administered for newly diagnosed GBM. Therapeutics targeting tumour angiogenesis concurrent with radiation is crucial to improve the overall therapeutic response in GBM (e.g. sunitinib, VEGF-inhibitors, glucocorticoids for edema-control, etc.) (Gerstner et al., 2009; Murayi and Chittiboina, 2016). However, with emerging complexities in cancer evolution, these approaches are not enough. For e.g., as a simple illustration, angiogenesis inhibitors are not effective against vasculogenic mimicry pathways.

Cancer is a systemic disease requiring multimodal therapies. Apart from surgery, radio-oncology is often the route of treatment in aggressive stages of cancers. Radiation therapy releases ROS (reactive oxygen species)/free radicals and inflammatory responses, which not only can mutate healthy cells but can increase the survival of adaptive cancer stem cells. GH (growth hormone) deficiency, neurological dysfunction and primary hypothyroidism are common after cranial irradiation and chemotherapy for nonpituitary-related brain tumors (Clarson and Del Maestro, 1999). Radiation and chemotherapy cause an acute depletion of (healthy) neural stem and progenitor cells. The long-term self-renewal capacity of neural precursor cells (NPCs) is diminished and NPC function is severely impaired by neurogenic - radiation-induced microglial inflammation. Moreover, these inflammatory pathways may increase the selection of more aggressive phenotypes in cancers.

Abnormal scar tissues are generated by radiation therapy along with arterial venous malformations/shunts, arrested growth and sometimes secondary tumors. For instance, HIF-1 (hypoxia-inducible factors) signaling increases in chemo-resistant cells and ALDH1 (aldehyde dehydrogenase) increases radiation resistance. Note: proton therapy is less damaging than x-rays when applicable. It must be investigated whether post-radiation therapy induced oncosomes can horizontally transfer adaptive traits such as radiation-resistance, as has been established for chemoresistance (Liauw et al., 2013). For instance, it is well-established NET (neutrophil extracellular traps) released from post-surgical stress and infections can induce cancer progression by collecting circulating tumor cells (CTC) and inducing the formation of metastases (Cools-Lartigue et al., 2013). In chemo-resistant patients, NETs shield tumor cells against the patient’s cytotoxic T-cells and thereby causes immunotherapy resistance. Recent findings suggest non-target specific chemotherapy causes cancer cells to communicate about
their microenvironmental changes by the release of pro-metastatic exosomes (Keklikogolu et al. (2019). Exosomes in the tumor microenvironment are hence mediators of cancer therapy resistance.

Childhood cancer therapies, such as cranial radiation, methotrexate and cyclophosphamide, induce persistent microglial reactivity which in turn induces neurotoxic astrocyte reactivity, disrupts oligodendroglial lineage dynamics, impairs myelination, blocks neuronal differentiation of hippocampal stem cells (neurogenesis) and contributes to aberrant pruning of synapses. Some of the current complications post-chemo/radiation therapies in neuro-oncology include: vascular-prothrombic neurological side effects, chemotoxicity, autoimmune responses, vasculitis, leukodystrophy, cognitive decline, delayed leukoencephalopathy with metabolic disorders, congenital issues, neuropathy, immunological inflammations, psychiatric symptoms, intravascular lymphomas, strokes, microangiopathy/cerebrovascular diseases, encephalitis, dystonia/dysmnesia, respiratory failures, secondary tumors/metastatic relapse, carcinomatosis, toxic metabolic encephalopathy, neoplasms/teratomas, etc. (Stewart et al., 1991). To overcome these health barriers, new ways of treating cancer must be devised and such is the attempt of cancer (stem cell) reprogramming.

Childhood cerebellar tumors have conserved fetal transcriptional programs where cancer stem cells resemble embryonic stem cells (Vladoiu et al., 2019). Lineage-specific mechanisms that regulate stem cell behaviour in the embryo may be subverted in cancer causing its hyper-proliferation and differentiation. For instance, single-cell transcriptomics-based cell lineage reconstruction of cerebellar tumors revealed mixed, temporal heterogeneity and four distinct targetable clusters in medulloblastoma. The four distinct subgroups were: the sonic hedgehog (Shh) subgroup, the Group 3 (Nestin +), Group 4 and Wnt medulloblastoma groups (Vladoiu et al., 2019; Northcott et al., 2017). Apart from differential gene expression signatures, CpG island methylation profiles were also distinct between the clusters as shown by t-SNE clustering.

Sox2 (a Yamanaka factor) is also an essential driver of cancer stem cell sub-populations in glioblastoma (Suva et al., 2014; 2014). PI3K/mTOR (mammalian target of rapamycin) and MEK/ERK pathways were shown to be critical to the self-renewal in glioma stem cells (GSC) and mediate cancer stemness in brain tumors (Sunayama et al. 2010). However, mutations in putative drug targets become distinct through the divergent, clonal evolution of tumors making their gene expression networks an intractable problem (Salloum et al., 2017). Regardless, key driver mutations such as oncohistone mutations/H3 post-translational modifications, IDH1, and BRAFV600E are tractable in all clusters. For instance, in glioblastoma, induction of GLI1 by Nanog increases the clonogenic and tumorigenic potential of the CD133-positive stem cell fraction (Suva et al., 2013; 2014). The epigenetic landscape of GBMs (Glioblastoma Multiforme) shows tremendous spatiotemporal heterogeneity. However, a core set of neurodevelopmental transcription factors (POU3F2, SOX2, SALL2, OLIG2) were identified to be essential for GBM propagation in brain cancer stem cells (Suva et al., 2014). Thus, many critical attractors of GRN circuitries are being revealed with recent multi-omics mapping of pediatric cancers. Yet what distinguishes a cancer stem cell in a given tumor population’s evolution remains an unsolved problem, requiring the study of fluid dynamics in cancer stem cell transitions and pattern formation equations. Currently, graph-theoretic approaches and machine learning algorithms are employed to map the cell fate trajectories, generally without an understanding of the underlying fluid dynamics.

To summarize, quantitative changes of lineage specific, Transcription Factors mediating cell fate decision-making can be revealed by single-cell RNA-sequencing and proteomics (e.g. MALDI-TOF mass spectrometry, gel electrophoresis, etc.) in early tumor progenitors (Taylor et al., 2019). A better
understanding of the reaction-diffusion of morphogens pertaining to cancer stem cells and their energy landscapes is crucial to approach reprogramming cancer phenotypes. As will be discussed, recent findings suggest simplified Navier-Stokes equations such as Darcy's law are required to better model the reaction-diffusion processes underlying GBM invasion and growth. The tumor flow equations are proposed to help understand the complexity of gene regulatory networks (GRN) and thereby approach cancer stem cell reprogramming. In principle, a Waddington landscape can be reconstructed from any tumor sample from multi-omics data and the fluid flow equations can map the evolution of distinct phenotype clusters and its single-cell fate trajectories.

ENERGY LANDSCAPES

Waddington (1942) first described the cell developmental topography as an energy landscape where cell fate bifurcations are classically visualized as balls rolling up hills (energy barriers) and down valleys (attractors) (Bhattacharya et al., 2011; Wang et al., 2011). The cell fate decision-making is governed by the gene expression dynamics of underlying gene regulatory networks. Due to the stochastic nature of molecules, the changes in molecular concentrations are generally defined by Langevin equations (Davila-Vederrain, 2015). The transitions between the attractors are considered as random walks on a network, following Boltzmann statistics (Perkins et al., 2014). The barrier heights and transition rates (epigenetic fluctuations) are in the order of thermal fluctuations $k_B T$, where $k_B$ is the Boltzmann constant and $T$ is the temperature.

As such, chemical reactions are often modelled by Hills-Langmuir equations or Michaelis-Menten kinetics, such that their rates follow the Arhenius law: $Rate \propto e^{-\Delta G_a/k_B T}$ where $G_a$ is the activation free-energy, and $T$ the temperature. Noise in chemical fluctuations are often considered as a Gaussian term. However, this only holds for simple chemical reactions, whereas biological reactions involve complex, nonlinear cascades with inhomogeneous energy potentials. Current models assume steady-state equilibrium for the reaction coordinates (Zhou et al., 2014). The dynamics of many biosystems are well described as random walks on a network, with three general possible scenarios for the path distribution: finite, stretched-exponential and power law (Perkins et al. 2014). The latter are used to describe the dynamics of the Lorenz attractor and G-protein folding. The Lorenz attractor is a strange attractor which gives the approximate description of a horizontal fluid layer heated from below. This is known as the Rayleigh-Benard convection, a simple testable model of experimental fluid turbulence and used in weather forecasts (Lorenz, 1963).

Turbulent/chaotic chemical oscillations are neglected in Waddington landscapes since low Reynolds number (laminar) flows are assumed within cells. However, turbulence can occur in very rapid bursts followed by laminar-like flows (intermittency). In complexity theory, symmetry-breaking in fluid transitions are NP-complete problems to which solutions can only be verified by brute-force searching (i.e. stochastic methods) or numerical approximations. Gene expression is often naively considered a stochastic process (Elowitz, 2002). If stochastic is to be taken as a synonym to probabilistic causality, then it may be fine. However, as will be discussed, recent evidences challenge these dogmas by demonstrating the paradoxical roles of chemical turbulence in the emergence of ordered structures in cellular pattern formation and protein folding.

An example of stochastic methods is evolutionary algorithms. Evolutionary algorithms are optimization algorithms used to study collective dynamics such as the swarming/flocking of many-body systems in Artificial Intelligence, without the specific details of the underlying fitness landscape (e.g. evolutionary game theory). Genetic algorithms, a sub-class of evolutionary algorithms are often employed as
optimization heuristics to searching problems, using the fundamental concepts of natural selection. They are global optimization algorithms which generate the minima in the parameter’s fitness landscape. Consider the replicator equation with a free energy of the system given by:

\[ F(q) = \sum_i q_i E_i - k_b T \sum_i q_i \ln q_i = \langle E \rangle - TS(q) \]

The first term describes the energy of the state and the second term is the Shannon information entropy for a reaction coordinate \( q \). At equilibrium, the Probability \( P \) of the system being in the \( i \)-th state is given by the Boltzmann factor and \( F \) is a measure of the fitness of the species (i.e. how it grows and spreads) assuming continuous Bayesian-hypothesis updating. Mutations help replicators evolve. The probability that a randomly chosen organism belongs to the \( i \)-th species is:

\[ P_i = \frac{P_i}{\sum_j P_j} \]

The fitness landscape well-defines current stochastic descriptions of the Waddington energy landscape.

The Langevin equation:

\[ m \ddot{r} = F(r) - \gamma \dot{r} + \sigma \phi(t) \]

is used as the mass-equation of the Waddington landscape, where the Fokker-Planck dynamics give the probability time evolution:

\[ \frac{\partial \rho(r, t)}{\partial t} = \nabla \cdot (\nabla D - F \frac{1}{\gamma}) \rho(r, t) \]

Where \( D \) is the diffusion coefficient, \( \rho \) the probability density, \( \gamma \) the friction/damping coefficient, \( F \) the force on the particle/molecule, \( r \) the relative polar coordinate, \( m \) is the mass of the molecule, and \( \sigma \phi(t) \) is the fluctuation-dissipation term. In the Langevin equation, a particle undergoing a random walk has two forces: the dissipative force (proportional to velocity) and another of random character. Inertial force is neglected due to viscous drag, assuming cells are highly viscous due to the macromolecule crowding, highly packed organelles and cytoskeletal filaments within. Hence, the protein flows in tissue patterning (reaction-diffusion) are considered laminar as inertial forces are dominated by viscous dissipation (i.e. low velocity fluctuations). The Freidlin–Wentzell theorem gives the probability estimate of an Ito diffusion, a stochastic differential equation, as deviating from the mean trajectory on the landscape (Chunhe and Wang, 2013; Wang, 2015). While the Freidlin–Wentzell theorem holds for Brownian motions on Euclidean space, the Gallavotti-Cohen fluctuation theorem is used to describe the phase space dynamics for ergostat dissipative systems (Evans and Searles, 2002). Yet the relative probability that entropy/information of a system far from equilibrium will increase or decrease in a given amount of time (i.e. the transition between two states) is given by \( e^{-\Delta F/kT} \) where \( F \) is the free energy difference of the path integral. The Liouville/Moyal equation is used to describe the density matrix evolution if the quasi-probability distribution is modelled as a Wigner landscape (i.e. probability amplitudes can take negative values from a threshold).

For an ensemble (cloud) of cells undergoing state bifurcations, the distribution of states on attractor space is given by the Fokker-Planck equations:

\[ \frac{\partial \rho}{\partial t} = -\sum_i \frac{\partial}{\partial x_i} [A_i(x) \rho] + \frac{1}{2} \sum_{ij} D_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} \rho \]

where the mean path is found from the Path integral formulation. Here, \( D \) is the diffusion matrix, \( A \) is the drift vector, \( x \) the spatial coordinate and \( \rho \) the probability measure. The mean path is determined
from an Onsager-Machlup functional (path integral) assuming a Gaussian kernel for the noise, as: 
\[ \rho(x, t) = \int Dx \exp[-S(x)] = \int Dx \exp[-\int L(x(t)) dt], \]
where the sum over all possible paths contributes to the trajectory. D is the diffusion coefficient matrix tensor, S is the action and L the Lagrangian or weight of each path. The conservation of probability flux is: 
\[ \frac{\partial \rho}{\partial t} + \nabla \cdot J = 0 \]
and the flux diffusion equation for the probability landscape is: 
\[ J = F\rho - D\nabla \rho, \]
which locally increases (source) or decreases (sink) the probability amplitudes. The fluctuation-dissipation driven cell fate commitments are often modeled using Hills equations and Gaussian noise terms (Wang et al. 2010; 2011).

A destabilization that results in the disappearance of an attractor state constitutes a bifurcation event in a nonlinear dynamical system. A single cell moves around an attractor basin over time, where the relaxation to a steady state potential is given by the Fokker-Planck equations (Li et al., 2016). However, certain markers/gene expressions can switch the cell fate from a basin of attraction into another, typically through chromatin remodelling (e.g. the binding of Yamanaka transcription factors during stem cell reprogramming). At the sudden critical shift of a system’s configuration of steady states \( x(t) \) (e.g. the nonexistence of the progenitor attractor), the effects of the bifurcation (control) parameter \( \mu \) is given by, 
\[ \dot{x}(t) = F(x(t), \mu). \]
In high-dimensional systems such as the Waddington landscape, the bifurcation parameters \( \mu \), are typically not known and used to denote qualitatively the critical transition state (Mojtahedi et al., 2016). In fluid dynamics, the bifurcation parameter is often considered to be the Reynold’s number (Ruelle, 1995).

The energy potential at a given point in the landscape (i.e. a cell fate) is 
\[ V(x) = -\ln \rho(x). \]
The potential \( V \) defines the attractor’s energy minimum (note: this is assuming steady-state potentials, otherwise \( \rho(x, t) \) is complex to model as in the case of strange attractors). By convention, attractors (energy minima) in cell fate decision-making are assumed to be fixed points or periodic orbits. However, in experiment, cell fates (even for non-cancer states), are dynamical systems fluctuating between attractors. Such fluctuations are characterized as stochastic processes (Huang and Kauffman, 2013). In argument, some of these attractors, especially in cancer ecosystems are proposed to be chaotic like the Lorenz attractor, where the cell fate can fluctuate back and forth between two valleys in an ‘apparent random’ motion. How to test the presence of Lorenz-like attractors in cancer/stem cell differentiation will be discussed in forth-coming sections.

