**INTRODUCTION**

*Claudin-Low Triple Negative Breast Cancer*

Triple Negative Breast Cancer (TNBC) is an aggressive, heterogeneous subtype of breast cancer (BC) accounting for 20% of BC cases1. Unlike other subtypes of BC which are characterized by the presence of estrogen receptors, progesterone receptors, or HER2 amplification, TNBC is characterized by the lack of all three of these markers. This, coupled with the disease’s aggressive nature, makes it difficult to treat TNBC with current targeted therapies. Currently, over 80% of clinical trials studying TNBC therapeutics are exploring combinations of targets.

Prior to 2007, there were four intrinsic subtypes of breast cancer: Luminal A, Luminal B, HER2-enriched, and basal-like. As gene expression technologies evolved, a new subtype accounting for 7-14% of invasive breast cancer was classified and termed Claudin-Low (CL) [SOURCE]. This subtype is characterized by low to absent expression of tight junction proteins Claudin 3, 4, and 7, as well as the cell adhesion molecule E-Cadherin. Thus, CL tumors are prone to epithelial to mesenchymal transition (EMT) and exhibiting stem-cell like characteristics, two hallmarks of cancer associated with high metastasis rates2,5. Furthermore, CL tumors progress through the cell cycle slower than other BC subtypes, making them resistant to traditional cytotoxic chemotherapies that depend of rapid cell division for selectivity. Due to its lack of response to targeted and cytotoxic therapies, this aggressive disease has a poor prognosis and low survival rates. New therapeutic strategies must be explored for the treatment of CL TNBC.

*Tumor Reversion*

Conventional cancer therapies are design to induce apoptosis in cancerous cells. These therapies work by targeting rapidly dividing cells or malignant cells with characteristic cell-surface receptors. The limitations to this approach lie in that it is difficult to selectively choose which cells to kill, and ultimately normal cells are damaged in the process. Furthermore, not all tumors are rapidly dividing or present cell-surface receptors that can be targeted. An alternative therapeutic approach for treating these diseases is tumor reversion. Tumor reversion is the biological process by which tumor cells lose a significant fraction of their malignant phenotype (source). It is unlike conventional therapies in that instead of inducing apoptosis of cancerous cells, tumor reversion is aimed at reverting them to normal-like cells.

Tumor reversion has been observed sporadically for over a century. It was first observed by Boveri and Rous in 1907 when ovarian teratoma cells spontaneously differentiated into a normal somatic cell lineage (Cho, 2017). Since then, it has been achieved *in vitro, in vivo,* and *ex vivo.* In particular, tumor reversion has been achieved *in vitro* with the Claudin-Low Triple Negative Breast Cancer cell line MDA-MB-231, and *ex vivo* in mice xenografted with MDA-MB-231 cells (sources).

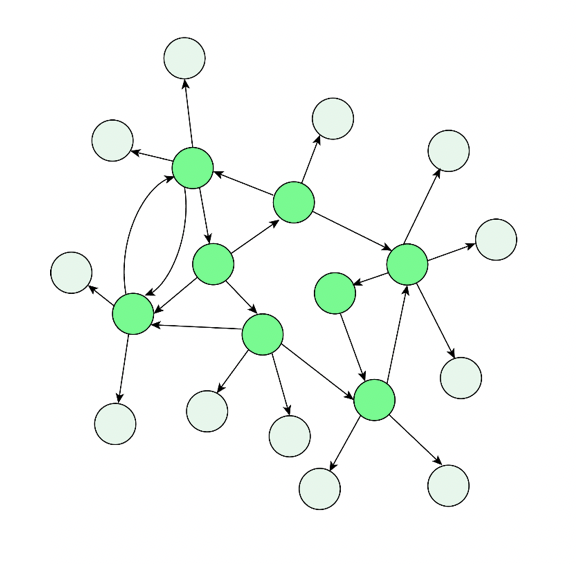
Tumor reversion is not merely the reversal of the development of a malignant cell. Instead, it is its own biological process involving cell reprogramming through epigenetic and genetic tools that supersede the tumorigenic characteristics of a cell. Molecular components of a cell such as transcription factors and master regulators can be targeted to induce the reversion of tumorigenic cells. These targets differ from tumor suppressor targets in that while both will suppress the cell’s malignant phenotype, only reversion targets will be able to trigger the shift from a malignant to a normal-like phenotype (Telerman, 2009). The field of systems biology is faced with the challenge of developing methods to systematically identify tumor reversion targets. This project takes a dynamical systems approach to do so.

