**Supplemental Text**

**Gene Set Enrichment Analysis (GSEA).**

GSEA compares gene expression in biological samples by ranking the genes that are differentially expressed between two or more phenotypes. A running score is calculated by moving down the ranked gene list and increasing the score when a gene in the ranked list is in the enriched gene set. The score is decreased when a gene is not in the list. The enrichment score represents the maximum deviation from zero encountered while walking the gene list. A positive enrichment score means that the phenotype is over-represented in the gene list and a negative enrichment score means the phenotype is under-represented in the gene list. Statistical significance is determined using an empirical phenotype-based procedure, which produces a p-value. A database of established gene sets exists in the Molecular Signatures Database 4.0 (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>).

**Network construction.**

Our strategy was the following: we knew that TGF is a prominent signal driving epithelial-to-mesenchymal transition in hepatocellular carcinoma, so we started with synthesizing a network for this process from the literature. Then we added the other known inducers of EMT and other processes and components that are known to be deregulated during EMT. This resulted in a network of 70 nodes and 135 edges, which we call the EMT network (Figure 3 and Figure 1). Supplemental Table 2 indicates the edges of this network, giving the source (upstream) node, the target (downstream) node, whether the edge represents an activating (positive) or inhibitory (negative) relationship, whether the edge corresponds to a direct physical or chemical interaction, and references where the relationship was reported.

*The destruction complex:*

The “destruction complex” (represented by the Dest\_compl node) is formed by APC, AXIN1 or AXIN2, and GSK3β. The APC protein, a constituent of the destruction complex, is not included as an individual node in the EMT network, as there is no compelling evidence of its differential regulation during the EMT process. This is equivalent with assuming that it is constitutively present. We use a single AXIN node, AXIN2, as there is strong evidence that suggests that AXIN1 and AXIN2 are functionally equivalent; however, they are differentially expressed and AXIN2 is transcriptionally induced by Wnt signaling through β-catenin and the TCF/LEF transcription factors(1).

**Boolean dynamic modeling.**

In Boolean models, each network node can be described by one of two qualitative states: ON or OFF. The ON state means an above threshold concentration of a molecular regulator whereas the OFF state means the below-threshold or inactivated form. The biological relationships defined by the network can be translated into mathematical equations using Boolean operators(2). Accordingly, we described each component in the network by a regulatory function that uses the Boolean logical operators OR, AND, and NOT and that reflects the known combinatorial effects of the regulators. Specifically, all rules involving multiple regulators were constructed with OR qualifiers, assuming that all regulators act independently, unless there was evidence of synergy or conditionality, in which case the regulators were connected by an AND qualifier. Here we explain the rules of three key nodes (TGF, E-cadherin and Destr\_compl) that are important for understanding the paper and to clarify our modeling methodology.

*The TGF rule* is as follows:

TGFβ\* = Goosecoid or SNAI1 or TWIST1 or GLI

When thinking about Boolean rules, it might be more intuitive if instead of ON or OFF, we consider each node as having the logic states TRUE or FALSE, respectively. According to this update rule, if all inputs are OFF (FALSE), then TGFβ will turn OFF (FALSE):

 TGFβ\* = FALSE or FALSE or FALSE or FALSE = FALSE

Now we consider the case when a single regulator, SNAI1, is ON (TRUE) at a previous time step:

 TGFβ\* = FALSE or TRUE or FALSE or FALSE = TRUE

Thus the TGF node will be ON (active) if any of its individual inputs is ON and it will be inactive if all of its inputs are OFF.

*The E-cadherin rule* is as follows:

E-cadherin\* = β-catenin\_memb and (not SNAI1 or not HEY1 or not ZEB1 or not ZEB2 or not FOXC2 or not TWIST1 or not SNAI2)

According to this rule, E-cadherin will turn OFF if β-catenin\_memb is turned OFF or all of the 7 negative regulators are turned ON. First, let’s demonstrate what happens when β-catenin\_memb is OFF (FALSE).