Moreover, it remains a debated query whether cancer stem cells are a small subset of a patient’s cancer population (hierarchial model) or any tumor cell can revert its phenotype to stem cell state or if all cancer cells are potentially stem cells depending on its tumor microenvironment/niche? There is growing evidence to the latter, especially in aggressive pediatric brain cancers. Herein it is proposed, cancer cell fates are ‘strange attractors’ in the Waddington landscape. A strange attractor is a chaotic attractor; bounded regions of phase space with positive Lyapunov characteristic exponents and a fractal dimension. Strange attractors are signatures of the frequency-spectra of turbulent hydrodynamics (Ruelle, 1995).

Waddington landscapes are modelled as spin glasses, where the gene regulatory networks form Discrete Boolean networks. Discrete Boolean networks show a probability distribution of states following a power law decay at critical transitions; a signature of scale-free, fractal organization (Font-Clos, 2018). The random walk on a network is assumed to follow the mean-square displacement law, where the anomaly parameter \( \alpha \) decides if the particle trajectory is clustered or dispersed \( \langle x^2 \rangle = D t^\alpha \). Sometimes fractal filamentous structures emerge from diffusion-limited aggregation. For Boolean networks, spin glass topologies form the ‘fitness landscape’. The fitness landscape is the graph-theoretic representation
of the energy landscape, which shows the mapping of vertices of a finite graph to real numbers (e.g. gene expression values). For e.g., consider the NK model, a spin-glass based energy landscape used in modelling complex biosystems. The NK model is a tunable rugged energy landscape used in the computational modelling of the combinatorial optimization of GRN (gene regulatory networks) and complex systems (Kauffman and Levin, 1987; Weinberger, 1991). When a critical value of average number of connections to the nodes of a Boolean gene regulatory network are exceeded, the Hamming distance between initially close two nodes grows exponentially in time (unstable regime) indicating the presence of chaotic attractors. These are NP-complete optimization problems (i.e. can the optimum global minimum be computed amidst the local minima, for a nonlinear dynamical system?).

Spin glasses are tools to model complexity in a system of many-interacting agents and study how phase transitions occur in initially disordered states (emergence) (Mezard and Montanari, 2009). One of the primary principles of complex systems, is the emergence of behaviours/properties not definable by the sum of its parts (local constituents). For e.g., EMT (epithelial-mesenchymal transition) states are reminiscent of glassy phase transitions. Emergent behaviours, especially in soft matter biophysics, pose scaling problems where the laws of physics become blurred (e.g. quantum vs. classical regime).

Artificial neural networks are fundamentally spin glasses. Lang et al. (2014), using single-cell gene expression data, demonstrated the Hopfield recurrent neural network can reconstruct the corresponding cell states of the Waddington landscape. Taking the orthogonal projection of key transcription factors’ gene expression creates a subspace defining an attractor. Correlation measures and z-scores creates the community structures (distinct clusters) corresponding to the gene regulatory networks in the spin glass framework. The energy minimization of the spin-glass’s Lyapunov function determines the landscape topography (Hopfield, 1982). Partially reprogrammed cell fates emerge as hybrids that co-express genes from multiple fates in the landscape (i.e. spurious attractors) (Lang et al., 2014).

The clonal evolution of cancer is currently modelled using stochastic processes, reminiscent of the spin glass methods in cell lineage and pseudo-time reconstruction. Branching processes (or random trees) are used to model population growth as a random walk and map the evolutionary dynamics of clonal hematopoiesis seen in leukemia models. The random trees representing the birth-death process (and mutation) can rapidly undergo exponential growth and bifurcate into lineages, whereby specific mutational subclones can be traced (Altrock et al., 2015). The hematopoietic system does not fully develop until adolescence, where it becomes a fixed number of mature cells. Hence, fixed number of cells are often used to study the fluctuations of hematopoietic stem cells’ fitness landscape. The Gillespie algorithm then stochastically simulates the branching processes in cell fate decision-making and displays cancer’s intra-tumor heterogeneity.

Criticality in spin-glass, phase transitions are reminiscent of chaotic bifurcations in pattern formation systems (Wolf et al., 2018). A simple reaction-diffusion model where complex patterns can emerge is the Fisher-KPP equation, where the cell fate trajectories are depicted as branching diffusion trees and hence, keep the random walk descriptions discussed above intact (Derrida and Spohn, 1988). According to this ‘directed polymers on a tree’ model, reaction-diffusion equations form travelling wave solutions. The speed of its wavefront is the free energy of the polymer and the minimal speed is the phase-transition into a glassy phase (Derrida and Spohn, 1988). The Cayley tree resembles the KPP-Fisher equation’s trajectories while the free energy of the tree’s wavefront is: \( G(t) = \langle \exp(-\beta x) Z(t) \rangle \), where \( \beta = \frac{1}{k_B T} \) and \( Z \) is the partition function (note: equivalent descriptions of criticality apply to Kauffman’s Boolean network transitions). The wavefront velocity is dependent of its initial condition and accepts
solutions of travelling waves given by $w(x - ut)$ where the velocity $u = \frac{1}{\beta} \log[k_B \int dV \rho(V) e^{-\beta c V}]$ and 

V is the volume, if $\beta > \beta_c$. Here, $\beta_c$ is the critical inverse temperature which is the effective noise 

parameter in Hopfield spin glasses as discussed by Lang et al. (2014). However, the probability density $\rho$ 

(network topography) is often assumed to be a Gaussian distribution, limiting the inhomogeneities and complexity of real energy landscapes. The result of these mathematics is a bifurcation tree with nodes at different points on the landscape (attractors) where the energies of the random walk change in time. At 

time $t$, there are $n(t)$ walks and the expectation value $\langle n(t) \rangle = e^{\lambda t}$, where $\lambda$ is the Lyapunov exponent or mean rate and $t$ depends on diffusion $D$. Again, attractors can be of fixed points, limit cycles, periodic orbits or chaotic (i.e. strange attractors). The latter is proposed to be the solution to understanding cancer cell fates.

PATTERN FORMATION

Tumorigenesis is a subset of morphogenesis. Although the theory of reaction-diffusion is well-developed, the model depends on the 'system'. Single cells exhibit molecular heterogeneity regardless of their pathological nature. Mathematical models of cancer generally revolve around simplified Navier-Stokes equations such as reaction-diffusion models, Darcy’s law and predator-prey dynamics. How accurately do the current molecular models in vitro depict the complexity of cancers in vivo (even within the same patient) remain questioned (i.e. context-dependence).

Turing (1952) first coined the term morphogen to describe the chemical oscillations that result in coat patterns of animals. The stripes, spots and spirals self-organized from Turing’s reaction-diffusion systems are given by nonlinear partial differential equations describing the dynamics of local morphogen concentrations. EMT-programs and related-proteins secreted by the CSC (cancer stem cell) niche remodel cancer phenotypes and the microenvironment through dynamic reciprocity. The general form of the corresponding 3D reaction-diffusion equation is given by:

$$\frac{\partial u}{\partial t} = D\nabla^2 u + R(u, t)$$

where $u$ is the velocity or concentration vector, $\nabla$ is the gradient and $R$ is the local reactions (coupling) term (note: this is Fick’s second diffusion-law with a reaction term $R$). The simplest reaction-diffusion equation pertaining to pattern formation is the Fisher- Kolmogorov–Petrovsky–Piskunov equation (FKPP) (Fisher, 1937; Kolmogorov et al., 1937). A simplified growth-invasion system is given by the FKPP equations:

$$\frac{\partial N_1}{\partial t} = r_1 N_1 \left(1 - \frac{N_1}{K_1} - \frac{N_2}{K_2} \alpha_{12}\right) - d_1 L_{N_1} + \nabla \cdot (D_{N_1}(N_1,N_2)\nabla N_1)$$

$$\frac{\partial N_2}{\partial t} = r_2 N_2 \left(1 - \frac{N_2}{K_2} - \frac{N_1}{K_1} \alpha_{21}\right) - d_2 L_{N_2} + \nabla \cdot (D_{N_2}(N_1,N_2)\nabla N_2)$$

$r_1$ and $r_2$ are the growth rates of normal and tumor tissue respectively, $d$ is the cellular susceptibility to 
excess acidic conditions, $N_1 = \text{normal tissue density}$, $N_2(r, t) = \text{density of the neoplastic and } L(r, t)$ 

is the excess proton [H$^+$] concentration (assumed normal tissue to be of logistic growth rate $r_1$ and carrying 
capacity $K$), $\alpha$ is the Lotka-Volterra (competition) strength, $D$ is the respective diffusion coefficients and
death rate is proportional to L (Gatenby and Gawlinki, 1996). For the neoplastic growth, it is the same equation denoted by a subscript 2 and a lack of death amidst excess acid (adaptive). Changes in oxygen and nutrient uptakes, especially in hypoxia (a signature of most glycolytic cancers), are modelled using such equations.

A two-species activator-inhibitor ($A$ and $B$) model with their respective diffusion coefficients $D$, from the Turing (1952) model is given by:

$$\frac{\partial A}{\partial t} = f(A, B) + D_A \nabla^2 A$$

$$\frac{\partial B}{\partial t} = g(A, B) + D_B \nabla^2 B$$

assuming the intracellular protein oscillations have membrane and cytosolic bound state-transitions. Reaction-diffusion equations are used to model ECM (extracellular matrix) degradative enzyme dynamics, cytoskeletal proteins and adhesion remodelling (Ramis-Conde et al., 2008; Domsuhke et al., 2014). In vitro models of cancer well-fit these equations, where given chemoattractant gradients, degradative enzymes and nutrients concentrations, the ECM-stroma dynamics can be modelled as Turing reaction-diffusion (invasion) systems:

$$D = \nabla(-V(x, t)) + r \nabla A(x, t)$$

The first term is the direction (gradient) of the potential $V$, the second term is the chemotactic gradient of species $A$, and $D$ is the diffusion coefficient. In a continuous description of tumor cell density, the change in number of cells $N$ is given by the partial-differential equation $\frac{\partial N}{\partial t} = D_N \nabla^2 N - \nabla(N \nabla E)$, where $E$ is the ECM density, $\frac{\partial E}{\partial t} = -\varphi ME$ and $M$ the degradative enzyme. Considering additional equations for each enzyme rates and oxygen supply, glucose level, nutrients, etc., a simplified model of cancer morphogenesis is permitted. However, these models do not account for spatiotemporal chaos, tumor microenvironment-specific fluctuations and phenotype adaptiveness/plasticity of the complex ecosystem, to name a few absent parameters.

For instance, the "Go or Grow" hypothesis proposes that cell division and cell migration are temporally exclusive events and that tumor cells defer cell division to migrate. The oxygen shortage (hypoxia) in the environment of a growing tumour may well apply for GBM (Corcoran and del Maestro, 2003). Type IV collagen is a major protein component of the vascular basement membrane and its degradation by MMPs (matrix metalloproteinases) and enzymes is crucial to the initiation of tumor-associated angiogenesis and invasion. Group of tumor cells even amidst distant locations can release collagen IV degrading enzymes to the site of invasion, questioning the (non)locality of the cancer stem cell niche.

Variants of the Turing model are currently in practice in (embryonic) developmental biology. In the Clock and Wavefront model proposed by Cooke and Zeeman (1976), the morphogen wavefront progresses slowly in the anterior-to-posterior direction of a population of oscillators (cells), initiating EMT transition. Wolpert (1969) developed the French flag model where morphogens autonomously form segmented concentration gradients (Garric and Bakkers, 2018). Cells then respond to a specific level of morphogen gradient with a specific differentiation route by positional information of the molecules through the nonlinear feedback loops of Gene oscillatory networks (GRN) (Lauschke et al., 2013; Greene and Sharpe, 2015). Once again, chaotic oscillations/couplings of the oscillators are not accounted for in
these models. After all, these are models for animal embryonic development and not necessarily cancer (stem) cells (i.e. model system-specific).

However, signals involved in the vertebrate embryonic development such as FGF (fibroblast growth factor), Shh (hedgehog), VEGF (vascular endothelial growth factor), MMPs (matrix metalloproteinases), E-Cadherins, Wnt, TGF-\(\beta\) (transforming Growth factor), Notch/Nodal pathway, etc. are core stem cell niche pathways and EMT programs. The cancer-stem cell (CSC) niche exhibits aberrant expressions of these embryonic morphogens (Battle and Clevers, 2017).

Chaotic spatial heterogeneity can occur in the Fisher-KPP equation (Pascaal et al., 1993). For instance, the Hagen-Poiseuille flow (laminar flow in a cylindrical pipe) and FKPP equation \(\frac{\partial c}{\partial t} = D \nabla^2 c + \rho c (1 - \frac{c}{c_{\text{max}}})\) have been used to model GBM growth, where \(c\) is the chemical species concentration and \(\rho\) is the tumor density (Cai et al., 2017, Altrock, 2015). The Fisher-KPP equation was used as a prognostic indicator in breast cancer patients undergoing neoadjuvant therapy. The reaction-diffusion equations well-described tumor growth in accordance to clinical imaging data (Weis et al., 2015). The model well-characterized the reaction-diffusion based drug delivery to tumors in clinical settings as well (Jarrett et al., 2018) The FKPP-model of GBM exhibits travelling wave solutions where the tumor invasion wavefronts have a velocity of \(v = 2\sqrt{DP}\) (Harko and Mak, 2015). However, cancer in vivo, is surrounded by turbulent systemic flows (i.e. turbulence at organ level can be \(Re > 1000\) and for exosomes in systemic flows). The flows are more viscous in lymphatic system than in circulatory system due to crowding by immune cells (neutrophils, macrophages, etc.), inflammatory proteins, ECM components, etc. Moreover, the Stokes-Einstein equations are used to simplify the Diffusion coefficients of spherical molecules at low Reynolds number (assuming the kinetic theory of a homogeneous, viscous fluid at equilibrium). The Stokes-Einstein equation is given by: \(D = \frac{k_B T}{6\pi\mu r}\), where \(D\) is the Diffusion coefficient, \(k_B\) the Boltzmann constant, \(T\) the temperature, \(r\) the spherical radius of diffusing molecule and \(\mu\) is the dynamic viscosity.

Linear stability analysis of local equilibria and Darcy's law have been used to model GBM patterning (Yan et al., 2017). Turing equations were shown to well-describe cancer-immune cell invasion dynamics (Zheng et al., 2018). Another recent GBM model is given by Darcy's law where the mass-averaged velocity of solid components \(u_s\) is defined as:

\[
 u_s = -\left(\nabla P - \frac{\delta E}{\delta \varphi_T} \nabla \varphi_T\right)
\]

Where \(\varphi_T = \text{total tumor volume}\), \(E\) the adhesion energy given a double-well potential and \(P\) is the solid pressure (Yan et al., 2017). Darcy's law can be derived from the 3D Navier-Stokes equations for a shallow flow between two plates.

Recently, a simulation modeled skin to behave like a 2D fluid where viscous forces determines pattern formation. Hydrodynamic interactions created patchy, isolated islands of cancerous patterns (melanomas) (Hoshino et al., 2019). Patchiness is a signature of turbulent flows as well (i.e. intermittency). The lateral inhibition in embryonic pattern formation has been described as optical phonons (quantized vibrations) on a lattice, hence, new physics are emerging in the description of morphogenesis (Negrete et al., 2019).
CHAOS THEORY

Chaos in physiological systems is well-established. An intensively studied system is the Mackey-Glass equation, a first-order nonlinear, time-delay differential equation used in many physiological feedback systems, especially hematopoietic control/diseases (e.g. leukemia). As the time-delay increase in the release of circulating WBC (white blood cells), state-dependent (chaotic) period-doubling bifurcations emerge as in the case of leukemia patients (Mackey and Glass, 1977). However, the time-delay dynamics of exosome release by cancer patients is not established and is crucial to approach cancer complexity. Furthermore, exosomes are nanoscopic structures. Hence, plausibly, quantum mechanics and edge of chaos/criticality dynamics may well apply to their emergent behaviours.

