TGF-β-Induced Epithelial-to-Mesenchymal Transition

Bryan Félix, Abbigael VanDusen, and Elizabeth Wilson

December 20, 2017

Epithelial-to-mesenchymal transition (EMT) is the cancer-associated process wherein an epithelial cell converts into a mesenchymal type cell. Mesenchymal cells lack apical polarity and secrete extracellular-matrix-degrading enzymes that provide them with enhanced migration capabilities. Moreover, EMT is closely related to cancer stem cell-like phenotype mutations, such as chemotherapy resistance and enhanced survival when deprived of nutrients. Thus, it should not come as a surprise that EMT has been linked as a promoter of cancer metastasis and is of clear medical relevance. Multiple factors may activate an EMT state in a cell; one of the most successful promoters is the transforming growth factor-β (TGF-β). Presence of TGF-β induces several changes in transcription and translation regulation within the cell. In particular, the production of the transcription factor SNAIL1 is upregulated upon increased TGF-β levels; SNAIL1, in turn, influences levels of transcription factors ZEB1, TWIST, and SNAIL2 (SLUG), among others. The overall effects of TGF-β tend towards increased repression of E-cadherin, a key marker of epithelial cells, and upregulation of N-cadherin, a key marker of mesenchymal cells.

As Tian, et al., have reported a model investigating the effects of these transcription factors on E-cadherin and N-cadherin levels (2013), this report aims to further this model by incorporating the effects of TWIST and SLUG transcription factors. In addition to influencing EMT through regulation of E-cadherin and N-cadherin, TWIST has also been shown to upregulate the production of endogenous TGF-β, further driving EMT. Additionally, TWIST increases SLUG production, which competes with the negative feedback system of ZEB1. By examining the influence of these two additional transcription factors on the previously reported system, this study will provide an interesting insight into how cancer is able to progress through EMT and shed light on therapeutic procedures relevant to cell migration, such as cancer metastasis.
1 Introduction

Epithelial-to-mesenchymal transition (EMT) is a biological process that is normally involved in embryonic development and wound healing. Epithelium and mesenchyme are basic tissue phenotypes in mammalian cells, with epithelium allowing for lining of various cavities and mesenchyme providing a more fluid and migratory lining. Therefore, EMT is highly important during embryonic development, where cells need to be able to lose cell attachment and gain motility in order to spread to other parts of the embryo during development [1]. EMT is also another important process during wound healing. Cells that border the wound undergo EMT to gain motility so the wound can heal. Cells then experience re-epithelialization to close the wound [2]. However, this transition also plays a crucial role in cancer metastasis. During this transition, cells lose their epithelial qualities, such as polarity, cell adhesion, and interaction with the basement membrane, and gain mesenchymal qualities such as spindle-like morphology and enhanced migration and invasiveness [3]. Mesenchymal cells lack apical polarity, lose intercellular tight junctions, release extracellular-matrix-degrading enzymes, and even have elevated resistance to apoptosis [4]. All of these traits promote migration of cancer cells, leading to cancer metastasis and therefore cancer progression.

There are a variety of pathways that contribute to EMT. Pathways such as transforming growth factor-β (TGF-β) and those involving tyrosine kinase receptors, ERK, MAPK, PI3K/AKT, and JNK can all contribute to EMT progression by propagating signals from exogenous growth factors [5, 6]. TGF-β is one of the most studied EMT inducers, playing an interesting role as both a growth suppressor of normal epithelial cells and a promoter of cancer metastasis [7]. TGF-β coordinates the first steps in the signaling propagation pathway. This growth factor binds to the TGF-β receptor which causes phosphorylation of Smads in the cytoplasm. These Smads then form complexes with each other and translocate to the nucleus, where they interact with and activate transcription factors such as ZEB1, SNAIL1, SNAIL2 (SLUG), and TWIST1 [4, 8]. TGF-β triggers EMT through a regulatory network involving these transcription factors, as well as microRNAs miR-34 and miR-200 [9]. There are several feedback loops that contribute to the core regulatory network of EMT and the expression of epithelial and mesenchymal markers. There is also evidence that a partial EMT state exists, expressing a epithelial-mesenchymal hybrid where cells express both epithelial and mesenchymal properties [10, 11]. This means that EMT is involved in a three-step transition process, requiring an intermediate or partial state between the epithelial and mesenchymal phenotypes [9].

This paper focuses on the transcription factor dynamics that follow Smad signaling; therefore, Smad signaling proteins such as Smad2, Smad3, and Smad4 and their nuclear-cytoplasmic dynamics will not be discussed further [12]. Further descriptions of current models will focus on transcription factor protein dynamics with miRNA repressors, which occurs post-Smad activation.

Currently there are two competing mechanisms of the TGF-β-induced EMT pathway: the Cascading Bistable Switches (CBS) model and the Ternary Chimera Switch (TCS) model. The Cascading Bistable Switches model, developed by Tian et al. [13], includes an autocrine negative feedback loop between TGF-β, ZEB, and miR-200, while the Ternary Chimera Switch model, proposed by Lu et al. [14], includes self-activation of ZEB but does not contain the negative feedback of miR-200 onto TGF-β. However, there is experimental
evidence for both proposed models. Analysis of these pathways in different cell types may contribute to the varying experimental data in the literature because TGF-β is expressed in a variety of cell types throughout the human body. Moreover, the two models have different roles for SNAIL1, ZEB1, miR-34, and miR-200.

More specifically, the CBS model depicts two SNAIL1 populations with changing TGF-β concentration, with miRNA levels acting as the switch. When both miRNAs are present at high concentrations, SNAIL1 is low, and cells are in the epithelial state, while high SNAIL1 concentration results in either a hybrid or mesenchymal phenotype [3, 13]. Determination between partial and mesenchymal states relies on the levels of miR-200, ZEB, and TGF-β at low concentration of miR-24. When miR-200 levels are high and the other two hold low concentrations, the cells are able to exist in the partial state, as miR-200 represses both ZEB and TGF-β levels [13]. Similarly, when miR-200 levels are low, repression is removed and ZEB and TGF-β levels are high and EMT is able to propagate.