*Cancer as a Developmental Disease*

So often is cancer reduced to its genetic component, that the larger picture of tumor cells in the context of development is lost. Normal cells, such as bloods cells and epithelial cells, differentiate into distinct phenotypes despite having identical genomes. The reason for their specificity has nothing to do with genetic mutations, but rather the microenvironment in which they are placed. The development of cancer can be viewed the same way. It has long been suggested that cancer cells are simply a cell type resulting from an error in the complex cell developmental processes that guide the differentiation of normal cells (Kauffman, 2009). In fact, tumors of the same type often have different mutational profiles, indicating that some other biological process is steering these tumors to the same malignant phenotype. Thus, the development of cancer can be seen as a systems-level dynamical process driven by an intracellular signaling network where signaling pathways have been dysregulated.

*Intracellular Signaling Networks and Waddington’s Epigenetic Landscape*

An intracellular signaling network is composed of molecular components of a cell (*nodes*) and their regulatory interactions (*edges*), and accurately represent large-scale signaling processes that occur within cells. Unlike mathematical models, which require knowledge of system dynamics such as kinetic parameters or logic formalisms, signaling networks can be constructed when this information is unknown or difficult to determine. They can be inferred from steady-state gene expression profiles at the genome-wide scale without bias and represent a cell-fate context (Cho, 2017).