The Rössler attractor is an example of a strange attractor observed in simple chemical reaction-diffusions (Rössler, 1976). The Sel'kov model for glycolysis is an important metabolic pathway in which glucose is broken down to make pyruvate. The model exhibits a Hopf/pitchfork bifurcation in glycolytic oscillations (Rensing and Jaeger, 1985). Chaotic oscillations of a single transcription factor can be used to up-regulate certain proteins, or specific protein complexes (Heltberg et al., 2019). Likewise, chaotic attractors emerge in the reaction-diffusion modelling of tumour growth depending on changes in the control parameters (i.e. oxygen concentration, glucose level, tumor volume, diffusion from surface and growth-parameters) (Itik and Banks, 2010). Lyapunov exponents and fractal dimension were calculated for cancer patterning. While the Lorenz attractor has a fractal dimension of 2.06, the cancer models showed ~2.03 indicating a chaotic attractor with Shilnikov-bifurcations (Itik and Banks, 2010).

Ivancevic et al. (2008) showed a Lorenz-like, chaotic attractor best describes the reaction-diffusion of cancer given by \( \frac{\partial \varphi}{\partial t} = \nabla \cdot (D(\varphi, v) \nabla \varphi(v, t)) \), where \( \varphi \) is the density of diffusing material (cancer cells). The emergence of Lorenz-like attractors indicate cancer cell fates are strange-attractors of the Waddington landscape. As opposed to being fixed in a single valley (attractor), a cell fate fluctuates back and forth between two or more valleys in an apparently unpredictable behaviour.

The study of strange attractors was inspired by the search for solutions of the Navier-Stokes equations. The Navier-Stokes equations for an isotropic, incompressible Newtonian flow (i.e. \( \nabla \cdot u = 0 \)) is given by:

\[
\frac{\partial u}{\partial t} = \nu \nabla^2 u - u \cdot \nabla u - \frac{\nabla P}{\rho} + f
\]

Where \( u \) is the velocity vector, \( \nu \) the kinematic viscosity, \( P \) is the pressure, \( \rho \) the fluid density and \( f \) is the external forces. The nonlinear, partial differential equation was reduced to a 1\(^{st}\) order-coupled ODE (ordinary differential equation) to form the Lorenz attractor in attempt to solve weather turbulence (Lorenz, 1963). Interestingly, a variant of FKPP describes Rayleigh-Benard convections when the reaction term \( R(u) = u (1 - u^2) \) (Newell and Whitehead, 1969).

Electric cell impedance recordings were performed in rat’s prostate cancers. The time-series, Fourier analysis of cancer micro-motions were assessed by Taken’s theorem (time-delay embedding) to detect patterns distinguishable from a random signal and chaotic attractor reconstruction. That is, the time-series data was embedded at least 2N+1 dimension (given N data dimensionality) with a fixed time-lag. The attractor reconstruction showed positive Lyapunov exponents in phase portraits (i.e. signature of chaos) (Posadas et al., 1996). Time-delay Hopf bifurcations can indicate tumor relapse (recurrence) and the emergence of aggressive stages of cancer (i.e. indicative of increased chaotic behaviour) (Khajanchi...
et al., 2018). Through time-series cancer data, signal processing methods such as wavelet analysis and object edge detection algorithms can be assessed with lag-time embedding to compute phase portraits of chaotic attractors. The time lapse imaging of reporter-intracellular proteins during pattern formation (e.g. reaction-diffusion of PAR cell polarity complexes in cancer cell division, tracking of EMT pathways, cancer stem cell niche factors, etc.) and time-series data of stem cell differentiation using single-cell RNA-Seq are ideal data sets to identify strange attractors (Briane et al., 2018). In principle, double-cell state reporters with fluorescence time-lapse imaging can map the Lorenz-like attractors in stem cells. Deep Learning algorithms may facilitate the search for these signatures.

Cancer modelling was performed using a three-body ecosystem consisting of host, immune and tumor cells. Assuming initially logistic growth, the Lotka-Volterra dynamics became chaotic as denoted by the period doubling cascades in the bifurcation diagrams (Letellier et al., 2013). The bifurcation analysis revealed Rossler-like attractors with a fractal Lyapunov dimension. Letellier et al. (2013) further proposed the understanding of tumor growth and metastasis as strange attractors will pave ‘dynamical therapies’ (i.e. how drugs alter the fractal complexity and chaotic structures of cancer?).

Hence, attractor reconstruction methods can track chemical oscillations of protein flows that are critical for cancer pattern formation. For e.g., do strange attractors emerge in the time-series imaging of cell polarity PAR complexes distribution in cancer cell division? The use of reaction-diffusion systems to predict cancer dynamics has been compared to forecasting weather patterns using fluid dynamics equations (Yankeelov et al., 2013; 2015; Tang et al., 2014). For centuries, there remains a fundamental roadblock in the study of fluid systems and flow transitions. An unresolved multi-scale flow problem, like cancer, is that of turbulence.

REPROGRAMMING

Reprogramming cell fates was an intractable problem until the pioneering works of 2012 Nobel Laureates Shinya Yamanaka and John Gurdon. Today, the Yamanaka factors can be replaced by a cocktail of small molecules. For e.g., Forskolin, 2-Me-5HT and D4476 in substitute for Oct4 and VC6T (combination of VPA (valproic acid), CHIR99021 (GSK3-inhibitor/Wnt activator), 616452 (TGF-β RI Kinase Inhibitor II) and Tranylcypromine) can reprogram mouse embryonic fibroblasts to iPSC (induced pluripotent stem cell) states (Hou et al., 2013). For e.g., Dppa2/3 facilitates the epigenetic remodelling during reprogramming to pluripotency; an overexpression of which increases reprogramming of fibroblasts to iPSC from 1-2% efficiency to near 80%. It targets chromatin decompactification via DNA damage responses and associated histone marks (\(\gamma H2Ax\)) (Hernandez et al. 2018). The chromatin remodelling allows the facilitated binding of Yamanaka transcription factors (TFs) and overcome the epigenetic barriers to dedifferentiate cells to iPSC states.

Of all current reprogramming techniques, the early retroviral induction presents the lowest efficiency. Micro-RNA-367/302 and synthetic mRNA encoding reprogramming factors generated iPSC clones claimed with up to 90% efficiency in human fibroblasts (Kogut et al., 2018). High efficient miRNA reprogramming methods such as the use of miR302, increases Oct 4 gene expression and suppresses HDAC2 (histone deacetylase 2). There is > 200- fold increase in reprogramming efficiency when the culture media contained antagonists of TGF-β, MEK/ERK inhibitors (mitogen activated kinases) and thiazovinin (ROCK inhibitor) (Xiong et al., 2019; Saito et al., 2019). Many alternate algorithms have been discovered for altering the epigenetic landscape and minimizing the trajectories towards pluripotency attractors (Rais et al., 2013). More importantly, iPSC stem cell reprogramming generates cancer stem
cells as well, indicating there are network dynamics leading to cancer attractors. Algorithms can extract the information of such attractors’ emergence.

LIF (leukemia-inhibitory factor) and 2i (2 small molecule inhibitors; GSK3β inhibitor/Wnt activator and MEKi (blocks MAPK pathway)) with TGFβ (transforming growth factor beta) signaling modulation can keep mouse Embryonic stem cells in a pluripotent state by upregulating the triad of Nanog/Oct4/Sox2 (similar algorithm for humans). For instance, a similar chemical cocktail of small molecules were shown capable of reprogramming human fetal astrocytes into neurons (i.e. change cell lineages) (Yin et al., 2019). It remains a query whether chemical conversion algorithms exist for reverting cancer phenotypes and is the purpose of this communication. However, in the context of cancer stem cells, signals effects vary on tissue type and microenvironment (O’Brien-Ball and Biddle, 2017). For e.g. TGFβ is oncogenic in some settings and tumor suppressive in others (Haigis et al., 2019).

Current efforts are aimed towards reprogramming somatic cells to totipotency (Barker and Pera, 2018; Etoc and Brivanlou 2019). Retrovirus reprogramming to embryonic stem cells can induce a small subset of 2C-like states (2-cell) which fluctuate in cell fate, partially due to histone-modifying enzymes (Macarlan et al., 2012). Such two-cell states, reminiscent of EMT transitions, is proposed as the strange attractors formed by cancer cell fates on the Waddington landscape. Chemical reprogramming platforms and single cell RNA-sequencing allows reconstructing the progression trajectories and cell fate dynamics of stem cells. Hi C Seq and time-series scRNA Seq are used to determine GRNs of reprogramming networks (Ranquist et al., 2017; Zhao et al., 2018).

What is a cancer stem cell? This is a complexity problem. In highly fluid cancers (e.g. leukemia), a cluster of propagating stem cells are identifiable. For instance, DNMT3A-mutant HSCs (hematopoietic stem cells) showed a multilineage repopulation accounting to heterogeneity and post-chemo remission in AML (Shlush et al., 2014). Relapse in AML is attributed to leukaemia stem cells (LSCs) exhibiting quiescence/dormancy and therapy resistance. The 17-gene stemness score (The LSC17 score) is a prognostic biomarker used for assessing AML relapse in clinical care (Ng et al., 2016). Current findings suggest CSC (cancer stem cell) plasticity is less constrained than believed (environment-dependent), highlighting the higher epigenetic burden of pediatric cancers. Transcriptome and proteomic analyses revealed that miR-126 regulates the PI3K/AKT/mTOR signaling pathway, and thereby confers LSC self-renewal, quiescence and therapy resistance (Lechman et al., 2016). MicroRNAs and their associated epigenetic dysregulation may also be biomarkers of cancer sub-types. For example, Embryonal tumors with multilayered rosettes (ETMRs) are rare pediatric brain cancers with microRNA cluster C19MC amplifications and associated DNMT3B deregulation (Kleinman et al., 2014). However, in more solid tumors, stem cell fates are highly unpredictable.

Stem cell fate transitions encounter many environmental fluctuations and molecular heterogeneities. For instance, Nanog heterogeneity arises from fluctuations in gene networks and sudden burst-like transcription (intermittency) in the coexisting states (Smith et al., 2017). The Myc network, one of the Yamanka factors, may be a common point between cancer and embryonic stem cells (Kim et al., 2010; Yan et al., 2018). Moreover, Myc is a global amplifier of gene expression. Single-molecule FISH (fluorescence in-situ hybridization) and live microscopy revealed that Myc increases transcription duration and heterogeneity. Thereby, Myc is a critical attractor in the cybernetics of phenotype transitions and conferring cancer cell fates (Patange et al., 2019).

Transcriptomics revealed glial-tumor interactions in the tumor microenvironment form neuroinflammatory signals regulating brain metastases, such as: upregulated EMT/MET pathways and
calcium signaling (Wingrove et al., 2019). Such pathways contribute to the dynamic switching in cancer stem cell fates. The identity of CSCs to drive tumor growth and resistance have been challenged in brain tumors. A recent finding shows cells expressing CSC-associated cell membrane markers in Glioblastoma (GBM) do not represent a clonal entity defined by distinct functional properties and transcriptomic profiles, but rather a ‘plastic state’ that most cancer cells can adopt (i.e. phenotypic heterogeneity arises from non-hierarchical manner) (Dirkse et al., 2019). Unlike the glycolytic cancers (Warburg effect), GBM rely on both angiogenesis and fatty acid oxidation for proliferation (Duman et al., 2019). In argument, cancer stem cells are plastic states with reversible fate transitions instructed by the microenvironment (Dirkse et al., 2019). Glioma cells in a subset of mesenchymal tumors were shown capable of losing their neural lineage identity, express inflammatory genes, and co-exist with marked myeloid infiltration, reminiscent of molecular interactions between glioma and immune cells. t-SNE projections of scRNA-Seq in GBM found cellular subpopulations resembling different expression subtypes co-occurring in the same tumor which adapt to heterogeneous phenotypes in disease progression (Patel et al., 2014; Yuan et al., 2018). Hence, all GBM cancer cells are claimed to be cancer stem cells as opposed to hierarchy-based, clonal subsections.

On the other hand, another paper’s findings suggest conserved neural trilineage, cancer hierarchy with glial progenitor-like cells at the apex (Couturier et al., 2018) A droplet-based scRNA-Seq of cells from a given patient, generated two or three cancer groupings by t-distributed stochastic neighbour embedding (tSNE) and the Louvain community detection algorithm, clustering different clones within a tumour. Following PCA (principal component analyses) to assess intratumoral heterogeneity in the enriched GSCs of each group, the data suggested that GSCs are organized into progenitor, neuronal, and astrocytic gene expression programs, resembling a developing brain. Strong correlations were found in pathways such as EZH2, FOXM1, and Wnt, as relevant to cancer stem cell self-renewal and tumorigenicity. Moreover, targeting E2F blockers were shown to preferentially affect progenitor GSC proliferation in xenografted mice (Couturier et al., 2018). However, the paper does not reject the micro-environment dependent phenotypic plasticity and interconvertibility of cell fates. Hence, whether there is a fixed hierarchial clustering in stemness or any cancer cell can acquire (or is in) a stem cell fate depending on its environmental cues, is a highly debated, unsolved problem. Cancer being a dynamical system, these data sets must screen time-series scRNA-Seq of cancers in different microenvironmental contexts at different time-points.

As mentioned, a core set of neurodevelopmental transcription factors (POU3F2, SOX2, SALL2, OLIG2) were identified to be essential for GBM propagation in brain cancer stem cells (Suva et al., 2013; 2014). GBM stem cells produce heterogeneity through wound healing pathways and EMT programs to migrate. These pathways are currently used to reprogram human skin fibroblasts to tumor-infiltrating cytotoxic stem cells for therapy (Bago et al., 2017). A note to keep in mind, stem cells plasticity on glass/in vitro and in vivo are highly different. Hence, these findings may not portray the actual complexity within a living system.

Moreover, tumor initiating cells (CSCs) have elevated methionine cycles. Hence, the high consumption of exogenous methionine (a precursor of methylation state of cancer cells) may be a newly emerging hallmark of cancer stem cells (Wang et al., 2019). Likewise, tryptophan metabolism is often linked to cancer’s immune system evasion. CSCs with the ability to form secondary tumours, have been shown to downregulate natural killer (NK) cell activator molecules when they transform into a quiescent state, through a mechanism involving autocrine inhibition of Wnt/β-catenin signalling by DKK1 (Malladi et al., 2016). These are merely examples to illustrate the dynamical complexity of cancer ecosystems and why reprogramming cancer cell fates remains an intractable problem.
EXOSOMES AND BIOMARKERS

The first-discovered, clinically relevant tumor biomarker is CEA (carcinoembryonic antigen). It is highly elevated in certain fetal tissues and found elevated in ~70% cancer patients’ blood tests (Gold and Freedman, 1965). Since then, the search for universal tumor biomarkers has now tuned into liquid biopsy targets: ctDNA (circulating tumor DNA), CTC (circulating tumor cells), exosomes and their and associated epigenetic/proteomic signatures.

Exosomes are usually 30-100 nm sized, heterogeneous packets of information released by cells forming cell-cell and cell-matrix communication networks. Tumor extracellular vesicles (EVs) mediate the communication between tumor and stromal cells, travel to distant regions from bloodstream or lymph vessels and form the pre-metastatic niche to promote metastases (Guo et al., 2019). Exosomes horizontally transfer therapy resistance to tumor cells and induce their aggressive development (Weixian et al., 2014; Hu et al., 2015; Zhang et al., 2018). It remains a query whether oncosomes exhibit nonlocal (quantum) effects due to their nanoscale.

Current EV characterization typically consists of three factors: size exclusion chromatography (nanoscale or density), shape under transmission electron microscopy (cup shape) and surface markers (e.g. TSG101, CD9, CD36, CD81, etc.). The technique of isolation affects the results as well (for e.g., impurities such as lipoproteins and immunoglobulins aggregates form in ultra-centrifugation, prior to extraction from column chromatography). Dynamic light scattering and nanoparticle tracking analysis are used for unlabelled exosome detection by studying its Brownian motion as well.