The TCS model, on the other hand, does not act as a binary switch but instead forms a stable SNAIL1/miR-34 loop module that predicts that SNAIL1 smoothly increases with TGF-β and can only define one population of cells [14]. ZEB1 is the key player in this model: differing ZEB1 concentrations at low, medium, or high levels denote epithelial, intermediate, and mesenchymal phenotypes, respectively, as predicted by the model. Ultimately, both models start with an input of TGF-β stimulation, and both the CBS and TCS models give an output of three stable equilibria which represent the epithelial, intermediate, and mesenchymal phenotypes [9].

Furthermore, in both the CBS and TCS models, there are two types of TGF-β in the system: endogenous and exogenous TGF-β. miR-200, as mentioned previously, has a double negative feedback loop with ZEB1, but it also represses endogenous TGF-β expression. Here, endogenous refers to the internal TGF-β concentration in the same cell and can therefore be influenced by other factors in the system. Exogenous refers to the TGF-β concentration outside of the cell, independent of other factors within the system. Biologically, this may represent a mutation in the TGF-β receptor that causes propagation of a larger-than-normal signal downstream through Smad proteins, upregulating expression of transcription factors. Exogenous TGF-β may also affect the system through an upregulation of TGF-β from other cancer cells, further inducing a greater response onto the cell.

Based off the literature, this paper focuses on the improvement of the CBS model developed by Tian et al. [13]. Currently, this model involves the upregulation of SNAIL1 by TGF-β, which in turn upregulates ZEB1 expression. Both contribute to an upregulation of mesenchymal cell makers and a downregulation of epithelial cell markers. This model also includes two double negative feedback loops, one between SNAIL1 and miR-34, where SNAIL1 can bind and inhibit miR-34 and miR-34 can bind to the mRNA transcript of SNAIL1 to inhibit its expression, and the other double negative feedback loop between ZEB1 and miR-200. miR-200 also inhibits endogenous TGF-β. As EMT progresses through cells, epithelial markers such as laminin-1 and entactin are downregulated, while mesenchymal markers are upregulated [17]. Such mesenchymal markers include vimentin and S100A4, which is also known as fibroblast-specific protein 1. EMT is also associated with an overexpression of β-catenin and a decreased expression of E-cadherin [18]. E-cadherin is indicative of cell adhesion. Cells undergoing EMT can also have an increased expression of N-cadherin, which has been shown to be associated with metastasis.

Although SNAIL1 and ZEB1 are key transcription factors, many other transcrip-
tion factors exist. While some are more prominent than others, TWIST1 is an important regulatory transcription factor that represses and activates certain components in the network. TWIST1 acts through the transcription factor SLUG to induce EMT, specifically through mediating the transcription of SLUG [15]. SLUG, also known as SNAIL2, is a paralog to SNAIL1. This transcription factor, like ZEB1, inhibits and is inhibited by miR-200 [14, 16]. Lastly, SLUG contributes to the crucial direct inhibition of epithelial phenotypes and upregulation of mesenchymal phenotypes that drives EMT. This report accounts for the addition of TWIST1 and SLUG, including the effects of inhibition and upregulation on their transcription and translation. A scheme of the modeled regulatory network is shown in the next section, Model Description. For simplicity’s sake, TWIST1 will further be referred to as TWIST, SNAIL1 as SNAIL, SNAIL2 as SLUG, and ZEB1 as ZEB. The main outputs of this model are the levels of E-cadherin and N-cadherin that correspond to epithelial markers and mesenchymal markers, respectively.

The modeling of TGF-β-associated EMT is complexed with crosstalk coming from various transcription factors and even other pathways such as the ERK and Wnt pathways [19]. The extension of the core regulatory network of TGF-β-associated EMT can provide new insight into EMT regulation. By incorporating transcription factors like TWIST and SLUG, the core regulatory network is further expanded. With these additions, the model should allow for the three stable equilibria shown on a bifurcation diagram of TGF-β against mesenchymal markers. Using the proposed model, the dynamics of the induced EMT can be analyzed, answering fundamental questions: How can the generated bifurcation diagram shift to the right to maximize the TGF-β needed to enforce this transition? What molecule can therapeutically accomplish this? At low TGF-β concentrations, the cell exists in the epithelial phenotype, but at higher TGF-β concentrations, the cell exists in the mesenchymal phenotype. If the stability of the bifurcation can shift to the right, this means that at higher TGF-β concentrations, the cell phenotype can still exist in the epithelial state. Through computational modeling, questions like these can begin to be answered, providing insightful information on how EMT, and therefore metastasis, can be minimized in cancer patients. Furthermore, understanding the mechanisms of how cells decide whether or not to undergo EMT is critical in improving the development of cancer treatments [14].

Computational biology has become a widespread and valuable tool in understanding cellular processes and pathways. By incorporating experimental data, mathematical models can provide a more detailed understanding of biological processes and transitions like EMT. These models are fundamental in achieving a quantitative comprehension in cancer biology and help predict experimental outcomes. This work is motivated by experimental data of SNAIL, ZEB, SLUG, and TWIST presented in the literature. The dynamic behaviors of these factors in response to exogenous TGF-β will be analyzed and presented, including bifurcation diagrams and concentrations over time. In the remaining sections, a description of the mathematical models and the key equations depicting the transcription factor dynamics of the TGF-β-induced EMT network are provided, followed by results and discussion of model implications.

2 Model Description

The proposed model is based on the model described by Tian et al., with the addition of the transcription and translation of TWIST and SLUG [13]. Figure 1 below provides a
simplified diagram of the TGF-β-induced EMT regulatory network. This cartoon schematic is simplified, showing miR-34 and miR-200 directly inhibiting SNAIL, ZEB, and SLUG protein; while in reality, these microRNAs are binding and inhibiting translation of the mRNA transcripts of SNAIL, ZEB, and SLUG rather than the protein itself.