E-cadherin\* = FALSE and (not SNAI1 or not HEY1 or not ZEB1 or not ZEB2 or not FOXC2 or not TWIST1 or not SNAI2)

“AND” rules logically require both states to be TRUE in order to produce a TRUE output. Thus, regardless of the state of the nodes in the parentheses, if β-catenin\_memb is OFF (FALSE), E-cadherin will be OFF. E-cadherin will also be OFF, if all of the 7 negative afferents are ON, regardless of the state of β-catenin\_memb. If SNAI1 is ON (TRUE), the expression “not SNAI1” is FALSE and similarly for all negative regulators of E-cadherin. In this case the E-cadherin rule is

E-cadherin\* = β-catenin\_memb and (FALSE or FALSE or FALSE or FALSE or FALSE or FALSE or FALSE),

This logically evaluates to FALSE, meaning that E-cadherin turns OFF.

If all 7 negative regulators are OFF and β-catenin\_memb is ON, then E-cadherin will be ON:

E-cadherin\* = ΤRUΕ and (TRUE or TRUE or TRUE or TRUE or TRUE or TRUE or TRUE) = TRUE

In fact, if any one of the negative regulators of E-cadherin is OFF, the parenthesis is equivalent with “FALSE or TRUE”, which evaluates to TRUE. Thus E-cadherin will be ON if β-catenin\_memb is ON and any of the 7 negative afferents are OFF.

*The Dest\_compl(destruction complex) rule* is as follows:

Dest\_compl\* = (GSK3 and AXIN2 and -catenin\_nuc) or (GSK3 and Dest\_compl)

According to this rule, the initial formation (ON state) of the complex requires GSK3, AXIN2 and -catenin\_nuc. However if the Dest\_compl is already present (ON), which it is in the epithelial state, then it will remain sufficiently stable even if AXIN2 and -catenin\_nuc are present at a below-threshold level (OFF). The only additional node that needs to be ON for the destruction complex to stay ON is GSK3.

*Boolean rule over-rides:* Biological dysregulations are simulated by forcing a node to be ON (positive dysregulation) or OFF (negative dysregulation). Positive dysregulation is analogous to over-expression, over-abundance, or mutational activation of a protein and negative dysregulation is analogous to a knockout, down-regulation, or mutational inactivation of a protein. To model the effect of an over-active TGFβ signal we use a positive over-ride in the simulation results shown in Figure 4, in the state transition network in Figure 5B, and to produce the motifs in Figure 7 and Supplemental Figure 3. We model the inhibition of SMAD by a negative over-ride, in the presence of an over-active TGFβ signal, to produce the simulation results in Figure 7B and the network motifs in Figure 7C.

**Asynchronous updating algorithm and node activity.**

To account for different timescales in signaling networks, asynchronous models may be used in which the state of each node is updated according to its own timescale. While we do not know the kinetics of individual events, we do know that signal transduction events occur substantially faster (with a timescale of seconds or faster) than transcriptional events (which have a timescale of minutes)(3). Thus, we apply a general asynchronous updating scheme with a ranking scheme. In the general asynchronous update, a single randomly chosen node is updated (i.e. its state is re-evaluated) at each time instant. It is possible that the same node chosen in two consecutive time instants.

In order to account for signal transduction events occurring faster than transcriptional events, we incorporate a ranking system. We update nodes regulated by signal transduction events with a greater probability than nodes regulated by transcriptional events. We tested probability values that differed by more than an order of magnitude and determined that the quantitative value of the ratio between the timescale of the slow and fast processes does not matter as long as it is significantly larger than one (Supplemental Figure 4). We define a time step as the average number of updates needed to update a slow (transcriptionally regulated) node. This time step thus corresponds to several minutes in real time. Altering the timing does not affect the possible steady states, nor does it affect the initial conditions that lead to these steady states. It does, however, alter the probability of specific trajectories and thus the relative rates at which nodes are affected by a signal in our network.

Because the ranked general asynchronous update is a stochastic process, different simulations that start from the same initial state can reach the different outcomes (steady states). Thus, a large number (10,000) of replicate simulations are performed and the average state of each node is calculated. We call the average state of a node the node activity. If a node’s activity over a large number of simulations stabilizes at 1, it means that each simulation led to a sustained ON state of the node. Conversely, stabilized node activity of 0 means that the node stabilized at OFF in each simulation. Node activities between 0 and 1 represent the fraction of simulations in which a node was ON. When running computer simulations it is important to run as many simulations as possible; however there are time constraints that arise when increasing the simulation size. At the same time, there is a diminishing return for doing more simulations. 10,000 simulations and 20 timesteps was deemed sufficient for most simulations explored in the EMT network, as averaged values over runs from these simulations produced consistent results with only small deviations.