Tumor biomarkers from liquid biopsies consist of circulating miRNA, ctDNA/cell free tumor DNA, circulating tumor cells and EV bodies/tumor-derived exosomes (as well their post-translational modifications, e.g. methylation pattern of ctDNA or nucleic acids within EVs). Proteins are usually identified by LC/MS-MS (liquid chromatography-mass spectrometry) in EVs extracted from patient’s blood plasma (body fluids) and hence, serve as diagnostic indicators for early detection. Exosomes can be used to monitor disease progression due to their unique cargo/biomarkers such as microRNA, nucleic acids, and pool of proteins found in/on the EVs. Exosomes are especially useful in relapse patients as prognostic indicators, to derive targeted therapies for the aggressive stages. The biomarkers change in type and concentration as the cancer evolves indicating the adaptive changes in the tumor microenvironment (Lane et al., 2018; Huang and Deng, 2019). For e.g., normal exosomes may contain abundant actin and ribosomal proteins, whereas, oncosomes express higher levels of oncoproteins and cancer-associated microRNAs. The nucleic acids within exosomes can be sequenced as well. With recurrence/relapse, the cargos are more specific to the patient and reminiscent of the adaptive tumor phenotypes. For instance, the EVs in relapse stage III melanoma were enriched with mutant BRAFv600E in the lymphatic exudate (Garcia-Silva et al., 2019).

Nanoflow cytometry shows EV output of glioma cells are highly heterogeneous. For instance, EVs released by EGFRvIII-transformed glioma cells were enriched for extracellular exosomes and focal adhesion related proteins. There was a high association of pro-invasive proteins (CD44, BSG, CD151) with the glioma, while the EGFR-negative cells released CD81 expressing EVs. Regardless of the markers, the cerebrospinal fluid showed EVs of EGFRvIII expressing glioma cells caused the GBM to spread and become more aggressive (Figueroa et al., 2017; Choi et al., 2018). Moreover, the EGFR/PI3K/Akt
pathway is a critical feedback loop allowing fatty acid metabolism and lipid synthesis in the survival of malignant gliomas (Guo et al., 2013). Clinically, exosomes derived from serum samples of patients with glioblastoma bear specific EGFRvIII (epidermal growth factor receptor variant 3), reinforcing tumor-derived exosomes as sources of biomarkers reflecting the status of origin cancer cells (Skog et al., 2008). The emergence of new heterogeneous sub-clones, post-conventional therapies, is a key signature of GBM. For e.g., some glioma will transform into gliosarcoma, another malignant cancer. Another complexity in most refractory cancers is senescence and transitory states from. Senescent cells are dynamical systems that can spontaneously reactivate from dormancy/growth arrest. However, exosomes are emerging as the interlinking mechanism by which adaptive therapy-resistance and heterogeneity are transferred to tumor cells (i.e. horizontal transfer of malignant traits).

Consider pancreatic ductal adenocarcinoma (PDAC), a highly lethal disease characterized by late diagnosis, adaptive heterogeneity, treatment resistance, lack of biomarkers and morbidity. Various Moffitt transcriptomic subtypes have been identified using Imaging mass cytometry and molecular profiling/RNA Seq with limited therapeutic applications. For e.g., GATA 6 (activates Wnt pathway) is high in the classical subtype of pancreatic cancers (Martinelli et al., 2017). The presence of hybrids or reversible, mixed types in relapse are still in query. However, senescence is a crucial factor behind PDAC intermittency and relapses. Pancreatic cancers are usually highly hypoxic and activate cancer associated fibroblasts (CAFs) to secrete extracellular matrices that confer its protective cancer stem cell niche. The ECM-niche (tumor microenvironment) signaling is what determines cancer cell fates and hence, can convert a transitory state or growing cancer cell into a cancer stem cell (and vice versa).

For instance, the pancreatic stellate cells produce vitamin A and vitamin C. In cancer, these convert into cancer associated fibroblasts (CAFs) and fuel cancer stemness. Hence, the supply of retinoic acid and vitamin D inducing drugs are used to revert these states. For example, CXCR2 inhibitors used in asthma/respiratory illnesses can disrupt tumor-stromal interactions in pancreatic cancers and inhibit NETs (neutrophil extracellular traps) transmissions from tumors and is emerging as a targeted combination therapy for PDAC and KRAS-mutant lung adenocarcinoma (Dart, 2016). Verapamil for instance, a calcium channel blocker and anti-hypertensive, is effective in treating as a combination drug for some pancreatic cancers. Regardless, pancreatic cancer remains highly morbid and difficult to treat alike GBM. Regardless of the cellular complexity, the pancreatic cancer exosomes are governing the pre-metastatic niche dynamics (Costa-Silva et al., 2015; Zhao et al., 2016).

The In vivo tracking of exosomes with multiplexed EV Reporters and intravital microscopy or bioluminescence imaging is another emerging clinical tool for early detection of cancer/relapses. Note, however, bioluminescence imaging may not be an accurate representation of the actual tumor size. For e.g., luciferase/Rluc or Gluc (biotin acceptor protein) combined with transmembrane domains, such as lactadherin, can label exosomes and be coupled with medical imaging-based techniques for tracking the spatiotemporal distribution of EVs in vivo (Chao et al., 2018). In a recent study, EV-GlucB were injected into athymic nude mice via the retro-orbital vein and live imaged. The tetraspanin marker and CD63-pHluorin (pH sensitive GFP derivative) with luciferase-biotinylation were used as exosome tagging methods, while total internal reflection fluorescence and dynamic correlative light–electron microscopy performed the in vivo tracking (Verweij et al., 2018; Lai et al., 2014; 2015). Moreover, exosomes are promising targeted drug delivery technologies as well (Gangadaran et al., 2018; Chao et al., 2018))
Exosomes are being recognized as diagnostic tools in other diseases too. For instance, neurofilaments and amyloid fibril proteins are found abundantly in Alzheimer’s patient exosomes (Sinha et al., 2018). Moreover, non-target specific chemotherapy is known to cause cancer cells to release pro-metastatic exosomes (Keklikogolu et al., 2019). Targeting exosomes should be the direction of future therapies in oncology. Hence, exosomes are heterogenous and adaptive, indicating alternative biophysical and complexity-based (machine learning) approaches are required for exosome screening/profiling. The emergence of nanofluidic technologies (NEMS-nanoelectromechanical systems) will pave point of care diagnostics with exosomes.

However, the most prominent feature of exosomes is as reprogramming machineries of cancer phenotypes. The human embryonic environment can suppress cancer phenotypes, via extracellular vesicles (Camussi et al., 2011; Zhou et al., 2017). An embryonic microenvironment might have the capacity to reverse the metastatic phenotype of cancer cells where exosomes from hESC (human embryonic stem cell) microenvironment has recently been shown to reprogram malignant phenotypes to healthy-like plastic states (Zhou et al., 2017). The reverse also holds true (Abdouh et al., 2017). An elevated expression of the Yamanaka factors in the vesicles were shown crucial for the plasticity transformation. Exosomes can reprogram distant cells/tissues and transform the microenvironment towards oncogenic microenvironments serving as hotspots for metastases. The understanding of the cooperation between tumor-derived exosomes and NETs in cancer-associated thrombosis is emerging as well. Nanoparticle tracking analysis and imaging microscopy show tumor stem cell-derived oncosomes prime neutrophils and NETs to promote cancer progression by developing the inflammatory conditions for a pre-metastatic niche (Leal et al., 2017; Hwang et al., 2019).

Multi-omics can map embryonic stem cells bifurcating to pluripotent states, where the phosphoproteome dynamics (kinase-substrate networks) precede the changes in epigenome and transcriptome governing cell fate transitions. The signaling dynamics of ERK, mTOR (mammalian target of rapamycin), etc. were shown to be critical in such early state bifurcations (Yang et al., 2019). Recent clinical trials show RNA-Seq and proteomic assessments increase the molecular targets in precision oncology (Le Toumeau et al., 2019; Rodon et al., 2019). The TARGET study shows analysis of ctDNA mutations matching with tumor biopsy provides more precise molecular targets in patient-care. Other research-platform based tools are emerging that may help clinical oncology as well. For instance, screen drug sensitivity assays for in vivo xenografted organoids has been demonstrated in ovarian cancers (Kopper et al., 2019).

Hence, to summarize, organoids, CRISPR screens, organs-on-chips (microfluidics/nanofluidics), proteomics and most importantly, multi-omics screening/scRNA-Seq analysis will provide precision oncology in Deep Learning Healthcare. The molecular profiling of patients’ tumor/liquid biopsies, microbiota, inflammatory markers, etc. can be fed into deep learning algorithms for optimal clinical decision making; a prototype of which is IBM Watson. Moreover, screening/sequencing the cargo of exosomes (proteins, RNA, DNA and lipids) may provide a more holistic picture of the disease progression and driver mutations. While the search for biochemical signatures across cancer exosomes is a current problem, there should be greater emphasis on the biophysical properties of oncosomes, such as their flow visualization, vibrational properties, spectroscopic (frequency) analyses, etc. A biomarker does not have to be an expressed molecule, driver mutations, methylome patterns in ctDNA, etc. It can be a characteristic pattern in the frequency spectra, a vibrational signature (e.g. Raman spectral band) or flow patterns (reaction-diffusion kymograms) in time-lapse imaging. Hence, even an equation that
distinguishes normal phenotypes from the pathophysiology of interest is a biomarker (i.e. a strange attractor/NSE equations as proposed for cancer). Hence, the biophysical properties of exosomes such as their flow patterns and spectroscopy bands (e.g. vibrational spectra) are emerging cancer biomarkers.

**ARTIFICIAL INTELLIGENCE**

*Complexity theory* is the study of complex systems, their emergence, self-organization, nonlinear dynamics and the multiscale flows of their interaction networks. Biocomplexity provides a holistic, systems biology approach to cancer ecosystems with interdisciplinary techniques for analysis (i.e. the whole is greater than the sum of the parts). Complexity is related to NP-completeness, as seen with decision-making and optimization problems in cancer ecosystems. Cancer, however, is a complex adaptive system. That is, the interacting parts respond to the stimuli of the environment and changes its behaviours through nonlinear feedback loops (signal transduction). The Hanahan-Weinberg hallmarks (e.g. autocrine feedback/self-regulated growth signals, telomerase overexpression, immune evasion/invasion, anti-apoptosis/death, etc.) highlight cancer adaptiveness.

**Question:** Is searching necessary to find a needle in a haystack? This depends on our tools, if we had a magnet, we don’t have to explore the vast space of possibilities to find the needle. This is an analogy of optimization. Finding the clique/master GRN (gene regulatory networks) controlling CSC cell fate transitions or modelling protein folding for instance are such searching problems that are combinatorially complex. The P vs. NP problem asks, can we solve searching problems without searching? NP (non-polynomial) problems are quickly checkable but solved by searching (difficult). Are easily checkable problems also easily solvable is the question. The problem was first informally addressed in a letter by Gödel to von Neumann (1956). If P=NP, then if you solve one problem, you solve a class of similar problems. The Navier-Stokes Equations (NSE) is a complexity problem, whereby the search for strange attractors is the algorithm (magnet). These optimization problems are studied in a branch of algorithmic information theory, known as Kolmogorov-Chaitin complexity. Algorithmic information theory and Kolmogorov complexity use tools in artificial intelligence such as machine/deep learning to predict causal trajectories in optimization problems. For e.g. searching, decision and optimization problems pertaining to graph theory/network flows are treated with information-theoretic measures and algorithms. Algorithmic information theory treats cancer as a computer program and hence, the information dynamics within the cancer ecosystem as problems of algorithmic complexity.

Phase transition (or, percolation) is when a few connections of a sparsely connected digraph system flips to a rich connectivity or vice versa (i.e. a network fragmented into small clusters). Phase transitions describe illness onset and immune network breakdowns (edge of chaos) (Bossamaier and Green, 2000). They also characterize the emergence of EMT switches or stem cell transitions in the Waddington landscape, as well as the transition from laminar to turbulent flows (Lemoult et al., 2016). Current approaches in modelling phase transitions mainly revolve around stochastic processes and spin glasses.

Identifying a minimum number of transcription factors to reprogram differentiated cell states to pluripotency was an NP-problem, experimentally solved by Yamanaka, Gordon, et al. The query now lies on whether cell states can be reprogrammed to earlier fates or more importantly, reprogramming stemness in a disease context (i.e. cancer stem cells to non-malignancy). Cancer phenotype is reprogrammable to non-malignancy by its environment, as has been demonstrated by the transfer of exosomes from sera (Zhou et al., 2017). However, identifying the minimal circuit to reprogram cancer cell states is currently a complexity problem.
Deep learning healthcare will allow AI (artificial intelligence)-assisted decision-making in precision oncology (Esteva et al., 2019; Topol, 2019). Digital pathology, cell-lineage reconstruction from biopsies, evidence-based drug screening (i.e. personalized pharmacogenomics), the use of machine-learning algorithms and data-science to identify novel therapeutic targets/biomarkers and the assessment of time-series, multi-omics to identify key GRNs regulating cell fate commitments are merely few examples of Deep learning healthcare (Silberbush et al., 2019; Yang et al., 2019).

Deep convolutional networks use multi-layered information processing and back-propagation to classify-predict complex signals such as images, speech and video data (LeCun et al., 2015). As discussed, fractal dimension analysis is a simple tool for distinguishing cancers from healthy tissues (Lennon et al. 2015) However, deep learning nets are revolutionizing clinical pathology and medical imaging reconstruction as it incorporates many layers of neural network architectures coupled to machine learning algorithms (Shan et al., 2019; Zhang et al., 2019). Unsupervised methods of machine learning can perform feature extraction and pattern recognition on newly presented cancer data, while supervised learning methods classify cancer from non-cancerous data sets based on trained databases. Deep Learning can be coupled to the latest of diagnostic imaging and pathology analyses. Three-dimensional, live-cell imaging of patient-derived tumor organoids are currently used for single-cell fate tracking and karyotype sequencing to study CIN (chromosomal instabilities) (Bolhaquiero et al., 2019). There are molecular cytogenetic techniques, such as comparative genomic hybridization (CGH), Fluorescence in situ hybridization (FISH) and spectral karyotyping (SKY) that study how chromosomal aneuploidies affect gene expression profiles (transcriptome) and tumor heterogeneity (Wangsa et al., 2018). Hence, many emerging techniques are capable of simultaneously tracking chromosomal structure anomalies and single-cell gene expression patterns in cancer, as dynamical systems.

Deep convolutional networks can assess complex drug interactions in patients. They are also used for quality assessment of protein folding and protein structure prediction from sequence (Tong and Altman, 2009). Wavelet analysis, object detection and similar machine algorithms are employed by deep learning neural networks for pattern recognition on cancer imaging data. Deep learning can iteratively group the learnt features and subsequently update the weights of the network towards an energy minimum. In principle, this can be applied as unsupervised training to recognize chaotic patterns in cancer time-series gene expression profiles and time-lapse imaging. Hence, the P vs. NP problem herein asks: Given Multi-omics map the cell fate transitions from stem cell populations, can Deep Learning networks find the minimum number of transcription factors and/or GRNs to reprogram cancer stem cells? (Fard and Ragan, 2017; Palii et al., 2019).