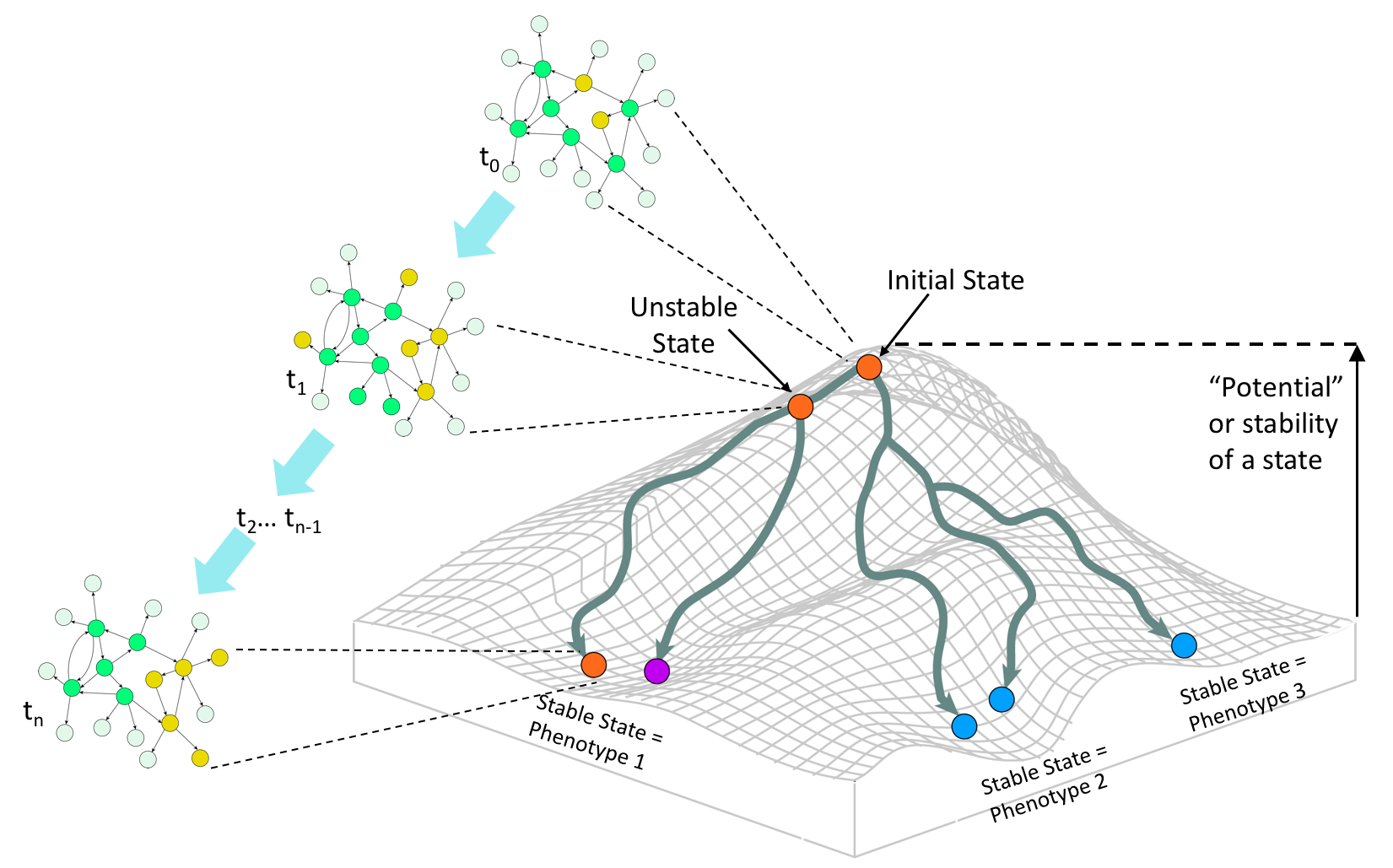
****

**Figure 1. Toy Example of Intracellular Signaling Network**

The state of the system evolves over time according to the regulatory relationships between each node until it reaches a steady state, known as an *attractor*. Each network state corresponds to a gene expression pattern of the cell, and each attractor corresponds to a gene expression profile associated with a cell fate (Kauffman). The state space of the intracellular signaling network is a high-dimensional space consisting of all possible states of the network. Attractor states are points in the state space characterized as “low energy states” at the bottom of a “potential well”. This potential well is the *basin of attraction*, and consists of unstable states that are close to the attractor in the state space and eventually lead to the attractor. An intracellular signaling network of N nodes has an N-dimensional state space, and cannot be visualized by the human eye. However, the use of Waddington’s Epigenetic Landscape can be used to visualize the difference in potential of network states.

Waddington’s Epigenetic Landscape was first developed in the 1950s as a metaphor for cell differentiation in developmental biology. However, it has since been formally explained as a generalized potential landscape for the state space of an intracellular signaling network, derived from its dynamics. The valleys of Waddington’s landscape each represent a cell phenotype, and correspond to basins of attraction of the intracellular signaling network. The attractor states are at the bottom of the valleys or basins of attraction. The height of a state in Waddington’s epigenetic landscape corresponds to how much “potential” the state has, or how unstable it is **(Figure 2b)**. Therefore, hills in the landscape correspond to unstable states separating attractors. In this framework, a switch between two cell phenotypes corresponds to the transition of a cell from one valley to another. It is important to note that the term “epigenetic” in Waddington’s epigenetic landscape does not refer to the molecular biology term used to describe chromatin modifications. Instead, it is more closely related to the notion of an “epigenetic state” from physics – a state that arises from genetic interactions (Kauffman, 2009).