Endogenous and exogenous TGF-β begin the pathway of EMT, promoting the transcription of snail mRNA, which is then translated to SNAIL protein. Lowercase writing denotes mRNA transcripts, while uppercase denote proteins. SNAIL, a transcription factor, enhances the transcription of zeb mRNA, which is then translated to ZEB. There is a double negative feedback loop between transcription factor SNAIL and miR-34 [20]. SNAIL inhibits transcription of miR-34, which in turn represses translation of SNAIL through binding of the microRNA to the mRNA transcript. Similarly, another double negative feedback loop exists between the transcription factor ZEB and miR-200. ZEB represses the expression of miR-200, while miR-200 represses the translation of ZEB [21]. The expression of endogenous TGF-β is also inhibited by miR-200 [13]. In the presence of a mutation in TGF-β, such as in the case of cancer, SNAIL and ZEB work together to upregulate mesenchymal markers like N-cadherin and to repress expression of epithelial markers such as E-cadherin. This ultimately causes the cell to exhibit a mesenchymal phenotype. Thus far, these conditions depict the model developed by Tian et al. [13].

Figure 1: Diagram of TGF-β-induced EMT regulatory network. TGF-β enhances the expression of SNAIL, which promotes ZEB expression. There are three double negative feedback loops involving SNAIL and miR-34, ZEB and miR-200, and SLUG and miR-200. TWIST allows for upregulation of endogenous TGF-/beta and acts through SLUG to provide enhanced expression of N-cadherin. The epithelial marker, E-cadherin, is inhibited by SNAIL, ZEB, and SLUG, while the mesenchymal marker, N-cadherin, is enhanced by SNAIL, ZEB, and SLUG.

The addition of TWIST and SLUG makes this regulation network unique. SNAIL enhances transcription of the twist mRNA, which is translated to create transcription factor, TWIST. TWIST upregulates the transcription of the slug mRNA, which is translated to
transcription factor SLUG [22]. TWIST also produces a positive feedback loop, upregulating expression of endogenous TGF-β and working against the inhibition provided by miR-200 [23]. Like ZEB, SLUG is involved in a double negative feedback loop with miR-200, repressing miR-200 expression, while miR-200 binds to the slug transcript to repress the translation of SLUG. Through upregulation of TWIST, SLUG acts with SNAIL and ZEB to cause an increased expression of N-cadherin and repression of E-cadherin [15].

TWIST needs SLUG to mediate this response. It has been confirmed that E-cadherin is inhibited by SNAIL, ZEB, and SLUG through repression of the E-cadherin gene, CDH1 [3]. Similarly, studies report the mesenchymal marker, N-cadherin, is promoted by SNAIL, ZEB, and SLUG through activation of the gene CDH2 [3]. The reported effect of TWIST on E-cadherin and N-cadherin, however, has been conflicting, with reports suggesting one of three outcomes: (i) direct downregulation of E-cadherin and upregulation of N-cadherin, (ii) action through either upregulation or downregulation but not both, or (iii) no direct effect on expression, but rather causing action only through other transcription factors. This model assumes the last case to be true with TWIST acting indirectly through SLUG; however, further research should be performed to confirm the role of this transcription factor.

A set of ordinary differential equations was developed to describe the dynamics of the transcription factors, mRNA transcripts, and miRNAs. The following thirteen ordinary differential equations were used to model the dynamics of the core TGF-β-induced EMT regulatory network in MATLAB. A definition of the variables and their values are provided in Table 1 in the appendix, Section 6, and the rules used to justify these parameter values are explained in Section 6.2. The corresponding MATLAB code is provided in Section 6.3. Lowercase snail, zeb, twist, slug denote the mRNAs that are translated to the transcription factors SNAIL, ZEB, TWIST, and SLUG, respectively. Regulation of transcription is modeled using the Hill function, which is common in transcription regulation [24]. Consistent with reports in literature, a value of 2 was used for the Hill coefficients to represent the cooperativity and nonlinearity between genes and protein [13]. Models of mRNA transcripts and miRNA contained terms for basal transcription rate, as mRNA and miRNA are constitutively produced to some extent, regardless of inhibition or lack of upregulating proteins. Inhibition is represented in the denominator of production terms at a given Michaelis rate. Upregulation is represented in the production terms, with the promoting molecules and their respective Michaelis constants appearing in both the numerator and denominator. All molecules were degraded at constant rates, independent of activation or inhibition.

$$\frac{d}{dt}[TGF_\beta] = k_0T + k_T \frac{[TWIST]^2}{M_{TW}^2} \left( 1 + \left( \frac{[mR200]}{M_{200}} \right)^2 \right) - k_{degT}[TGF_\beta]$$ (1)

$$\frac{d}{dt}[snail] = k_0sn + k_{sn} \frac{[TGF_\beta + T_0]^2}{M_{sn}^2} \left( 1 + \left( \frac{[mR34]}{M_{34}} \right)^2 \right) - k_{degsn}[snail]$$ (2)

$$\frac{d}{dt}[SNAIL] = k_S[snail] \frac{1}{1 + \left( \frac{[mR34]}{M_{34}} \right)^2} - k_{degS}[SNAIL]$$ (3)
In attempt to simplify the model, the transcription and translation equations of the transcription factor proteins were combined, accounting for the inhibition by miRNAs and upregulation by other transcription factors. This was done by moving the inhibition terms to the differential equations for the transcription factors and incorporating the mRNA production and degradation rates into the protein production rates [25]. This simplification would allow for fewer equations and fewer parameters. However, doing so removed the possibility of a partial epithelial-mesenchymal state, a phenomenon repeatedly confirmed in literature. Additionally, the transition from epithelial state to mesenchymal state could not be observed at the reported time point of around 100 hours [26]. These issues were attributed to the assumption that the mRNA transcripts existed constant at a steady-state value without inhibition or upregulation. Therefore, it became clear the mRNA transcripts are essential to create the three steady-states and reflect the nature of the system. The quantity of mRNA
transcripts must be able to vary over time to create the double negative feedback loops between the transcription factors and the miRNAs.