For e.g., Deep learning nets cluster-classified the diverse subgroups in PDAC (pancreatic cancer) heterogeneity. Six molecular and clinically distinct subtypes of PDAC were identified with 160 subtype-specific markers (Zhao et al., 2018). Time-series multi-omics data and machine learning is being used to predict metabolic pathways in cancer resistance (Castello and Martin, 2018). LC-MS (liquid chromatography-mass spectrometry) analysis of >30 cancer types show distinct metabolic drug targets (Li et al., 2019). Deep learning can predict microsatellite instabilities in gastric cancers and thus, improve immunotherapy decisions in patients (Kather et al., 2019). Bayesian hierarchical clustering of heterogeneous stem cells with droplet-Seq technology is feasible (Shun et al., 2019). Nanofluidic Drop-Seq and machine/deep learning will also pave the future of point-of-care diagnostics and personalized interventions. These are merely examples to illustrate oncology is transitioning towards Deep learning healthcare.
However, the most prominent scope of machine learning algorithms in oncology is its applicability in dissecting the GRNs and reconstructing the Waddington energy landscape of cancer stem cells, thereby paving the route towards reprogramming cell fates (Saelens et al., 2019). Recently, convergent cross-mapping with cytokine networks within the hematopoietic system was used to identify causal trajectories in regulatory transcriptional networks to construct a hierarchical, directed graph (Krieger et al., 2018). Cytokine immune networks were inferred from delay embedding mapping with Pearson coefficients to validate their network connection strengths. Such time-delay embedding techniques can be used to infer from the nearest neighbors of multi-omics data (i.e. cell-gene expression matrix), causal maps of cancer stem cell networks in state space. In other words, higher-dimensional embedding techniques analogous to Taken’s theorem can be applied for attractor reconstruction from time-series, single-cell RNA-Seq data. As proposed, cancer pattern formation consists of turbulent flows. Hence, fluid dynamics grid systems are proposed as means of identifying the strange attractors, since analytical solutions do not exist for such complex, intractable problems (i.e. Navier-Stokes equations). Hence, approximate solutions of strange attractors can be inferred from coupling time-delay embedded deep learning to computational fluid modelling (CFD) grid schemes.

The general approach to scRNA-Seq based cell lineage clustering and pseudotime inference is based on data filtering (pre-processing) and machine learning algorithms. CyTOF mass cytometry can be used for protein expression quantification in single cells (i.e. heterogeneity assessment of tissues), whereas RNA-Seq is used for observing differentially expressed genes in cells. The scRNA-Seq counts are usually processed into readable data, where false counts remains a current problem. Hence, the raw data reads are filtered and usually selected for most differentially expressed genes in the cells. The single-cell RNA Seq data is then fed into the machine learning algorithms as a table/spreadsheet of cells (columns) by genes (rows) (i.e. expression matrix), followed by dimensionality reduction techniques like PCA (principal component analysis) or t-SNE, etc. Then, network-graph theoretic approaches are used to reconstruct the spatial neighborhood of cells with statistical inference (e.g. k means clustering (unsupervised) or kNN graph (supervised)). Following, optimization algorithms are used to find trajectories and regulatory modules. Correlation analysis (of activity over time or distance) assigns a score to potential gene-gene interactions, after unsupervised community detection algorithms and partial-information decomposition algorithms are employed. Dispersion cell -cell variability measures generally use covariance, correlation measures, Bayesian-statistical inference and Shannon entropy (mean entropy generally increases in transitions) (Chan et al., 2017).

The discussed algorithms can be easily adapted to any data set, especially with transfer learning for large data sets (Stein-O’Brien et al., 2019). However, these methods can still be misleading as a lot of data filtering is done prior to analysis and valuable information may be lost. We assume cell to cell variability on the reduced latent components from high dimensional data, whereas spatiotemporal chaos can be extracted from the lost data. Low-gene expression levels can result in chaotic fluctuations in cell activity (i.e. sensitive dependence on small perturbations). For example, dynamic Bayesian network analysis is an algorithm, where given protein concentration changes in time, the factors most connected in the regulatory networks are determined to construct the circuitry. Dynamic Bayesian networks model dynamical systems as steady states through probabilistic Boolean networks. However, cancer stem cells are not steady (equilibrium) states.

Gene expression is for most part of literature considered as a stochastic processes. An example, Piecewise-deterministic Markov processes (PDMPs) were shown to well-capture genetic switching in regulatory networks with the same performance measures as Monte Carlo approaches for large number
of chemical kinetics. The probability (stochasticity) that a cell will move to the next microstate depended on how long it spent in the current macrostate (Lin et al., 2018). However, these interpretations are model-system specific (e.g. normal mouse cells vs. cancer stem cells, in vitro). Gene expression variability increases intermittently during differentiation. Hence, stochastic interpretations are not well understood and for most-part speculative. Stochasticity may reflect ‘flickering’ when systems pass through a critical transition point, where the same definition holds for chaotic systems such as burst-like transition to turbulence (i.e. intermittency) (Stumpf et al., 2017).

Topslam estimates pseudotime by mapping individual cells to the surface of a Waddington-like landscape with probabilistic dimensionality reduction (Bayesian Gaussian process latent variable model/GPLVM) (Zwiessele and Lawrence, 2017). These methods project high dimensional data into 2 or 3 components, whereby distances are interpreted as cell-cell variability for cluster analysis and neighborhood graphs. However, such statistical, machine learning methods are sensitive to fluctuations in gene expression data (environmental, intrinsic, etc.) and identify key regulators without understanding the system dynamics nor complexity driven by molecular interactions (e.g. reaction diffusion patterns, epigenetic modifications, etc.). This is generally the downfall for stochastic modelling of complex biosystems.

Cell Router is a graph theoretic, flow network-based trajectory detection algorithm (Lummertz da Rocha et al. 2018). First, dimensionality reduction (PCA, t-SNE, Diffusion map, etc.), is performed on the single-cell data set. Following, a kNN Graph (k-Nearest neighborhood) is assessed. Then, the Jaccard index finds the similarity between two cells (if they belong to the same cluster, high correlation) and the Louvain community (structure) detection algorithm configures the populations (assess weights of the graph by similarity of cell-cell) forming the flow network (i.e. source to sink, directed graph to map the GRNs). In a directed graph, the nodes are connected by flow arrows with the weighting indicating their capacity. The trajectories are found with Bellman-Ford algorithm (cost flow optimization algorithm) and ranked by total flow, cost and length (between vertices). Then, the minimized trajectories are ranked by GRN scores (Pearson-Spearman correlation). Corresponding heat maps and dynamic curves of regulators are projected to obtain trajectories of the identified clusters, displayed as a GRN flow network. The Bellman-Ford algorithm computes the shortest paths from a single source vertex to all vertices in a weighted graph. It is slower than the Dijkstra’s algorithm but can handle graphs with edge weights that are negative.

Seurat is another algorithm that can be used as a pre-processing tool of cancer data sets. Seurat is a scRNA-Seq correlation and clustering computational tool (Butler et al., 2018). The identified clusters can be further sorted using algorithms like CellRouter into flow networks. As discussed earlier, there is growing evidence challenging the hierarchial model of cancer stem cells. In argument, GBM shows that any cancer cell can be a stem cell depending on its microenvironment. ScRNA-Seq cell lineage and Waddington landscape reconstruction algorithms discussed herein can validate these claims.

Single cell Energy path (scEpath) is a method for mapping the quantitative energy landscape of single-cell dynamical processes. It reconstructs cell lineages and pseudotime (cell fate trajectory) inference with information-theoretic measures and statistical-machine learning methods (Jin et al., 2018). scRNA-seq data is pre-processed/ filtered from low gene expressions, then using Spearman correlation of the adjacency matrix, scopath builds a GRN, calculates the normalized energy between expressions, and performs a linear dimensionality reduction with PCA. Data then undergoes structural clustering of cells using an unsupervised framework called single-cell interpretation via multikernel learning (SIMLR). From
the cell clusters, Boltzmann–Gibbs distributions are used to find the transition probabilities. To infer cell lineages, scEpath first constructs a probabilistic directed graph with maximum probability flow (equivalent to finding a minimum directed spanning tree (MDST) (Edmonds’ algorithm)). Then, scEpath uses the R “princurve” package to fit a principal curve of these core cells (centroid) to compute the pseudotimes (trajectories). The pre-processing is similar to Slingshot. Slingshot is a trajectory-cell lineage classification algorithm using dimensionality reduction techniques and cluster analysis (Street et al., 2018). Diffusion maps, PCA, etc. are used as dimensionality reduction followed by model-based clustering (e.g. expectation-maximization algorithm (i.e. maximum likelihood measures) and Bayesian inference) and/or k-means clustering, and lastly MST (minimum spanning trees-Prim’s algorithm) trajectories in inferring cell fate branching. Some of the best pseudo-time inference approaches in mapping cell fate transitions consist of t-SNE followed by MST on single-cell data.

Norbert Wiener and J.J. Hopfield (1982) pioneered the field of cybernetics. The Hopfield network is a recurrent neural network. The Perceptron (a supervised learning algorithm of binary classifiers) is the predecessor of most neural networks. Convolutional neural networks are usually tools for image classification and feature extraction/pattern recognition, whereas recurrent neural networks are typically useful in time-series analysis (Szedlak et al., 2014). Neural networks are being used in reconstructing Waddington landscapes from scRNA-Seq data. Neural nets can perform hierarchical clustering on heterogeneous protein–protein (PPI) and protein–DNA interactions (PDI) datasets, with kNN supervised learning or unsupervised learning (on unlabelled data sets) including on cancer ecosystems (Lin et al., 2017).

Many classification and regression algorithms exist for small sample sizes of image data sets. For e.g. SVM (support vector machines), regression techniques, t-SNE based dimensionality reduction, clustering methods, wavelet analysis, etc., are such classification methods. A neural network useful in image classification of tumor pathologies are GANs (generative adversarial networks). The general goal of neural networks is to train AI for pattern recognition. The adversary (discrimination method) trains the datasets to iteratively optimize between the real data set and a randomly generated false image via gradient descent learning. Likewise, Saliency methods are also tools for image classification from gradient learning. However, for larger datasets as is the case for cancer networks reconstruction, deep learning methods are used.

Hopland is a continuous Hopfield neural network-based algorithm which interprets cell RNA-Seq or qPCR data for pseudo-time estimation and Waddington landscape reconstruction (Guo and Zheng, 2017). The Hopfield algorithm takes in the gene expression data matrix and feeds it into the Hopfield network to construct the landscape’s topography. The cell fate attractors are constructed based on gene to gene expression correlation. First, isomap dimensionality reduction is performed. Isomap generally consists of a kNN nearest neighborhood graph followed by geodesic (shortest path between two nodes) calculation from the Dijkstra’s algorithm and multidimensional scaling. The fast-marching algorithm devised for solving the Eikonal equation (describes wavefront propagation in geometrical optics) was used to calculate the geodesic distances on the landscape as the weights of edges connecting the cells. Each gene is modelled as a neuron in the network, the cumulative energy values of which add up to each cell fate on the landscape. A Gaussian-mixture optimization algorithm was used to infer the parameters, the mean values of the outputs (gene expression values) from the cell lineages at different time points.

The Gradient descent algorithm optimizes the Hopfield network (alike backpropagation in Deep Learning) whereby the activation values (weights of the edges) undergo a relaxation process in the Lyapunov energy function (i.e. find the minima). The Lyapunov energy function gives the energy values
corresponding to the cell states (i.e. high energy corresponds to less differentiated states (hills), while low energy indicates differentiated states/energy minimization (valleys)). In classical theory, cell differentiation paths follow the lowest potential energy in landscape. However, cell fate fluctuations as seen in hybrid states or complex adaptive systems (cancer phenotypes) are not considered here. Moreover, the emergence of complexity theory and soft-condensed matter physics indicate biosystems, especially at the mesoscopic and nanoscales, do not necessarily conform to the laws of classical mechanics, especially, if such systems are chaotic, dynamical systems. Hence, why strange attractors and chaotic fluctuations are proposed to better account for modelling cancer dynamics herein. Rather current approaches model the cell fate trajectory as a path integral/random walk between fixed, stable attractors. The Lyapunov energy function is given by:

\[
E = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij} U_i U_j + \sum_{i=1}^{N} I_i U_i + \sum_{i=1}^{N} \delta_i \int_{0}^{u_i} g^{-1}(u)du
\]

Here \( g \) is the activation function (sigmoid in this case, yet can be ReLu, tanh, etc. such that it monotonically increases), \( U_i = g_j(V_j) \), where \( V \) is the neuron outputs for \( N \) genes, \( W \) is the weights from gene expression data, \( \delta_i \) is the gene signal degradation rate, \( I \) is the combination of propagation delays and regulations/noise from environment. Following, to simulate dynamic trajectories, Euler’s method (\( 1^{st} \) order Runge-Kutta) is used to solve the updating of neurons as an ODE (ordinary differential equation), while the Euler-Maruyama method is used for solving SDEs (stochastic differential equations)/Langevin equations underlying the network dynamics. These two techniques apply for high molecular concentrations, whereas, the Gillespie algorithm is used by the neural network for low copy number (i.e. categorized as stochastic fluctuations) (Guo et al., 2017). To visualize the landscape/network topography, GPLVM, a probabilistic nonlinear dimensionality reduction technique is used onto a grid/stencil with a triangulated mesh, where the fast-marching algorithm computes the geodesics based on cell-cell variability in gene expression clusters and an MST algorithm identifies the minimum trees connecting clusters.

Again, the efficiency of the algorithms discussed are model-dependent and limited by the assumption of classical, steady state- equilibrium statistical mechanics. For instance, Topslam outperformed Hopland in dissecting mouse embryo development, single cell RNA-Seq data without the Hopfield element (using only GPLVM and MST). Moreover, the data sets in these algorithms were highly filtered/processed and are static. Thus, the complex bifurcations and patterns arising in chaotic dynamical systems are not observed.

In proposition, the algorithms can be improved by adopting fluid dynamical grid systems- like LBM (lattice Boltzmann method) or FEM (finite elements method). Network-theoretic approaches to discretize fluid flow exist (Gustafson, 1984; Gustafson and Hartman, 1985). Neural network representations of FEM are theoretically established as well (Takeuchi and Kosugi, 1994). While neural network frameworks of computational fluid dynamics exist, they are not in practice within biological networks reconstruction. Deep learning and machine learning algorithms that can model complex fluid flows is herein proposed as the solution to better model the Waddington landscape reconstruction from time-series multi-omics data. Thereby, fluid-grid based time-delay embedding and clustering techniques are predicted to capture ‘strange attractors’ in cancer ecosystems at multi-scales (i.e. genes expressions, protein patterning of cells, cluster dynamics, etc.).
The most advanced of current approaches in scRNA-Seq lineage reconstruction is scDeepCluster (Tian et al., 2019). scDeepCluster uses a model-based cluster analysis through multi-layered neural networks (i.e. deep learning) and hence outperforms the above-discussed algorithms. Multi-kernel spectral clustering methods and community detection methods are amidst the most commonly used unsupervised learning methods in cell lineage reconstruction. However, such methods do not account for the false zero (low RNA capture) counts and scalability (dimensionality reduction) issues in scRNA-Seq cluster analysis (2019). Deep learning embedded clustering was proposed as a solution. Autoencoder, a type of deep neural networks (DNNs) was used to replace the mean square error (MSE) loss function with a zero-inflated negative binomial (ZINB) model-based loss function. The autoencoder performs a nonlinear function mapping of the read count matrix of scRNA-Seq data to low dimensional latent space. Following, clustering analysis is performed by the Kullback–Leibler (KL) divergence which characterizes the relative Shannon entropy (highly useful in time series data). The KL divergence was assessed by the ‘deep embedded clustering’ (DEC) algorithm along with noise-reduction techniques (Tian et al., 2019).