**Figure 2. Waddington’s Epigenetic Landscape**



*Cancer Attractors*

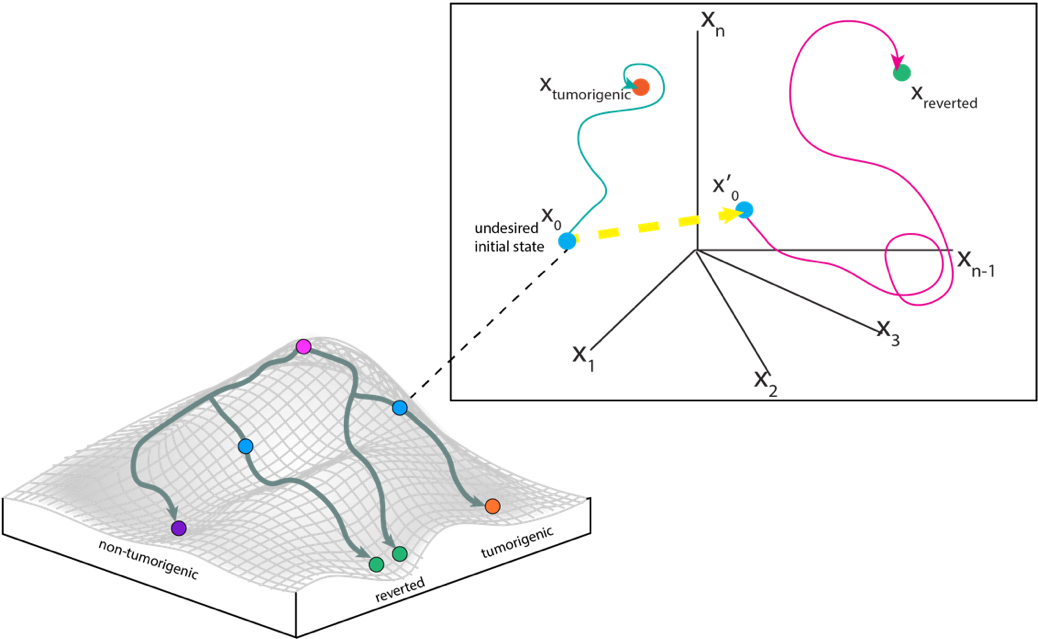
Since attractors correspond to cell types, and cancer is an abnormal cell type, then there must be attractors of the intracellular signaling network representing cancer cells. Kauffman described the cancer attractor in 1971 as a preexisting attractor of the intracellular signaling network that cannot be accessed by normal cells. This stems from the idea that an intracellular signaling network can have many attractors that are not associated with normal physiological cell types but rather a byproduct of its complex dynamics. Many of these attractors may correspond to non-viable gene expression profiles, but some may be associated with a viable, proliferative phenotype, representing cancer. Normal developmental trajectories steer cells away from these attractors, but when those processes are dysregulated, as they are in the development of cancer, cells can fall into the cancer attractors.

The introduction of the cancer attractor allows for the consideration of genetic mutations in the context of development. In a signaling network without genetic mutations, cancer attractors are not accessible. However, genetic mutations alter the structure of the intracellular signaling network by removing a node or link in the case of loss-of-function mutations, and strengthening or gaining additional links in the case of gain-of-function mutations. These changes in the intracellular signaling network alter the topography of the epigenetic landscape by lowering barriers to reach a cancer attractor, or in the case of a bifurcation event, creating a *de novo* cancer attractor. These mutations increase the probability that developing cells falls into a cancer attractor.