The initial conditions and the values of production and degradation rate constants are provided in the appendix (Section 6). If a rate constant could not be identified, its value was varied around a value of a similar molecule to match the reported characteristics of the system. Using these thirteen differential equations, the exogenous TGF-β input was varied in order to assess the effects on the epithelial and mesenchymal phenotypes. Specifically, the effect of TGF-β on N-cadherin and E-cadherin concentrations were plotted to provide a quantitative depiction of concentration on pathway outputs. Following this initial evaluation of the differential equations, the effect of the various transcription factors were analyzed to determine the importance of each. Plots were generated to demonstrate the effect of the various transcription factors, particularly TWIST and SLUG, on EMT. Lastly, bifurcation diagrams were generated to visualize the progress of EMT, including its epithelial, partial, and mesenchymal states, as well as to assess whether any parameters largely affected the system. These results will provide interesting insight into concentrations of proteins necessary to induce partial epithelial/mesenchymal traits as compared to full mesenchymal traits. This also allows for further understanding of how treatment can influence the changes between these phenotypes, which is necessary to prevent metastasis and cancer progression.

3 Model Results

This modeling system was a deterministic model that provided information of the dynamics of the process, specifically how each variable affects the outputs. The first focus of the model results was the production of three steady-state values of the system, verifying the model accurately reflects the reported system. Epithelial (high E-cadherin levels), mesenchymal (high N-cadherin levels), and partial states have been observed experimentally, and are therefore necessary in the model results as well. The concentrations of E-cadherin and N-cadherin were compared over time with varying endogenous TGF-β (T0) concentrations; the results are shown in Figure 2 below.

Consistent with Tian et al, T0 represents a unitless value, providing insights into general trends of varying endogenous TGF-β concentrations. When T0 is zero, E-cadherin expression is very high, and N-cadherin expression is zero, representing the epithelial state. When T0 holds a slightly greater value of 1.8, N-cadherin expression becomes greater than E-cadherin expression; although, there are still significant levels of E-cadherin expression. This is the partial state, where the cell phenotypically express both epithelial and mesenchymal markers. When T0 is 5, N-cadherin expression is further increased, and E-cadherin expression is further decreased, approaching a value of 0 µM. This represents the full mesenchymal state where the cell has become more spindle-like and holds a higher ability to migrate. These results confirm that there are three steady-states within our model and that increasing levels of endogenous TGF-β promote the transition between these states.

With the addition of TWIST and SLUG, the dynamics of the modeled system are upheld and create three steady states. Moving forward, the next focus was the effect of each of the transcription factor concentrations over time. Concentrations of the four transcription factors over time are shown in Figure 3. It is immediately notable that SNAIL increases the fastest in the first 100 time steps of the simulation. ZEB, TWIST, and SLUG all first lag before less drastic increases in concentration. From this, SNAIL appears to be the main
inducer of EMT, while ZEB, TWIST, and SLUG each help sustain the transition to the mesenchymal state. To further support this claim, the SNAIL production rate was set to zero. When SNAIL protein is not produced, the system does not undergo EMT, and the epithelial characteristics remain (data not shown). However, if ZEB, TWIST, or SLUG is set to a production rate of zero, this simplified system proceeds toward a partial EMT state and mesenchymal characteristics increase. Further analysis involving the complete system of pathways is important in determining the extent of importance for these results.

Figure 2: Effect of T0 on epithelial and mesenchymal markers. E-cadherin (E-cad) is shown in blue, while N-cadherin (N-cad) is shown in red. The solid lines show phenotypic markers when T0 is 0, the dashed lines show when T0 is 1.8, and the dotted lines show when T0 is 5.

Through further analysis of the concentrations of transcription factors over time, Figure 3 also shows that SLUG protein has the largest delay. SLUG has the smallest rate of increase because its production is associated with activation by TWIST. SLUG cannot be produced at high levels until TWIST has been fully activated. However, TWIST and SLUG concentrations continue to hold relatively low values compared to SNAIL and ZEB. TWIST is most likely three-fold lower than SNAIL concentrations because it acts through SLUG, meaning TWIST concentrations do not need to be all that high to upregulate SLUG. This does not explain why SLUG concentrations remain low, however, as multiple SLUG proteins can be produced by one TWIST molecule. Both ZEB and SLUG have double negative feedback loops with miR-200. It appears that SLUG may be more inhibited by miR-200, causing less repression on zeb translation. ZEB is able to increase with a concentration approximately 2.5-fold greater than SLUG. This is supported by the Michaelis-Menten constants of translation inhibition by miR-200. The corresponding value for zeb translation is 0.06 µM and is half the value for slug translation at 0.03 µM, mathematically providing stronger repression on the production of SLUG than on the production of ZEB.
Moreover, the effect of the Hill coefficient is shown through these plots of Figure 3. The increases in concentration do not follow direct Michaelis-Menten kinetics, but rather show some cooperativity. Higher Hill coefficients would provide much more switch-like responses, rather than graded responses. Further research into the extent of cooperativity would prove interesting as to how these transcription factors actually behave within a cell.

Lastly we analyzed the stability properties of our system. Bifurcation plots were generated using MATCONT [33]. Figure 4 and Figure 5 show how the expression of E-Cadherin and N-Cadherin, respectively, vary against introduction of exogenous TGF-β. Congruent to results seen in [13], our model admits multiple steady states for certain values of T0 both in the expression of E-cadherin, [Figure 4] and N-cadherin [Figure 5]. These correspond to epithelial, mesenchymal, and partial states. Furthermore, stability analysis confirms that those states representative of fully epithelial or fully mesenchymal are stable. However, we remark that, as opposed to that of Tian et al. [13], our steady state corresponding to partial EMT is unstable. Biologically, the implication is that achieving a partial EMT state is not expected, and that cells tend to be in either full epithelial or mesenchymal states.

Furthermore, we investigated the effect of the perturbation for rate constants. Surprisingly, as it is not addressed in the literature, factors other that the introduction of TGF-β exhibit similar bifurcation properties. Figure 6 shows some of our results. It is worth noting that our model admits limit point and Hopf steady states. The former refers to a state that vanishes as the parameter is perturbed, whereas the later indicates that some species exhibit oscillatory properties in its expression for a specific range of parameters. The parameters analyzed in figure 6 include the basal transcription rate, degradation rate, transcription rate dependent on activators/repressors, and the Michaelis constant of translation of snail transcripts. These could cause oscillatory properties of E-Cadherin because SNAIL is a major inducer in the model and is therefore so heavily upregulated in the first simulated time points. Of these four parameters of snail, the $k_{sn}$ and $M_{sn}$ parameters shown in equation 2,
Figure 4: **Effect of exogenous TGF-β on steady-state E-cadherin levels.** Resulting bifurcation diagram of epithelial cell markers with varying T0 concentrations.