A finite element method (FEM) or Boltzmann machines in addition to the above-discussed deep learning-embedding and clustering algorithms can better capture fluid dynamical transitions in cell states. The mapping of complex transitory states such as hybrid/EMT states or dynamical cell fates fluctuating between distinct clusters (strange attractors), require: 1) time-series datasets and, 2) fluid dynamical equations into the framework of these deep embedding algorithms (e.g. time-delay embedding and reservoir computing). Moreover, scRNA-Seq technologies are rapidly improving with droplet-based micro and nano-fluidics (e.g. Drop-seq) and hence, the capturing of time-series datasets is better improved.

The lattice Boltzmann method is an efficient algorithm for mapping fluid turbulence in reaction diffusion systems (Chen and Doolen, 1998). Stochastic neural networks such as Boltzmann machines are simple fluid computers yet powerful tools to model fluid dynamics and complex wavefunctions (including many-body quantum systems). For e.g., the finite volume lattice Boltzmann method (FVLBM) based on cell center grids computes turbulent flow dynamics. Dynamic factor graphs are analogous to restricted Boltzmann machines and use gradient-based inference in processing time-series data, yet, function algorithmically very similar to recurrent neural networks. However, Deep learning networks are preferred as they can be coupled to multiple machine learning algorithms and layers of signal processing in the classification of complex time-series datasets. In principle, deep learning algorithms consisting of time-delay embedding layers on a fluid dynamics grid (stencil) will allow the reconstruction of strange attractors seen in complex fluid flows (i.e. turbulence) (Tran and Hasegawa, 2019; Cestnik and Abel, 2019).

Another embedding approach is to use domain translation methods (i.e. image-to-image translation). First, RNA-Seq data reads must be generated at different time points, where fluorescence reporters/labelling can help track cells or genes in time across different sample distributions. Gene expression matrices are then mapped across the different distributions with domain adaptation/translation, which trains a neural network for residual mapping (minimize the distance between domains, then iteratively for the following time-points). Cycle GANs is generally useful for unlabelled/unpaired data sets to map potential loops (cyclic trajectories). This approach is similar to the time-delay embedding for time series datasets using an image-pixel difference or gene expression residual classification. The cycle GAN and labelling methods discussed work well for real world data modelling (i.e. complex systems). For e.g., GibbsNet is an iterative adversial network used for image-to-image translation inferences in complex datasets such as traffic flows (Lamb et al., 2017).
Object detection and tracking can classify multiple channels of fluorescence-labelled gene expression or protein flows in time lapse imaging of cells. This is a counting problem. Image density maps are predicted using supervised learning methods (optimizing the loss based on the MESA-distance (Maximum Excess over SubArrays)). Trajectories are optimally learned by applying a Gaussian peak to label the centroid of cells and predict the flow/location of cells by mapping the center from frame to frame (Lemptisky and Zisserman, 2010). Regression networks with convolutional redundant counting are now used by machines to tackle the counting problems in different heterogeneous populations (e.g. traffic flows, crowds, and cells). Generic Matching Network (GMN) architectures can count any objects in a class of complex systems, including cancer cells tracking (Cohen et al., 2017; Lu et al., 2018).

Moreover, the optimization algorithms (e.g. Bellman-Ford, Dijkstra’s, fast marching) discussed in the graph-theoretic flow networks herein, can be replaced by Hamilton-Jacobi-Bellman (HJB) equations pertaining to fluid models (Farsikov et al., 2000). The solution of the Hamilton-Jacobi-Bellman equation is a partial differential equation that gives the optimal/minimum cost flow for a dynamical decision problem. HJB is a necessary and sufficient condition to find optimal time paths (i.e. finding local minima, attractors and singularities) of control variables in feedback loop systems (i.e. GRN). It is analogous to the Hamilton-Jacobi equation with the energy term minimized with respect to a weight-parameter w(t), given as:

$$\frac{dS(q, t)}{dt} + \min(H(q, p, t; w)) = 0$$

Where, \( p = \frac{ds}{sq} \), H is the Hamiltonian (energy), q and p are coordinates and momenta of the system, t is time and S is the Hamilton’s principal function (upper limit of action integral). This is a complexity equation since min () adaptations are often (currently) intractable. Optimal control problems are generally nonlinear and without analytic solution, hence, falling under the P vs. NP problem (i.e. cannot be resolved in reasonable time as in the current status of solving the Navier-Stokes smoothness and regularity). Therefore, BHJ is often handled with stochastic methods as is the case for all current computational methods in assessing cell fate decision making.

**Question:** Given limited time-series measurements can we predict the future state of the dynamical system? A dynamical system is chaotic if in a bounded phase space, two nearby trajectories diverge exponentially, characterized by the spectrum of Lyapunov exponents. The forecasting of long-term prediction of chaotic systems is difficult due to cascading butterfly effects. This is a complexity problem. As mentioned, time-delay embedding may be the key to overcome such problems in biological networks reconstruction. While Taken’s time-delay coordinate embedding is useful for low-dimensional systems, reservoir computing is a machine-learning algorithm that trains recurrent neural networks to find the Lyapunov exponents of high-dimensional dynamical systems. Reservoir computing permits the chaotic attractor reconstruction in complex fluid flows (Nakai and Saiki, 2018). Both knowledge-free (hidden Markov-like) models and hybrid machine learning models (e.g. causal state modelling, agent-based modelling (e.g. genetic algorithms), etc.) showed good prediction on short-term forecasting of chaotic systems to several multiples of Lyapunov time (Pathak et al., 2018). The Lyapunov time is the time over which the distance of two initially close bifurcations (trajectories) increases by a factor of e (Euler’s number 2.78128…) raised to a characteristic Lyapunov exponent (i.e. exponential divergence). For instance, the reservoir computer (RC) can predict the time-series of the Kuramoto-Sivashinsky equation,
a chaotic system whose pattern formation closely resembles that of fluid turbulence (Pathak et al., 2018).

Current Waddington landscape reconstruction techniques do not account for intermittency in gene expression and the presence of chaotic attractors (i.e. a cell fate fluctuates reversibly between otherwise point/fixed attractors in an apparently random manner). Hence, a Deep Learning net clustering algorithm consisting of CFD (computational fluid) techniques and time-delay embedding layers is proposed as the optimal approach to: map the complex cell fate transitions and test for ‘strange attractors’ in the cancer Waddington landscape. Even simple Hopfield networks can map strange (hidden) attractors. With an activation function of \( \tanh(x) \) to approximate the switches in time-series trajectories, with as little as three or four neurons, strange attractors were mapped in the Hopfield (Li et al., 2005). The 3-Neuron Hopfield can exhibit chaotic attractors for different parameters shown by its bifurcation plot and Lyapunov spectrum (Yang and Huang, 2016). Hence, observing strange attractors in neural networks is feasible. However, tracking them in a highly complex, cancer interactome is easier said than done.

Another promising scope for machine learning is early detection of cancer through liquid biopsies. Machine learning algorithms applied on Intraoperative Raman spectroscopy can distinguish brain cancer stem cells from normal brain tissue (including both invasive and dense cancers) with an accuracy of 92%, sensitivity of 93%, and specificity of 91% (Jeremy et al., 2015). Background subtraction algorithms (e.g. rolling ball algorithm) and autofluorescence removal algorithms were employed to distinguish the cancer stem cell Raman bands from those of healthy tissues (Brusatori et al., 2017; Zhao et al., 2007). Machine learning can help assess liquid biopsies of patients, especially in the emerging field of exosome characterization. Patient-derived exosomes were analyzed using Raman spectroscopy with a 785 nm laser at 5mW irradiance onto calcium fluoride slides. The cluster analysis and interpolation resolved Raman spectroscopy found that cancer and healthy exosomes clustered differently when subjected to simple dimensionality reduction techniques like PCA (Gualerzi et al., 2017; 2019). The clustering also depends on the purity of the technique used to isolate them, whereby, analytical ultracentrifugation had less purity than column chromatography. Similarly, surface enhanced Raman scattering (SERS) signals of exosomes from normal and NSCLC (Non-small-cell lung carcinoma) cells was performed in another study. The Raman spectra of cancerous exosomes showed unique peaks in the Raman vibrational bands distinguishing NSCLC exosomes from healthy clusters when subjected to PCA (Shin et al., 2018). Other works using peak fitting algorithms, background subtraction methods and MCR-ALS algorithm (multivariate curve resolution- alternating least squares) with Raman spectral analysis to cluster-classify EV exosomes on gold-nanoparticle plated SERS substrates have been reported as well (Banaei et al., 2017; Shin et al., 2018).

Such clustering methods can be implemented on the signals acquired from other spectroscopic methods as well (e.g. FT-IR (infrared), Brillouin spectroscopy, etc.) and the distinguished spectral bands can serve as potential biomarkers. For e.g., NMR (nuclear magnetic resonance) spectroscopy can detect exosomes (Ko et al., 2016). Microfluidic chips using \( \mu \)NMRI can sensitively detect circulating exosomes from GBM patients as potential point-of-care diagnostics. GBM-derived MVs exhibit elevated expressions of EGFR (epidermal growth factor receptor), EGFRVIII, PDPN and IDH1 R132H (isocitrate dehydrogenase variant) (Recall: EGFRVIII mutations is the most frequent characteristic signature associated with GBM) (Shao et al., 2012). If the NMR spectra show distinct bands for healthy exosomes vs. cancer-derived exosomes or whether they cluster differently in machine learning algorithms remains undetermined. For e.g., Infrared spectroscopy can characterize exosomes by amide and C-H vibrations. Atomic-force microscopy enhanced infrared spectroscopy and plasmonic biosensors are emerging technologies for quick,
facilitated exosomes classification from liquid biopsies (Chiang et al., 2019; Mihaly et al., 2017; Qiu et al., 2019). Hence, spectroscopy and interferometry may pave the future of early cancer detection with exosomes and allow point-of-care diagnostics with microfluidics.

On a final note, a recent study showed that RNA MuTect found environmental mutagens generate many mutations in skin, lung and esophagus (i.e. air we breathe, our surrounding contact, the food we consume, etc.) in normal tissues, mostly in cancer promoting genes (Yizhak et al., 2019). This further raises questions about our lifestyle choices and environmental conditions. Moreover, this outlines the above-discussed methods can be implemented in many levels of oncology care (i.e. pre-clinical/preventative, diagnostic and prognostic/relapse).

TURBULENCE: THE BEGINNING

Turbulence is Universal. It occurs in all scales from the very large to the very small; within the astrophysical, geological, ecological, biochemical and quantum regimes. It is observed in the heat flow of coffee cups, the motion of galaxies, flickering solar flares, gusts of wind, the swirls of Van Gogh’s *The Starry Night*, pattern formation of clouds, predator-prey dynamics, blood flows in arteries, the dance of exosomes, airflow in the respiratory tract and the paintings of children (Lorenz, 1963; Wensink et al., 2012; Arts and Culture, 2019). Likewise, turbulence also describes the flow of plastic pellets seeping into our oceans, the transmission of carcinogens in air, the vortex streets of cigarette smoke and the clouds of atmospheric pollutants enacting as the precursors to cancer incidence.

For instance, the simulation of an ecological system was performed with three trophic levels: Nutrient, Prey and Predator, corresponding in the fluid system to laminar flow, turbulence and zonal flow, respectively. The ecological predator-prey dynamics were shown to be well characterized by intermittent puffs in a pipe for a turbulent fluid flow. The dynamics are driven by the competition between different wavenumber (frequency) selection mechanisms and concluded to correspond to the Navier–Stokes equations (NSE) (Shih et al., 2016). Hence, the NSE equations are observed across all scales.

*What is the NSE global regularity, smoothness and existence problem?* The problem poses: Given an initial velocity vector, elucidate the existence of a velocity field and scalar pressure field which are both smooth and globally defined that solve the equations? Whether the Navier–Stokes equations allow solutions that develop singularities in finite time remains unresolved. However, there is immense body of experimental and theoretical works confirming the topologies of turbulent flows consist of strange attractors (Ruelle, 1980; 1995; Landford, 1982; Miles, 1984; Brandstater and HL Swinney, 1987).

Leray (1934) showed the existence of weak (viscous) solutions of NSE in the sense of distributions satisfying the Cauchy-Schwarz energy inequality by spectral decomposition. According to L.F. Richardson, turbulent motion is a superposition of eddies undergoing an energy cascade (i.e. bifurcate into a fractal hierarchy of smaller eddies). The eddy size $\lambda$ corresponds to a fluctuation of wavenumber $k \sim \frac{1}{\lambda}$. The structure of turbulence remains an unsolved problem. Experimentally, even the turbulence of smallest eddies is patchy and intermittent. Regardless, the $E(k)dk \sim \epsilon^{2/3} k^{-5/3} dk$ (the Kolmogorov energy distribution) holds well with the spectra measured in experiments for low order moments where $\epsilon$ is the energy flux and $E$ is the energy density (Kolmogorov, 1941). In experiments, turbulence arises in sudden large bursts, one after the other, followed by a period of relative quiescent behaviours (i.e. 29
Turbulent phenomena in fluids are characterized by strong fluctuations and power-law spectra suggestive of correlations observed in critical phase transitions (Goldenfeld, 2006).

In topological fluid dynamics, Moffat’s theorem states that helicity (knottedness of vortex lines in a flow) can be interpreted as knot invariants (Arnold and Khesin, 1991; Ricco, 2000). That is, the vortex lines of turbulent flows form complex knotted structures (i.e. vortex tangles) that are stationary solutions to the Euler-Navier-Stokes equations and remains unchanged with transformation. **Theorem:** Suppose that $t^*$ is the largest time of existence of a smooth solution of the initial value problem for Euler’s equation in $\mathbb{R}^3$, then \[ \int_0^{t^*} \max |\omega|(s)ds \to \infty \text{ as } t \to t^*, \] where $\omega$ is the vorticity (local rotation of a fluid). Consider the diffusion of the vortices in a turbulent flow given by the vorticity equation:

\[
\frac{D\omega}{Dt} = (\omega \cdot \nabla)u + v\nabla^2 \omega
\]

Where $v$ is the kinematic viscosity and $u$ is the velocity, $\omega$ is the vorticity defined as $\omega = \nabla \times u$ (i.e. Vorticity is the curl of the flow velocity vector $u$). In 2D, it is assumed ergodicity results in globally smooth solutions to the initial value problem of Euler-NSE equations, yet in 3D no analogous results are known due to the loss of regularity of initial data in finite time (chaos). Vorticity in 3D encounters a blow up as it becomes: $\frac{D\omega}{dt} = u \cdot \nabla \omega \to |\omega|\omega$ and the blow-up time becomes $\frac{1}{\text{mod}(\omega)}$.

Tao (2016) used logic gates and wavelet analysis to design a system of ODEs similar to the NSE, and construct a blow-up solution given the averaged version of Euler equations behaves like a von Neumann machine (a simple self-replicating computer used in fluid modelling). Blow-up solutions for the Euler equations are known to exist for sufficiently small scaling exponents as may be the case for biosystem turbulences (i.e. low Reynolds systems). Current numerical approaches in approximating the solutions to NSE consist of using deep learning neural networks and Boltzmann machines to model complex fluid flows (Kutz et al., 2017). Deep learning in fluid dynamics will pave better flow optimization techniques for the previously discussed graph-theoretic, cell fate reconstruction algorithms (Brunton et al., 2019). Whether ‘quantum machine learning’ algorithms can further elucidate the complexity of such problems remains theoretical (Sarma et al., 2019). Regardless, the fluid computers/machines discussed herein are tools that can be implemented in the Waddington landscape reconstruction from cancer stem cell, time-series datasets.