Transient non-genetic perturbations to the network state can move it from one site in the landscape to another, ultimately changing which basin of attraction it falls into. Alternatively, constitutively controlling target nodes can alter the dynamics of the intracellular network, and therefore the topography of the epigenetic landscape, making it possible for a cell to move from the cancerous to normal attractor. Thus, tumor reversion can be viewed as an optimal control problem in dynamical systems theory to identify targets that can trigger the shift out of the basin of the cancer attractor and towards that normal attractor. The challenge facing systems biologists is in determining ho

*Control Of Intracellular Signaling Networks*

Structure based control methods study the controllability of systems based solely on the structure of the network (SOURCES). Attractor based control methods aim to restrict the target states of a network to attractors. Structure-based attractor-based control methods aim to do both. Many structure based control methods are designed for linear systems or require dynamics to be linearized when applied to nonlinear systems. These approaches are not appropriate for applications to intracellular signaling networks, which are governed by nonlinear dynamics. Recently, a structure-based attractor-based control method for systems with non-linear dynamics, Feedback Vertex Set Control (FC), has been proposed. FC is a an extension of mathematical control theory applied to dynamical systems focused on the controllability of the system by restricting the target states to attractors.



The Feedback Vertex Set (FVS) of a network is a subset of nodes whose removal leaves the network without directed cycles. Mochizuki et al. mathematically proved that, for a network governed by non-linear dynamics, controlling the dynamics of the FVS is sufficient to drive the dynamics of the whole system to converge to any dynamical attractor. Overriding the state variables of the FVS nodes into the trajectory specified by a given dynamical attractor ensures that the network will approach that desired dynamical attractor.

Zanudo further expanded this framework to include source nodes (nodes with no incoming edges) as necessary control targets. Mochizuki and colleagues assume that source nodes converge to a unique trajectory and do not need independent control, but Zanudo asserts that source nodes can denote external stimuli a cell responds to, which could result in different attractors for source node states. The combination of the FVS and source nodes constitutes the FC set. Therefore, the identification of overrides of FC nodes that steer the network towards the normal associated attractor is equivalent to the identification of perturbations that should be applied to cancerous cells to steer them to a normal-like phenotype.

*Estimation of Network Dynamics*

In order to determine the effect of perturbations on the long-term behavior of a signaling network, an estimation of the system dynamics is necessary. Though signaling networks do not describe the temporal behavior of a system, dynamical systems based approaches for network analysis have recently been developed. One such approach is the estimation of network dynamics based on network structure. Lee and Cho developed an algorithm to estimate signal flow based on the network topology and predict network attractors from user provided initial conditions. The activity of a node depends on its basal state, the activity of the nodes regulating it at the previous time step, and the weight of the interaction between the node and its regulatory elements. This is defined in the signal propagation algorithm as:

$a\_i(t+1) = (\prod\_ja\_j(t)^{W\_{ij}})^\alpha a\_b(i)^{1-\alpha}$

Where a\_j(t) is the activity of node j regulating node I at the previous timestep, W\_ij is the weight of the interaction between the two nodes, a\_b(i) is the basal activity of node I, and \alpha is a hyper parameter between 0 and 1. The algorithm assumes that input stimulation does not change according to time. Taking the logarithm and using matrix notation produces:

$x(t+1) = \alpha W x(t) + (1-\alpha)b$ .

Where x is log(a), b is log(a\_b), and W is a weight matrix. This equation can be solved iteratively until it convergences at some tolerance level, or the steady state solution of the equation can be solved with:

$x\_s = (1-\alpha)(I-\alpha W)^{-1}b$.

The link weight matrix, W, is defined as:

$W = D\_{in}^{-1/2}AD\_{out}^{-1/2}$

$(D\_{in})\_{ii} = \sum\_j\abs{A\_{ij}}$ and $(D\_out)\_{jj} = \sum\_j\abs{A\_{ij}}$

Where A is the adjacency matrix of the network, and D\_{in} and D\_{out} are diagonal matrices consisting of the indegree and out degree for each node, respectively.

To compare the effects of multiple perturbations, the difference between log steady state attractor values can be calculated as

$x^{fold} = x^{c\_1} – x^{c^2}$

Where x^c\_i denotes the log steady state activity for each condition. Therefore, x^{fold} is essentially a log fold change. The sign (positive or negative) of x^{fold} denotes if the activity is up or downregulated in the compared conditions, rather than predicting the accurate amount of change.