Figure 5: **Effect of exogenous TGF-β on steady-state N-cadherin levels.** Resulting bifurcation diagram of mesenchymal cell markers with varying T0 concentrations.
Figure 6: Bifurcation diagram of E-Cadherin against rates of snail. E-Cadherin expression is plotted against the (a) basal transcription rate, (b) the degradation rate, (c) the transcription rate dependent on activators/repressors, and (d) the Michaelis constant for translation of snail.

represented as $k_2$ and $m_1$ in Figure 6, are of the greatest importance. These define the transcription rate and Michaelis constant, respectively. Further experimentation needs to be done in order to assert with great confidence that our parameter values are correct. These parameters affect snail transcription, which leads to SNAIL translation, and then ultimately to the full induction of EMT onto the system.

4 Discussion

Epithelial cells form dense layers in structures found all over the body. Individual epithelial cells are able to form tight junctions in order to create a continuous and immobile tissue. In contrast, mesenchymal cells are highly motile and loosely connected to one another. Due to the complete changes in phenotype, it is expected that the transition from each of these states is tightly regulated. In wound healing, epithelial cells respond to stimuli and convert to mesenchymal cells. After this process is complete, the cells then re-epithelialize. This means the regulatory network must be able to respond to changes in stimuli in order to reversibly convert cells between epithelial and mesenchymal states but also be able to
withstand minor fluctuations within the network.

Our results show that exogenous TGF-β (T0) is an important player in the process of EMT. As exogenous TGF-β increases, the cell undergoes epithelial-to-mesenchymal transition. Biologically, this is the mutation or upregulation occurring in a cancerous cellular system. Theoretically, with these increases in exogenous TGF-β comes upregulation of other mesenchymal cell markers and downregulation of other epithelial cell markers. In normal wound healing, the cell would proceed to EMT but then re-epithelialize after migration. However, in cancerous cells, the extended exposure to exogenous TGF-β causes cells to remain in the mesenchymal state until they migrate to a new area of the body, where TGF-β levels may be lower and the cells can return to the epithelial state through the reverse process mesenchymal-to-epithelial transition (MET). This allows for cell proliferation and therefore tumor growth in a new area of the body.

The concentrations of each transcription factor over time, shown in Figure 3, provided further details regarding the system dynamics. The results emphasize the importance of SNAIL protein production in cancer development, providing a potential target for future therapeutics. Further research would benefit in confirming SNAIL as the main promoter of ZEB and TWIST production. This could be accomplished by knocking out the gene for SNAIL protein and determining the extent of EMT progression through this and other pathways. If other pathways do not cause drastic upregulation of ZEB, TWIST, or SLUG in the absence of SNAIL, therapeutics targeting SNAIL would be a powerful tool in slowing cancer metastasis. Moreover, our bifurcation analysis strongly suggests that EMT transition if regulated by factors other than the introduction of TGF-β, in particular, those concerned with the dynamics of transcription factors. Moreover we witnessed oscillatory properties in the concentrations of some of the species, a finding that needs to be verified experimentally. Future research are necessary to confirm the parameters generated in this work. While most parameters should be near their exact values, it is possible a few may be off, balanced by other imprecise values, to provide the necessary characteristics of EMT. This is one major limitation to this model, and the parameters cannot be confirmed without thorough experimentation. Another limitation to this model is the size of this pathway. In reality, several pathways with many more transcription factors play a role in EMT, and given more time and experimental data, it would be both interesting and highly important to account for their roles in this process. One addition would include some sort of feedback loop on TWIST. In the reported model, SNAIL, ZEB, and SLUG are all regulated by double negative feedback loops, but no modulator was provided for TWIST, which is highly unlikely to be true in vivo. Therefore, future models should account for this. Similarly, the extent of cooperativity should be analyzed to examine this effect on model outputs. Moreover, bifurcation and stability analysis could be performed by introducing exogenous concentrations of different species. Our team conjectures that a coupled perturbation of reaction constants and species concentrations might yield more insight into EMT transition.

5 Conclusion

We are confident that our results are of relevance to the modeling of epithelial to mesenchymal transition dynamics. However, many of these findings are yet to be experimentally verified. Specifically, a more robust model for transcription factors need to be incorporated, and at the same time, rate constants need to be experimentally established.
Moreover, future research should address those findings from our bifurcation analysis, such as the soundness of vanishing steady states and oscillatory properties of certain species. Lastly, a more comprehensive network description for EMT transition that accounts for the vast amount of recent findings would increase the rate at which mathematical models converge to the true biological dynamics. Further research in this direction will shed light in the manipulation of cell mutation. At the outset, we aim for a better understanding of diseases where cell mutation is involved, such as cancer, and in particular, metastasis. The authors thank the assistance of Dr. David Odde. (University of Minnesota) for his insight, comment and suggestions when developing this project.
## 6 Appendices