A decay function of time correlation is observed in experimental turbulence for which only for Axiom A diffeomorphisms, it is proven that the intensity $F(t)$ decreases exponentially at infinity (i.e. continuous frequency spectrum- which itself is the accumulation of large number of independent frequencies). An Axiom A diffeomorphism is defined as follows: let $\varphi: M \to M$ be a $C^1$ map between compact, differential Riemannian manifolds that have a differential inverse. Then, $\varphi$ satisfies Axiom A if its non-wandering set $\Omega(\varphi)$ is hyperbolic and its periodic points are dense in $\Omega(\varphi)$. An example, is the Smale Horse-shoe map. A continuous frequency spectrum is a signature observed in all cases of experimental fluid turbulence. The frequency spectra (i.e. square amplitude of each frequency of a fluid) can be used to detect strange attractors through attractor reconstruction methods and peak analysis. There is no direct test to ‘sensitive dependence on initial condition’ in hydrodynamics. Hence, a frequency analysis of the fluid velocity (square amplitude vs. frequency) must be performed (Ruelle, 1973; 1995). These methods can be implemented on claimed systems of chemical turbulence in cell pattern formation.
Given the time-evolution of a fluid as a dissipative dynamical system, 
\[ \frac{d}{dt} X(t) = F_\mu \left( X(t) \right), \quad X \in \mathbb{R}^m \]
where \( X \) is the state vector and \( F \) is the intensity function defined above, assume the bifurcation parameter \( \mu \) is the Reynolds number. For small \( \mu \), steady state solutions exist while at larger \( \mu \), the oscillations asymptotically approach strange attractors (Ruelle, 1995). A small \( C^2 \) perturbation of a quasi-periodic flow on a 3-torus can produce strange axiom A attractors. An (ergodic) Axiom A attractor is non-trivial (strange) if it doesn't consist of a single periodic orbit and has a fractal dimension (Ruelle and Takens, 1971).

The competition between the inertial forces (velocity) and the viscous forces defines the Reynolds number. This presents a scaling problem, as what Reynolds number produces turbulent flows varies from one system to another. The (normal) cell cytoplasm is often assumed to be dominated by viscous forces due to macromolecular crowding (i.e. cytoskeletal filaments, organelles, etc.), thereby limiting the movement of intracellular protein to Brownian motion. Diffusion acts to redistribute the conserved quantities between neighboring compartments. However, phase-transition to chemical turbulences can occur in reaction-diffusion systems. Chemical turbulence in protein-mediated cell patterning was recently confirmed experimentally (Brauns et al., 2018). A single-mode lateral instability can lead to chemical turbulence at the onset of pattern formation in cells. Paradoxically, following turbulence transition, coherent patterns emerge through diffusively coupled local equilibria. Such long-range orders (standing and traveling waves), far from onset of the (subcritical) lateral instability are observed in morphogenesis and cell division.

The formation of cancers in organs and their metastases demonstrates a scaling over many orders of magnitude, from exosome flow to circulating tumor cells (CTC) in blood vessel/lymphatic networks. Such biosystems comprise of multi-scale patterns of flow. For instance, at the organs level, the Reynolds number may be around 1000, in the aorta up to 10,000, etc. (and depending on the pathophysiology, the turbulence may increase to higher levels). Exosomes are in the nanoscale. Thereby, in principle, their flows must account for a few fold differences in Reynolds numbers in comparison to circulating tumor cells, in systemic flows. The Reynolds number is a dimensional analysis argument. In a cylindrical tube, whether we consider the length scale as the radius or diameter will change the Reynolds number by a factor of 2. The critical Reynolds number were shown to be in the order of 100-200 for Taylor-Couette and Rayleigh–Benard convection systems, which is relatively feasible in/around cellular biosystems (Ruelle, 2012; 2014).

Systemic circulation exhibits turbulence, yet the turbulent kinetic energy increases with ageing and pathophyslogies (Hojin et al., 2018). Hematopoietic iPSC (induced pluripotent) stem cells were shown to increase their production rates in systemic flows by many orders of magnitude under turbulence (Ito et al., 2018). Low levels of turbulence generated in turbines increased the levels of platelets produced by the iPSC-derived megakaryocytes significantly. Turbulence stimulated macrophage migration inhibitory factor (MIF), insulin growth factor binding protein 2 (IGFBP2) and nardilysin, accounting for the increased platelet generation. Hence, such mechanisms may be utilized by cancer stem cells to undergo rapid clonal expansion during metastatic flow. Cancer stem cells of different tissues and molecular groups must be placed in turbulent flows as within this experiment, to observe what gene expression pathways are activated in turbulent flows. Moreover, the use of turbulence to increase platelet and HiPSC stem cells may help progress with the clinical therapies in hemato-oncology.

Cancer metastasis, growth and therapy response is governed by fluid dynamics (Koumoutsakos et al., 2013; Goetz, 2018). Using shear-stress tensor in terms of viscosity and the fluid velocity gradients, focal...
adhesion dynamics can be investigated in cancer metastasis (Thamilselvan and Basson, 2004; Chivucula et al., 2014). The hemodynamic shear forces modulate the formation and location of pre-metastatic niches through regulating CTC flow. Microfluidic experiments on 3D ovarian cancer models show that fluidic streams induce a motile and aggressive tumor phenotype. Hence, higher fluid flow were shown to induce EMT transitions and promote aggressive phenotypes that can metastasize (Rizvi et al., 2013; Ketene et al., 2012). Furthermore, Huang et al. (2018) showed that non-laminar (turbulent) shear stress may increase the adhesive ability of cancer cells in metastatic invasion.

Microfluidic devices/flow chambers are reshaping experiments on tumor dynamics (i.e. lab/organ on a chip). For instance, carcinogen exposure from food (gastric juices) or cigarette smoke can be mimicked on layers of patient-derived gastric or lung cancer cells in the chip. Videomicroscopy then maps the dynamics in the chamber along with appropriate biochemical techniques (e.g. flow cytometry and RNA-Seq) to profile the adaptive changes. Exosomes profiling and virtual drug screening can also be performed in such microfluidic, and nanofluidic chips.

Blood flow is often modelled using the Hagen-Poiseuille equation assuming vessels to be cylindrical pipes and incoming flows to be laminar. However, tumor vasculature does not conform to this naïve assumption and most biological fluids are anisotropic/non-Newtonian. The complexity of tumor vasculature and tissue organization can be quantified in terms of fractal dimensions, where such blood vessels showed only intermittent flows (Jain and Baish, 1998; Nasu et al., 1999). Mutifractals is a tool to study the geometry of turbulence (Sreenivasan and Meneveau, 1986; 1991). Hence, the application of Box counting algorithms on tumor vasculature and morphology can study how treatments alter its geometric complexity (Coffey, 1998; Yakhota and Sreenivasan, 2004).

The fluid shear stresses experienced by CTC cells consist of turbulence. Blood flow velocities in the circulation system can range from 0.03 to 40 cm/s depending on the vessel size. How individual CTCs and CTC emboli can withstand these intercellular interactions and persist in the fluid flow is not yet fully understood. However, the Navier-Stokes equations are imposed on Cartesian fluid grids in current modelling approaches and the fluid velocities must be solved from the NSE to measure CTC dynamics. As such, the blood plasma is modeled as a viscous, incompressible Newtonian fluid governed by Navier-Stokes equations (Rejniak et al., 2016). With appropriate boundary conditions, the steady, incompressible Navier-Stokes and continuity equations governing the flow are solved using finite-volume CFD (computational fluid dynamics) to model circulating melanoma cells and their adhesion dynamics (Behr et al., 2015). Doppler techniques can track turbulent blood flow detected in liver tumor patients, in the hepatic vein portal, where the most malignant tumors had greater turbulence (Yasuhara et al., 1997).

Levinthal’s paradox states protein folding is an NP-problem with a puzzling time scale. Although misfolded aggregates have long range interactions with many protein complexes and are aided by chaperones, protein folding displays combinatorial complexity. However, nature resolves the folding problem within seconds. Recent evidences show protein folding may best be described by the turbulence of Kolmogorov-Richardson’s energy cascade, where turbulent eddies and vortices breakdown into smaller fractal hierarchies. The protein folding flows of an SH3 domain protein model have fractal nature and are filled with 3D eddies containing strange attractors, at which the tracer flow paths behave as saddle trajectories (Kalgin and Chekmarev, 2011; Andryushchenko and Chekmarev, 2017). Recall, the SH3 domains are crucial in focal adhesion complexes regulating cancer metastasis and EMT transitions. Likewise, folding dynamics of villin subdomain HP-35 protein in a FRET experiment
was shown to obey the $\beta$-model of turbulence with many orders of turbulent flow transitions for eddies in the 3D conformational space (Andryushchenko and Chekmarev, 2016; Chekmarev, 2018).

Turbulence has been observed in bacterial swarming and the collective dynamics of cellular microorganisms. Active fluids are a rapidly evolving research field inspired by the biophysics of dense suspensions of motile cells (Dunkel et al., 2013). Their hydrodynamic interactions give rise to the emergence of meso-scale vortex patterns reminiscent of two-dimensional turbulence. Cell tissues and reconstituted cytoskeletal solutions exhibit active turbulence, where emergent scaling behaviours were observed. Bundles of microtubules and their associated kinesin motors deposited on an oil–water interface behaved like 3D vortex tangles in a turbulent fluid (Sumino et al., 2012). 2D distortions of nematics consisting of microtubules show quadratic variants of the Navier-Stokes equations in growth (Martinez-Pratt et al. 2019). Self propelled particles like cytoskeletal flocking, follow continuum mechanics like fluids and can exhibit turbulence even at low Reynolds (Marchetti et al., 2013; Baggaley, 2016). Active fluids such as actin and microtubules can be described using the Navier-Stokes equations (James et al. 2018\textsuperscript{a,b}). Coupling a mass-conserved Turing-like reaction–diffusion system for polarity proteins to an active-gel description of the actomyosin cortex, revealed active turbulence in \textit{C. elegans} embryos (Gross et al., 2019). Paradoxically, many meso/nanoscopic systems exhibit emergent properties requiring an explanation from many-body quantum physics and turbulence theory (Doostmohammadi et al., 2018).

Cell division is orchestrated by intracellular protein patterning, mainly from cytoskeletal filaments and cell polarity complexes. Abnormal cell division is the primary signature of cancer. Usually low intracellular concentration of proteins and their corresponding dynamics are modelled as stochastic fluctuations. Classical theory states these chemical systems are close to equilibrium and inertial effects are negligible. However, a recent theory by Halatek and Frey (2018) challenged the dogma and using finite elements method (FEM) simulations predicted chemical turbulence (spatiotemporal chaos) at the onset of the pattern-forming instabilities in tissues. Cytosolic diffusion constants $D_c$ were in the order of $60 \mu\text{m}^2\text{s}^{-1}$ in 1D, where the MinD-ATP/ADP in bulk were given by the reaction-diffusion equations:

$$\frac{\partial}{\partial t} u_D(z,t) = D_c \nabla^2 u_D - F_{u_D}$$
$$\frac{\partial}{\partial t} u_T(z,t) = D_c \nabla^2 u_T + F_{u_D}$$

Where, $u_D$ and $u_T$ define the cytosolic density of MinD- ATP/ADP conformations and assumed the MinD-ATP binds to the membrane via nonlinear coupling rate constants. The resultant kymograms demonstrated turbulent flows at low MinE/MinD ratios. Usually low-levels of gene expression patterns, protein concentrations or signal fluctuations (e.g. in RNA Seq) are filtered out in the algorithms computing the cell lineage tracking discussed earlier. Hence, the strange attractors may be relevant to the low-level protein fluctuations, otherwise classified as ‘noise’.

The theoretical predictions of Halatek and Frey (2018) have been confirmed experimentally (Denk et al., 2018). This must be experimentally repeated in the PAR cell polarity complexes (the mammalian equivalence of the Min proteins) and other associated cytoskeletal remodelling proteins (e.g. cdc42-RhoGTPases/Rac system of actomyosin contractility and focal adhesion dynamics with ECM, morphogens in the stem cell niche, etc.) in cancer (stem) cells during cell fate bifurcations. According to these findings, chemical turbulence in pattern formation is used as a synonym for spatio-temporal chaos, i.e. a broad distribution in the power spectrum and a low spatial correlation length reminiscent of the Kolmogorov spectrum. But none of these terms are strictly/unambiguously defined in the literature. The term was adapted from the work by Nobel laureate Gerhard Ertl on reactions of heterogeneous
catalysis (Kim et al., 2001). According to this work, during chemical turbulence, both the amplitude and the phase of local concentration oscillations are strongly fluctuating creating spiral waves as seen in the Denk et al. (2018) data with Min proteins pattern formation. This diffusion-induced chemical turbulence is typical for oscillatory surface chemical reactions and has been observed under special conditions in the Belousov-Zhabotinsky reaction (Kim et al., 2001). Hence, turbulent chemical oscillations can occur in cellular patterning (reaction-diffusion) systems (Mecke, 1996). The Belousov-Zhabotinsky reaction is often a toy-model system for turbulent pattern formations (Ouyang and Swinney, 1991; Ruelle, 1995).

Turbulence was treated as a heat flow problem using nonequilibrium statistical mechanics of moving fluid particles in three-dimensional lattice boxes (Ruelle, 2012). As mentioned, the Reynolds number is not a well-defined quantity. Consider

\[ \xi_p = \frac{P}{3} - \frac{1}{\ln k} \ln \Gamma \left( \frac{P}{3} + 1 \right) \]

Where \( \xi_p \) is the intermittency exponent, \( \Gamma \) is the gamma function, \( P/3 \) is the Kolmogorov term and \( (\ln k)^{-1} \sim 0.32 \). The average change in velocity \( \vec{u} \) over the small length of a turbulent fluid in these boxes is roughly:

\[ \langle |\Delta \vec{u}|^p \rangle \sim l^{\xi_p} \]

Thus, the onset of turbulence is shown to occur below a Reynolds number of \( \sim 100 \) with numerical methods (independent of geometry) (Ruelle, 2012). Then, the macroscopic fluid transitions on the grid system are given by Boltzmann-Gibbs distributions. Hence, the probability distribution of turbulent eddies in a lattice cube and their lifetimes gives the canonical distribution:

\[ \rho(E) \sim \exp(-\beta |u_n|^3) d^3 u_n \]

Using the second law of thermodynamics, we get the state transition from an attractor A to B as:

\[ \frac{\rho(A \to B)}{\rho(B \to A)} \geq \exp[-\langle S \rangle - \beta \langle Q \rangle] \Rightarrow \rho \sim \exp(-\beta \Delta G) \]

Where \( \Delta Q \) is the heat enthalpy transferred to the active heat bath, \( \beta \) is the coldness of the system \( \left( \frac{1}{k_B T} \right) \), \( S \) is the entropy and \( G \) is the Gibbs free energy. Such a description well applies to the metastable cell states of the Waddington attractor landscape (Ruelle, 2012). As opposed to a heat flow (thermodynamic) approach to entropy, complex systems theory studies entropy as a measure of uncertainty or change in the biosystem’s information flow.

**QUANTUM CHAOS**

Quantum chaos is the quantum description of classically chaotic systems. Although fundamentally all processes are quantum dynamic, a newly emerging field called quantum biology proposes that biosystems may exhibit macroscopic quantum effects at ambient temperature (Lambert et al., 2012). The most controversial postulate of this field, yet most pertinent to cancer, is the plausibility of nonlocal information processing in microtubule networks (Hameroff and Penrose, 2014). However, there is a lack of rigorous experimental evidence to support these claims.

Microtubules are examples of nanoscale, soft matter systems. Other soft condensed matter systems include, polymers, proteins and exosomes. In soft matter physics, the energy scales are comparable
with room temperature. However, in contrast to conventional predictions, the emergence of many complex behaviours require quantum descriptions. An example is the emergence of vortices in turbulent flows. For instance, recently it was shown, thermal fluctuations are sufficient to remodel microtubule lattices’ dissipative dynamics. The growth not only occurs on the extremities as previously thought but dimers can be added anywhere in the entire rod (Shaedel et al., 2019). The tetrapolar spindle generated through cytokinesis defects are used in cancer as a transitory/intermediate state on the route to aneuploidy and genetic diversification in cell division (Lens and Medema, 2019).