**State space analysis.**

Biological systems like the EMT network can be modeled as dynamic systems; these models can explore the full range of possibilities in terms of outcomes (e.g., steady states) the system can reach and ways in which these outcomes can be reached (trajectories). State space analysis entails exploring every possible state of a network of size N (2N states) and how these states can lead to other system states, as governed by the Boolean rules and updating scheme defining the system. All sequences of consecutive states will finally reach an attractor (a steady state or a small set of states) that the network cannot leave. The set of system states directly reaching an attractor are called the attractor’s basin of attraction.

**Network reduction.**

Because the size of the state space in Boolean models grows exponentially (2N, where N = the number of nodes in the network), it becomes impossible to map all the trajectories in large networks (e.g. 270 >1020 possible states in the EMT network). Network reduction techniques have been developed to reduce the size of the network without limiting its number of outcomes. We used a two-step network reduction method. The first step is to identify and eliminate stabilized nodes. Because the focus of our work is TGF induction of EMT, all other signals (e.g. EGF) were set as OFF. Signal nodes unaffected by crosstalk (e.g. EGF, HGF) maintain their OFF states, and the effect of this OFF state iteratively fixes the state of certain downstream nodes. These signals and their fixed-state downstream nodes were identified analytically and removed as their effect was not targeted by TGF signaling. Our second network reduction step is to iteratively collapse mediator nodes that have one regulator and in turn regulate a single node. This method conserves the attractors of the system(4-6). In these cases the node directly upstream of the removed node was connected to the node directly downstream of the reduced node, and the state of the upstream regulator was inserted in the downstream-regulated node’s Boolean update rule. Thus using these reduction techniques we can reduce the size of the network from 70 nodes and 135 edges to 19 nodes and 70 edges and perform a state space analysis on a smaller network (Reduced EMT dynamic model, Figure 5 and Figure 1) that is representative of the full EMT network. This network reduction was shown to have no effect on the permitted dynamic behaviors (e.g. steady states) of a system(4)*.*

**Stable motif analysis.**

We used stable motif analysis to identify key feedback loops regulating EMT. A stable motif of a Boolean model consists of a set of nodes and their respective states which satisfy two characteristics: 1) the nodes form a strongly connected component (in which each node is reachable from any other node) in a directed network representation of the Boolean model, and 2) the specified states are steady states of the nodes of the stable motif, independent of the state of any node outside of the stable motif(7). These two characteristics have a topological and dynamical interpretation. Topologically, they imply that the nodes of the stable motif form a type of positive feedback loop. Dynamically, they imply that the states associated to the stable motif act as a checkpoint or point of no return in the dynamics of the network. Previous work has shown that an iterative algorithm which finds the stable motifs and uses the influence of these motifs on the rest of the network can be used to find all the attractors of asynchronous Boolean models(7).

**References**

1. Dao DY, Yang X, Chen D, Zuscik M, O'Keefe RJ. Axin1 and Axin2 are regulated by TGF- and mediate cross-talk between TGF- and Wnt signaling pathways. Ann N Y Acad Sci. 2007;1116:82-99.

2. Saadatpour A, Albert R. Boolean modeling of biological regulatory networks: a methodology tutorial. Methods. 2013;62:3-12.

3. Papin JA, Hunter T, Palsson BO, Subramaniam S. Reconstruction of cellular signalling networks and analysis of their properties. Nature reviews Molecular cell biology. 2005;6:99-111.

4. Saadatpour A, Albert, R., Reluga, T. A Reduction Method for Boolean Network Models Proven to Conserve Attractors SIAM J Appl Dyn Syst. 2013;12:1997-2011.

5. Saadatpour A, Albert I, Albert R. Attractor analysis of asynchronous Boolean models of signal transduction networks. J Theor Biol. 2010;266:641-56.

6. Saadatpour A, Wang RS, Liao A, Liu X, Loughran TP, Albert I, et al. Dynamical and structural analysis of a T cell survival network identifies novel candidate therapeutic targets for large granular lymphocyte leukemia. PLoS Comput Biol. 2011;7:e1002267.

7. Zañudo JGT, Albert R. An effective network reduction approach to find the dynamical repertoire of discrete dynamic networks. Chaos. 2013;23:025111.