### 6.1 Model Inputs

Table 1: Parameters used in the Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_0$</td>
<td>Exogenous TGF-$\beta$</td>
<td>0.5 unit</td>
</tr>
<tr>
<td>$k_{0T}$</td>
<td>Basal production rate of TGF-$\beta$</td>
<td>0.06 $\mu$M/hr</td>
</tr>
<tr>
<td>$k_T$</td>
<td>Production rate of TGF-$\beta$</td>
<td>1.2 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{T200}$</td>
<td>Michaelis constant of TGF-$\beta$ expression inhibition by miR-200</td>
<td>0.06 $\mu$M</td>
</tr>
<tr>
<td>$M_{T200TW}$</td>
<td>Michaelis constant of TGF-$\beta$ expression activation by TWIST</td>
<td>1 $\mu$M</td>
</tr>
<tr>
<td>$k_{degT}$</td>
<td>Degradation rate of TGF-$\beta$</td>
<td>0.6/hr</td>
</tr>
<tr>
<td>$k_{0sn}$</td>
<td>Basal transcription rate of snail</td>
<td>0.0006 $\mu$M/hr</td>
</tr>
<tr>
<td>$k_{sn}$</td>
<td>Transcription rate of snail</td>
<td>0.03 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{sn}$</td>
<td>Michaelis constant of snail translation by TGF-$\beta$</td>
<td>1.2 $\mu$M</td>
</tr>
<tr>
<td>$k_{degsn}$</td>
<td>Degradation rate of snail</td>
<td>0.09/hr</td>
</tr>
<tr>
<td>$M_S$</td>
<td>Michaelis constant of SNAIL translation inhibition by miR-34</td>
<td>0.08 $\mu$M</td>
</tr>
<tr>
<td>$k_{degS}$</td>
<td>Degradation rate of SNAIL</td>
<td>1.66/hr</td>
</tr>
<tr>
<td>$k_{0z}$</td>
<td>Basal transcription rate of zeb</td>
<td>0.003 $\mu$M/hr</td>
</tr>
<tr>
<td>$k_z$</td>
<td>Transcription rate of zeb</td>
<td>0.06 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_z$</td>
<td>Michaelis constant of zeb transcription by SNAIL</td>
<td>3.5 $\mu$M</td>
</tr>
<tr>
<td>$k_{degz}$</td>
<td>Degradation rate of zeb</td>
<td>0.09/hr</td>
</tr>
<tr>
<td>$k_{0tw}$</td>
<td>Basal transcription rate of twist</td>
<td>0.0017 $\mu$M/hr</td>
</tr>
<tr>
<td>$k_{tw}$</td>
<td>Transcription rate of twist</td>
<td>0.06 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{tw}$</td>
<td>Michaelis constant of twist transcription by SNAIL</td>
<td>3.5 $\mu$M</td>
</tr>
<tr>
<td>$k_{degtw}$</td>
<td>Degradation rate of twist</td>
<td>0.09/hr</td>
</tr>
<tr>
<td>$k_{0SL}$</td>
<td>Basal transcription rate of slug</td>
<td>0.006 $\mu$M/hr</td>
</tr>
<tr>
<td>$k_{sl}$</td>
<td>Transcription rate of slug</td>
<td>0.03 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{sl}$</td>
<td>Michaelis constant of twist transcription by TWIST</td>
<td>3 $\mu$M</td>
</tr>
<tr>
<td>$k_{degsl}$</td>
<td>Degradation rate of slug</td>
<td>0.09/hr</td>
</tr>
<tr>
<td>$k_{SL}Z$</td>
<td>Translation rate of ZEB via slug</td>
<td>17 $\mu$M/hr</td>
</tr>
<tr>
<td>$k_{degSL}$</td>
<td>Degradation rate of ZEB</td>
<td>1.66/hr</td>
</tr>
<tr>
<td>$k_{0SL}$</td>
<td>Basal transcription rate of SLUG</td>
<td>0.001 $\mu$M/hr</td>
</tr>
<tr>
<td>$k_{SL}$</td>
<td>Translation rate of SLUG via slug</td>
<td>5.6 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{SL}$</td>
<td>Michaelis constant of slug translation by SNAIL</td>
<td>0.03 $\mu$M</td>
</tr>
<tr>
<td>$k_{degSL}$</td>
<td>Degradation rate of SLUG</td>
<td>5.66/hr</td>
</tr>
<tr>
<td>$k_{034}$</td>
<td>Basal production rate of miR-34</td>
<td>0.0012 $\mu$M/hr</td>
</tr>
<tr>
<td>$k_{34}$</td>
<td>Production rate of miR-34</td>
<td>0.012 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{34S}$</td>
<td>Michaelis constant of miR-34 production inhibition by SNAIL</td>
<td>0.15 $\mu$M</td>
</tr>
<tr>
<td>$k_{deg34}$</td>
<td>Degradation rate of miR-34</td>
<td>0.035/hr</td>
</tr>
<tr>
<td>$k_{0200}$</td>
<td>Basal production rate of miR-200</td>
<td>0.0002 $\mu$M/hr</td>
</tr>
<tr>
<td>$k_{200}$</td>
<td>Production rate of miR-200</td>
<td>0.012 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{200Z}$</td>
<td>Michaelis constant of miR-200 production inhibition by ZEB</td>
<td>0.9 $\mu$M</td>
</tr>
<tr>
<td>$M_{200SL}$</td>
<td>Michaelis constant of miR-200 production inhibition by SLUG</td>
<td>1.5 $\mu$M</td>
</tr>
<tr>
<td>$k_{deg200}$</td>
<td>Degradation rate of miR-200</td>
<td>0.035/hr</td>
</tr>
<tr>
<td>$k_{es}$</td>
<td>Production rate of E-cadherin by SNAIL</td>
<td>0.6 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{es}$</td>
<td>Michaelis constant of E-cadherin production inhibition by SNAIL</td>
<td>0.2 $\mu$M</td>
</tr>
<tr>
<td>$k_{ez}$</td>
<td>Production rate of E-cadherin by ZEB</td>
<td>0.4 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{ez}$</td>
<td>Michaelis constant of E-cadherin production inhibition by ZEB</td>
<td>0.5 $\mu$M</td>
</tr>
<tr>
<td>$k_{esSL}$</td>
<td>Production rate of E-cadherin by SLUG</td>
<td>0.4 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{esSL}$</td>
<td>Michaelis constant of E-cadherin production inhibition by SLUG</td>
<td>0.5 $\mu$M</td>
</tr>
<tr>
<td>$k_{degE}$</td>
<td>Degradation rate of E-cadherin</td>
<td>0.5/hr</td>
</tr>
<tr>
<td>$k_{en}$</td>
<td>Production rate of N-cadherin by SNAIL</td>
<td>0.6 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{en}$</td>
<td>Michaelis constant of N-cadherin production by SNAIL</td>
<td>0.2 $\mu$M</td>
</tr>
<tr>
<td>$k_{enZ}$</td>
<td>Production rate of N-cadherin by ZEB</td>
<td>0.4 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{enZ}$</td>
<td>Michaelis constant of N-cadherin production by ZEB</td>
<td>0.5 $\mu$M</td>
</tr>
<tr>
<td>$k_{enSL}$</td>
<td>Production rate of N-cadherin by SLUG</td>
<td>0.4 $\mu$M/hr</td>
</tr>
<tr>
<td>$M_{enSL}$</td>
<td>Michaelis constant of N-cadherin production by SLUG</td>
<td>0.5 $\mu$M</td>
</tr>
<tr>
<td>$k_{degN}$</td>
<td>Degradation rate of N-cadherin</td>
<td>0.5/hr</td>
</tr>
</tbody>
</table>
Table 2: Table of initial conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Initial value (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[TGFβ]</td>
<td>Concentration of TGF-β</td>
<td>0.16</td>
</tr>
<tr>
<td>[snail]</td>
<td>Concentration of snail</td>
<td>0.01</td>
</tr>
<tr>
<td>[SNAIL]</td>
<td>Concentration of SNAIL</td>
<td>0.01</td>
</tr>
<tr>
<td>[zeb]</td>
<td>Concentration of zeb</td>
<td>0.03</td>
</tr>
<tr>
<td>[ZEB]</td>
<td>Concentration of ZEB</td>
<td>0.01</td>
</tr>
<tr>
<td>[twist]</td>
<td>Concentration of twist</td>
<td>0.01</td>
</tr>
<tr>
<td>[TWIST]</td>
<td>Concentration of TWIST</td>
<td>0.01</td>
</tr>
<tr>
<td>[slug]</td>
<td>Concentration of slug</td>
<td>0.01</td>
</tr>
<tr>
<td>[SLUG]</td>
<td>Concentration of SLUG</td>
<td>0.01</td>
</tr>
<tr>
<td>[mR34]</td>
<td>Concentration of miR-34</td>
<td>0.38</td>
</tr>
<tr>
<td>[mR200]</td>
<td>Concentration of miR-200</td>
<td>0.35</td>
</tr>
<tr>
<td>[Ecad]</td>
<td>Concentration of E-cadherin</td>
<td>2.8</td>
</tr>
<tr>
<td>[Ncad]</td>
<td>Concentration of N-cadherin</td>
<td>0</td>
</tr>
</tbody>
</table>
6.2 Justification for Model Parameters