Biological systems (especially on nanoscales and below), show there exists processes for which a quantum mechanical description is necessary to fully characterize the behaviour of the relevant subsystem (Marais et al., 2018; McFadden and Al-Khalili, 2018). If further validated, certain proteins in cancer cell patterning, cytoskeletal filaments in cell division (cause of aneuploidy/CIN) and exosome flows are few examples pertaining to the scales at which such macroscopic quantum effects may apply.

Quantum coherent beatings are detected in photosynthetic systems for femtoseconds in the FMO (Fenna-Matthews-Olson) complex, suggesting energy transfer is best described as a linear superposition of chromophores’ energy states (Engel et al., 2007; Thyrhaug et al., 2018). Craddock et al. (2014) propose a longer coherence time occurs for the Tryptophan (Trp) networks in tubulin via Forster dipole–dipole energy transfer, forming coherent beats at 600 femtoseconds. Forster resonance energy transfer (FRET) via dipole-dipole coupling of Trp is predicted as the mechanism responsible for the extended quantum coherence and other plausible nonlocal effects in microtubule networks at ambient temperature.

Frohlich (1968) proposed that a set of biological oscillators can condense to the lowest frequency vibrational mode by the supply of biochemical energy, which he defined as pumped phonons. The pumped phonons are predicted by some to account for the plausibility of macroscopic quantum effects in microtubules in 4 or 5 orders of magnitude of self-similar (fractal) organization (Hameroff and Penrose, 2014). Davydov (1985) tried to find spatial localization of such low-frequency vibrational energy using solitons in DNA, proteins, etc., while Frohlich sought for frequency selection. Recent experimental evidences are suggesting long-range, long-lived coherence of low frequency vibrational modes due to Frohlich-like condensation in the hen-egg white lysozymes (Lundholm et al., 2015). Given the phosphorylation energy of GTP hydrolysis at MAPs (microtubule-associated proteins) binding site, simulations of microtubules as coherent phonons suggest that phonon maxima correspond to MAP binding sites (Prodan and Prodan, 2009). Raman spectroscopy of metabolically active (E. coli) cells claimed to have measured such coherent vibrations as well (Del Giudice et al., 1982). Recent evidence suggests Rabi oscillator-like quantum entanglement with photons can be observed in photosynthetic bacteria at ambient conditions (Marletto, 2018).

Solving the Navier-Stokes equations (NSE) is a complexity problem. Quantum mechanics has fluid-like properties and hence, is interconnected to the NSE problem. Although quantum machine learning may help infer approximate solutions to the NSE, both quantum mechanics and turbulence can only be understood by exploring the underlying complexity of fluid topologies (i.e. vortex tangles) and addressing the fundamentals of quantum chaos. There are alternative formulations of quantum mechanics by Madelung, de Broglie and Bohm that describes nature in terms of fluid dynamics (Madelung, 1927; Bohm, 1952; de Broglie, 1987). The Hamilton-Jacobi equations is the closest formulation of classical mechanics where the motion of a particle can be represented as a wave (i.e. transition to quantum mechanics). The de Broglie-Bohm theory depicts the Schrodinger wave equation as modified Hamilton-Jacobi equations and thereby, allow a quantum analogue of the Navier-Stokes
equations (Harvey, 1965; Bohm and Hiley, 1995). Cosmological models such as SVT (superfluid vacuum theory), superfluid dark matter, etc. propose the fundamental constituent of space-time geometry as a fluid. Such fluid interpretations of quantum theory allow the emergence of macroscopic quantum effects as well.

Bohm metaphorically interprets the fabric of reality as a multi-dimensional ocean known as the Holomovement (Bohm, 1980). Note: The Waddington landscape is a metaphor too. Turbulence in the quantum scale consists of complex vortex tangles, interwoven knots, loops and strings (i.e. analogous to the topologies observed in current theoretical approaches to quantum gravity such as Loop quantum gravity and string theory) (Baggaley et al., 2012). Topological fluid dynamics and braids theory address the same problems in classical scales. Hence, an understanding of turbulence is a foundational, multiscale problem.

The Madelung hydrodynamics equations are quantum Euler equations:

\[
\frac{du}{dt} = \frac{\partial u}{\partial t} + u \cdot \nabla u = -\frac{1}{m} \nabla (Q + V)
\]

The Madelung transformation (1927) and de Broglie-Bohm theory allows the Schrodinger equation to be in hydrodynamic form. Given the polar form of the wavefunction \( \psi(x, t) = R(x, t) e^{iS(x, t)}/\hbar \), and respecting the conservation of probability (fluid density): \( \partial \rho(t) + \nabla (\rho u) = 0 \), by substituting the Schrodinger equation we derive

\[
\frac{\partial S}{\partial t} + \frac{1}{2} \frac{m u^2}{\rho} + Q(x, t) + V(x, t) = 0,
\]

where \( u(x, t) \) is the quantized velocity of the (probability) fluid, \( \rho = m |\Psi|^2 \) is the fluid (probability) density, \( m \) is the mass of the fluid system, \( R = \sqrt{\rho} \) the amplitude of \( \psi \), \( S \) is the phase of the fluid wavefunction (also pertains to the action \( S \) in Feynman path integral formulation) and \( Q \) is Bohm’s Quantum potential (Holland, 1993; Bohm and Hiley, 1995), given by:

\[
Q = -\frac{\hbar^2}{2m} \frac{\nabla^2 \sqrt{\rho}}{\sqrt{\rho}}
\]

The expectation value (most probable value) of the quantum potential is then \( \langle Q \rangle = \frac{1}{2} \left[ \frac{\hbar}{2m} \nabla (\ln \rho) \right]^2 \rangle = \frac{1}{2} \left( \frac{\hbar}{2m} \right)^2 I \), where \( I \) is the Fisher information given as \( I = \int \frac{1}{\rho} (\nabla \rho)^2 \, dv = \int \rho (\nabla (\ln \rho))^2 \, dv \). The Pressure tensor \( P \) is related to the Quantum potential by their respective gradients: \( \frac{\nabla P}{\rho} = -\nabla Q \) (Caldeira and Leggett, 1983; Jungel et al., 2011). The fluid flow is quantized and only accepts solutions of discrete vortices given by:

\[
\oint u \cdot dl = 2\pi \hbar n
\]

where \( u \) is the velocity field and the solutions \( u(x, t) = \frac{\nabla S(x, t)}{m} \) are integer quantized (\( n=0,1,2,3,... \)). Note how the hydrodynamic description of quantum mechanics allows probabilities to be viewed as fluid densities, thereby, deviating from the conventional probabilistic interpretation of quantum mechanics towards one of deterministic chaos.

Most biomolecules are tuned at the edge of chaos/transition point, for e.g., cell membrane fluidity. Power law decay is seen across spontaneous symmetry-breaking and self-organization processes in biosystems. Such emergent behaviours are termed ‘edge of criticality’ processes. Vattay et al. (2014) demonstrated that the transition from quantum to classical chaos is best described as a power law decay. Some biological systems can stay quantum coherent for long time at room temperature near
critical quantum chaos. The decoherence exhibits a power law decay as opposed to exponential decay where above the critical point, chaotic dynamics emerge (Vattay et al., 2014). The prior-discussed far-from equilibrium pattern formation and emergence of coherent structures from the onset of chemical turbulence in cells is an edge of criticality process.

Most biological networks are close to being scale-free, where their node connectivity follow a power law, i.e. the probability that a node is connected with k other nodes (the degree distribution p(k) of a network) decays as a power law. Systems that are fine-tuned at the critical transition point can withstand environmental decoherence and thereby, permit long-lived coherence and transport to coexist for several scales/orders of magnitude. Such processes are observed in the newly emerging field of quantum biology (e.g. photosynthetic energy transfer of FMO complexes, magnetoreception of birds, tunneling in DNA, tunneling in enzymatic catalysis, etc.) (Lambert et al., 2012).

Furthermore, Vattay et al. (2015) used a fractal-box counting method to compute multifractal wavefunctions for proteins (e.g. myoglobin) describing the energy level statistics with the extended Huckel model and random matrix theory. Moreover, experiments of millimetric fluid droplets on vibrating fluid baths show the droplet walkers exhibit chaotic fluid motions. The collective ensembles exhibit quantum mechanical properties similar to those of the double slit experiment’s interference patterns (Harris and Bush, 2013; Moláček and Bush, 2013;). Hence, evidences show macroscopic quantum hydrodynamics can be experimentally conceived supporting the pilot wave theory (i.e. fluid dynamical interpretations of quantum mechanics) (Bush, 2015). Moreover, many soft matter, complex systems including cancer pattern formation may be currently intractable problems as they paradoxically require explanations from: turbulence theory and quantum mechanics.

**CONCLUSION**

The discussed mathematical tools of machine/deep learning, graph theory and experimental fluid dynamics will pave the future of dynamical therapies in oncology and in principle, approach ‘cancer reprogramming’. Deep-learning cluster algorithms and time-delay embedding based on gridding schemes compatible with computational fluid dynamics are proposed as means of testing for ‘strange attractors’ and mapping complex flow transitions in cancer cell fates. Time-series, multi-omics data (scRNA-Seq) and time-lapse imaging of identified intracellular protein patterning within cancer stem cells are proposed as ideal data sets to test the presence of strange attractors. The single-cell transcriptomics and proteomics of cancer stem cells under (simulated) turbulent flows in the microenvironment must also be investigated.

“Big whorls have little whorls Which feed on their velocity, And little whorls have lesser whorls And so on to viscosity”- L.F. Richardson (1922)

**ACKNOWLEDGEMENTS:** Thanks to Dr. Mario D’Amico (Concordia University), Dr. Mohan Malleshaiah (UdeM) and Dr. Rolando Del Maestro, Dr. Arena Goffredo, Dr. Mohamed Abdouh, and Dr. Lorenzo Ferri of McGill University for granting me the knowledge and tools discussed herein.

**DECLARATIONS:** I, Abicumaran Uthamacumaran, am the sole author of the submitted paper. There are no competing interests.
REFERENCES

• The cancer genome atlas research network Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nat. 455: 1061-1068 (2008)
• A deterministic map of Waddington’s epigenetic landscape for cell fate specification, S Bhattacharya et al. BMC Systems Biol. 2011, 5:85
• Quantifying the Waddington landscape and biological path for development and differentiation J Wang et al. PNAS 108 (20):8257-8262 (2011)
• Lorenz, E.N. Deterministic nonperiodic flow. J. Atm. Sci. 20:130-141 (1963)
• Wang, J. et al. 2010 the potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. Biophys J. 99, 29-39
• Li, Q. et al. Dynamics inside the cancer cell attractor reveal cell heterogeneity, limits of stability, and escape. PNAS. 113 (10) 2672-2677 (2016)
• Font-Clos, F. Topography of epithelial-mesenchymal plasticity. PNAS. 1-6 (2018)
• Y.I. Wolf et al., Physical foundations of biological complexity. PNAS, 2018
• Green, J.B.A and Sharpe, J. Positional information and reaction-diffusion: two big ideas in developmental biology combine. Development 2015 142: 1203-1211


- P Hou et al. Pluripotent stem cells induced from mouse somatic cells by small molecule compounds, Science. 341 (6146) (2013)
- Yuan, Jinzhou et al. “Single-cell transcriptome analysis of lineage diversity in high-grade glioma.” Genome medicine vol. 10,1 57. 2018
• Martinelli, Paola et al. “GATA6 regulates EMT and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer.” Gut vol. 66,9 (2017): 1665-1676
• F.J. Verweij et al., Quantifying exosome secretion from single cells reveals a modulatory role for GPCR signaling. JCB 2018
• P. Gangadaran et al., An update on in Vivo imaging of extracellular vesicles as drug delivery vehicles. 2018
• Kopper et al. An organoid platform for ovarian cancer captures intra- and interpatient heterogeneity. Nat Med. 2019
• T.R.J. Bossamaier and D.G. Green, *Complex Systems*. Cambridge Univ. Press, 2000
• J. N. Kather et al. Deep Learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med*. 2019


Gualerzi, Alice et al. “Raman spectroscopy uncovers biochemical tissue-related features of extracellular vesicles from mesenchymal stromal cells” Scientific reports vol. 7,1 9820. (2017)


Shin H, Correlation between Cancerous Exosomes and Protein Markers Based on Surface-Enhanced Raman Spectroscopy (SERS) and Principal Component Analysis (PCA). ACS Sens., 2018, 3 (12), pp 2637–2643


• J. Leray. Sur le mouvement dun liquide visqueux emplissant lespace, Acta Math. 63 (1934), 193248
• Nigel Goldenfeld, Roughness-Induced Critical Phenomena in a Turbulent Flow. PRL 96, 044503 (2006)
• V. I Arnold and BA Khesin, Topological methods in hydrodynamics. Springer 1991
• Thamilselvan V, Basson MD. Pressure activates colon cancer cell adhesion by inside-out focal adhesion complex and actin cytoskeletal signaling. Gastroenterology. 2004;126(1):8–18
• M. James, W.J.T. Bos, M. Wilczek, Turbulence and turbulent pattern formation in a minimal model for active fluids, Phys Rev. Fluids 3, 061101(Rapid Communication) (2018)
• M. James, M. Wilczek, Vortex dynamics and Lagrangian statistics in a model for active turbulence, EPJE 41, 21 (2018)
• Q Ouyang and H.L. Swinney, transition to chemical turbulence. Chaos, 1991. 1, 411
• Interface 11: 20140677 (2014)
• Frohlich, H. Long range coherence and energy storage in biological systems. Int. J Quantum
Chem. 2(5):641-649 (1968)
(1977)
• Lundholm, I.V. et al. Terahertz radiation induces non-thermal structural changes associated with
• E. Prodan & C. Prodan. Topological phonon modes and their role in dynamic instability of
• Del Giudice, E. et al. A collective dynamics in metabolically active cells. Physica Scripta,
26(3)(1982)
• Marletto, C. et al. Entanglement between living bacteria and quantized light witnessed by Rabi
• Bohm, D. A suggested interpretation of the Quantum Theory in terms of “hidden variables”. I.
Phys. Rev. 85(2):166-179 (1952)
• de Broglie, L. Interpretation of quantum mechanics by the double solution theory. Annales de la
Fondation Louis de Broglie, Volume 12, no.4, 1987
• Harvey, R.J. Navier-Stokes analog of Quantum mechanics. Phys Rev 152, 1115 (1966)
• Bohm, D. & Hiley, B. The Undivided Universe: An ontological interpretation of Quantum Theory
(Routledge, 1995)
• Andrew W. Baggaley et al. Vortex-Density Fluctuations, Energy Spectra, and Vortical Regions in
• Holland, P.R. The Quantum theory of motion: an account of de Broglie-Bohm causal
121:4857 (1983)
• Jungel, A. et al., A new derivation of the quantum Navier-Stokes equations in the Wigner-
One. 9(3): e89017 (2014)
• Daniel M. Harris and John W. M. Bush, The pilot-wave dynamics of walking droplets. Physics of
Fluids 25, 091112 (2013)
• Jan Moláček and John W. M. Bush. Drops walking on a vibrating bath: towards a hydrodynamic
Figure 1: Waddington landscape- The Waddington landscape shows a stem cell S bifurcating to various cell fates represented by the blue balls. As seen in red, the flows of gene expression underlying the differentiated cell states S1 and S2 seem more laminar. However, cell fates are reversible as indicated by the dotted line. Multiple bifurcation routes exist towards a local energy minimum X. The attractor X is a chaotic cell fate (i.e. cancer) and shown as a Lorenz-like strange attractor on the developmental landscape. Note the two-way (revertible) flow from the initial stem cell state to the cancer state.