1. The half-life of SNAIL is approximately 25 minutes [27]. Because there was no other literature reference for the other transcription factors, it is assumed that ZEB, TWIST, and SLUG are the same as SNAIL.

2. The half-life of mRNA degradation is approximately 7-9 hours [28],[29].

3. The half-life of miRNA is considerably longer at approximately 119 hours (5 days); however, miRNA concentration is decreased due to growth related dilution from cell division and is therefore given a half-life of 20 hours, which is about the length of one cycle of cell division [30],[31].

4. The concentration of transcription factor protein is one-fold greater than the concentration of mRNAs and miRNAs [13].

5. Each gene has a basal transcriptional rate that is much smaller than the rate that is dependent upon transcription factors.

6. EMT occurs in a matter of days rather than the minute or hour timescale [32].
6.3 MATLAB Code

```matlab
function rhs = ODE_13_v4(t,C,T0)

% 1. TGF
% 2. snail
% 3. SNAIL
% 4. miR34
% 5. zeb
% 6. ZEB
% 7. miR200
% 8. E-cadherin
% 9. N-cadherin
% 10. twist
% 11. TWIST
% 12. slug
% 13. SLUG

% Equation parameters for TGF
k1_t = 0.06 ; % Basal production rate of TGF-B
k2_t = 1.2 ; % Production rate of TGF-B
m1_t = 0.06 ; % Michaelis constant of TGF-B inhibition by miR200
m2_t = 1 ; % Michaelis constant of TGF-B activation by TWIST
k3_t = 0.6 ; % Degradation rate of TGF-$\beta$

% Equation parameters for snail
k1_s = 0.0006 ; % Basal transcription rate of snail
k2_s = 0.03 ; % Transcription rate of snail
m1_s = 1.2 ; % Michaelis constant of snail translation by TGF-B
k3_s = 0.09 ; % Degradation rate of snail

% Equation parameters for SNAIL
k1_S = 17 ; % Production rate of SNAIL by snail via translation
m1_S = 0.08 ; % Michaelis constant of SNAIL translation inhibition by miR34
k2_S = 1.66 ; % Degradation rate of SNAIL

% Equation parameters for miR34
k1_3 = 0.0012 ; % Basal production rate of miR-34
k2_3 = 0.012 ; % Production rate of miR-34
m1_3 = 0.015 ; % Michaelis constant of miR-34 production inhibition by SNAIL
m2_3 = 0.16 ; % Michaelis constant of miR-34 production inhibition by ZEB
k3_3 = 0.035 ; % Degradation rate of miR-34

% Equation parameters for zeb
k1_z = 0.003 ; % Basal transcription rate of zeb
k2_z = 0.06 ; % Transcription rate of zeb
m1_z = 3.5 ; % Michaelis constant of zeb transcription by SNAIL
k3_z = 0.09 ; % Degradation rate of zeb

% Equation parameters of ZEB
k1_Z = 17 ; % Translation rate of ZEB via zeb
m1_Z = 0.06 ; % Michaelis constant of zeb translation inhibition by miR
```
\[ k_{2-Z} = 1.66 \quad ; \quad \text{Degradation rate of ZEB} \]

%Equation parameters of miR-200
\[ k_{1-2} = 0.0002 \quad ; \quad \text{% Basal production rate of miR-200} \]
\[ k_{2-2} = 0.012 \quad ; \quad \text{% Production rate of miR-200} \]
\[ m_{1-2} = 1.25 \quad ; \quad \text{% Michaelis constant of miR-200 production inhibition by SNAIL} \]
\[ m_{2-2} = 0.9 \quad ; \quad \text{% Michaelis constant of miR-200 production inhibition by ZEB} \]
\[ m_{3-2} = 1.5 \quad ; \quad \text{% Michaelis constant of miR-200 production inhibition by SLUG} \]
\[ k_{3-2} = 0.035 \quad ; \quad \text{% Degradation rate of miR-200} \]

%Equation parameters of E-cadherin
\[ k_{1-e} = 0.605 \quad ; \quad \text{% Production rate of E-cadherin (by SNAIL)} \]
\[ m_{1-e} = 0.2 \quad ; \quad \text{% Michaelis constant of E-cadherin production inhibition by SNAIL} \]
\[ k_{2-e} = 0.4 \quad ; \quad \text{% Production rate of E-cadherin (by ZEB)} \]
\[ m_{2-e} = 0.5 \quad ; \quad \text{% Michaelis constant of E-cadherin production inhibition by ZEB} \]
\[ k_{3-e} = 0.4 \quad ; \quad \text{% Production rate of E-cadherin (by SLUG)} \]
\[ m_{3-e} = 0.5 \quad ; \quad \text{% Michaelis constant of E-cadherin production inhibition by SLUG} \]
\[ k_{4-e} = 0.4 \quad ; \quad \text{% Production rate of E-cadherin by TWIST} \]
\[ m_{4-e} = 0.4 \quad ; \quad \text{% Michaelis constant of E-cadherin production by TWIST} \]
\[ k_{5-e} = 0.5 \quad ; \quad \text{% Degradation rate of E-cadherin} \]

%Equation parameters for N-cadherin
\[ k_{1-n} = 0.61 \quad ; \quad \text{% Production rate of N-cadherin by SNAIL} \]
\[ m_{1-n} = 0.2 \quad ; \quad \text{% Michaelis constant of N-cadherin production by SNAIL} \]
\[ k_{2-n} = 0.4 \quad ; \quad \text{% Production rate of N-cadherin by ZEB} \]
\[ m_{2-n} = 0.5 \quad ; \quad \text{% Michaelis constant of N-cadherin production by ZEB} \]
\[ k_{3-n} = 0.4 \quad ; \quad \text{% Production rate of N-cadherin by SLUG} \]
\[ m_{3-n} = 0.5 \quad ; \quad \text{% Michaelis constant of N-cadherin production by SLUG} \]
\[ k_{4-n} = 0.4 \quad ; \quad \text{% Production rate of N-cadherin by TWIST} \]
\[ m_{4-n} = 0.4 \quad ; \quad \text{% Michaelis constant of N-cadherin production by TWIST} \]
\[ k_{5-n} = 0.5 \quad ; \quad \text{% Degradation rate of N-cadherin} \]

%Equation parameters for Twist
\[ k_{1-w} = 0.0017 \quad ; \quad \text{% Basal transcription rate of twist} \]
\[ k_{2-w} = 0.06 \quad ; \quad \text{% Transcription rate of twist} \]
\[ m_{1-w} = 3.5 \quad ; \quad \text{% Michaelis constant of twist transcription by SNAIL} \]
\[ k_{3-w} = 0.09 \quad ; \quad \text{% Degradation rate of twist} \]

%Equation parameters for TWIST
\[ k_{1-W} = 14 \quad ; \quad \text{% Translation rate of TWIST via twist} \]
\[ k_{2-W} = 5.6 \quad ; \quad \text{% Degradation rate of TWIST} \]

%Equation parameters for slug
\[ k_{1-g} = 0.006 \quad ; \quad \text{% Basal transcription rate of slug} \]
\[ k_{2-g} = 0.03 \quad ; \quad \text{% Transcription rate of slug} \]
\[ m_{1-g} = 3 \quad ; \quad \text{% Michaelis constant of twist transcription by TWIST} \]

20
k3_g = 0.09 ; % Degradation rate of slug

% Equation parameters for SLUG
k1_G = 10 ; % Translation rate of SLUG via slug
m1_G = 0.03 ; % Michaelis constant of slug translation inhibition by miR200
k2_G = 5.66 ; % Degradation rate of SLUG

rhs=NaN(13,1);

% 1. TGF
rhs(1) = k1_t + k2_t*(((C(11)/m2_t)^2)/(1+(C(7)/m1_t)^2)+(C(11)/m2_t)^2) - k3_t*C(1);

% 2. snail
rhs(2) = k1_s + k2_s*(((C(1)+T0)/(m1_s))^2)/(1+((C(1)+T0)/(m1_s))^2) - k3_s*C(2);

% 3. SNAIL
rhs(3) = k1_S*C(2)*((C(1)+T0)/(m1_S))^2 - k2_S*C(3);

% 4. miR34
rhs(4) = k1_3 + k2_3*(((C(3))/(m1_3))^2)/(1+((C(3))/(m1_3))^2) - k3_3*C(4);

% 5. zeb
rhs(5) = k1_z + k2_z*(((C(3))/(m1_z))^2)/(1+((C(3))/(m1_z))^2) - k3_z*C(5);

% 6. ZEB
rhs(6) = k1_Z*C(5)*((C(1)+T0)/(m1_Z))^2 - k2_Z*C(6);

% 7. miR200
rhs(7) = k1_2 + k2_2*(((C(3))/(m1_2))^2)/(1+((C(3))/(m1_2))^2) + ((C(6))/(m2_2))^2 - k3_2*C(7);

% 8. E-cadherin (needs to be revised)
rhs(8) = k1_e*(1/(1+((C(3))/(m1_e))^2)) - k3_e*C(8);

% 9. N-cadherin (needs to be revised)
rhs(9) = k1_n*(((C(3))/(m1_n))^2)/(1+((C(3))/(m1_n))^2) + k2_n*(((C(6))/(m2_n))^2)/(1+((C(6))/(m2_n))^2) - k5_n*C(9);

% 10. twist
rhs(10) = k1_w + k2_w*(((C(3))/(m1_w))^2)/(1+(C(3))/(m1_w))^2) - k3_w*C(10);

% 11. TWIST
rhs(11) = k1_W*C(10) - k2_W*C(11);

% 12. slug
rhs(12) = k1_g + k2_g*(((C(11))/(m1_g))^2)/(1+(C(11))/(m1_g))^2) - k3_g*C(12);

% 13. SLUG
rhs(13) = k1_G*(((C(7))/(m1_G))^2) - k2_G*C(13);

C_ini = [3.16;0.01;0.01;0.38;0.03;0.01;0.35;3.2;0];
[time, C] = ode15s(@ODE,[0,480],C_ini);
References


