

UNIVERSIDADE FEDERAL DE SANTA MARIA  
CENTRO DE CIÊNCIAS NATURAIS E EXATAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

Daner Acunha Silveira

**MODELAGEM DA DINÂMICA DO PROCESSO DE TRANSIÇÃO  
EPITÉLIO-MESENQUIMAL EM CÂNCER**

Santa Maria, RS  
2021

**Daner Acunha Silveira**

**MODELAGEM DA DINÂMICA DO PROCESSO DE TRANSIÇÃO  
EPITÉLIO-MESENQUIMAL EM CÂNCER**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Física, Área de Concentração em Sistemas Complexos, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Doutor em Física**.

ORIENTADOR: Prof. José Carlos Merino Mombach

Santa Maria, RS  
2021

**Daner Acunha Silveira**

**MODELAGEM DA DINÂMICA DO PROCESSO DE TRANSIÇÃO  
EPITÉLIO-MESENQUIMAL EM CÂNCER**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Física, Área de Concentração em Sistemas Complexos, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Doutor em Física**.

**Aprovado em 10 de agosto de 2021.**

**Comissão examinadora:**

---

**José Carlos Merino Mombach, Dr. (UFSM)**  
(Presidente/Orientador)

---

**Solange Binotto Fagan, Dra. (UFN)**

---

**Marialva Sinigaglia, Dra. (ICI)**

---

**Jonas Maziero, Dr. (UFSM)**

---

**Lucio Strazzabosco Dorneles, Dr. (UFSM)**

Santa Maria, RS  
2021

## DEDICATÓRIA

*À minha família*

## **AGRADECIMENTOS**

*Gostaria de agradecer a todas as pessoas que contribuíram de alguma forma para que este trabalho de mestrado fosse realizado:*

*Aos meus pais, Gilberto de Castro Silveira e Isabel Cristina Acunha Silveira, pela confiança que me deram durante toda essa jornada;*

*Ao meu orientador, Prof. Dr. José Carlos Merino Mombach, pela paciência, amizade, confiança e ensinamentos.*

*Ao Dr. Shantanu Gupta, por toda amizade e colaboração que desenvolvemos durante todo esse período;*

*Aos meus outros colegas de laboratório, pela amizade, força e apoio;*

*Aos professores que participaram da minha formação acadêmica, por toda a sabedoria que me passaram durante todo esse processo;*

*A Universidade Federal de Santa Maria, pelo suporte durante a pesquisa;*

*Ao CNPq pelo apoio financeiro.*

*"If I have seen further it is by standing  
on the shoulders of Giants."*

*(Isaac Newton)*

## RESUMO

### MODELAGEM DA DINÂMICA DO PROCESSO DE TRANSIÇÃO EPITÉLIO-MESENQUIMAL EM CÂNCER

AUTOR: Daner Acunha Silveira

ORIENTADOR: José Carlos Merino Mombach

A transição epitélio-mesenquimal (TEM) é um processo celular fundamental baseado na conversão de células epiteliais em mesenquimais, que desempenha um papel essencial na disseminação do câncer através da metástase. A evolução do conhecimento em termos das moléculas e interações que governam o complexo sistema da TEM tem requerido a utilização de abordagens computacionais que proporcionam resultados não intuitivos sobre a complexidade de sua dinâmica. Assim, este estudo objetiva analisar recentes informações bioquímicas publicadas na literatura através da modelagem computacional lógica de maneira a amplificar o conhecimento dinâmico da transição. Os dois modelos propostos no presente estudo são relacionados à molécula ativadora da TEM TGF- $\beta$ . Seus resultados foram comparados com dados teóricos e experimentais para a linhagem de células epiteliais da mama humana MCF10A apresentando excelente concordância. Coletivamente, os resultados demonstraram a funcionalidade dos circuitos durante a transição e, dessa forma, os microRNAs 129, 340 e 1199 podem ser considerados fatores de estabilidade fenotípica do estado híbrido. Adicionalmente, os dados revelaram as probabilidades fenotípicas na região de coexistência e constatam que o estado híbrido é o mais provável na dinâmica tri-estável, concordando com dados experimentais. Portanto, o presente estudo demonstra novos mecanismos relacionados à dinâmica da TEM em resposta à molécula TGF- $\beta$  em células epiteliais da mama e tais resultados auxiliam no aprimoramento de novas estratégias terapêuticas para o câncer de mama.

**Palavras-chave:** TEM. Modelagem lógica. MicroRNA.

## ABSTRACT

### MODELING THE DYNAMICS OF THE EPITHELIAL-MESENCHYMAL TRANSITION PROCESS IN CANCER

AUTHOR: Daner Acunha Silveira

ADVISOR: José Carlos Merino Mombach

The epithelial-mesenchymal transition (TEM) is a fundamental cellular process based on the conversion of epithelial cells into mesenchymal cells, which plays an essential role in the dissemination of cancer through metastasis. The evolution of knowledge in terms of the molecules and interactions that govern the complex TEM system has required the use of computational approaches that provide non-intuitive results on the complexity of its dynamics. Thus, this study aims to analyze recent biochemical information published in the literature through logical computational modeling in order to amplify the dynamic knowledge of the transition. The two proposed models are related to TEM activating molecule TGF- $\beta$ . The data were compared with theoretical and experimental data for MCF10A human breast epithelial cell line presenting excellent agreement. Collectively, the results demonstrated the functionality of the circuits during the transition and, thus, microRNAs 129, 340 and 1199 may be considered factors of phenotypic stability of the hybrid state. Additionally, the data revealed the phenotypic probabilities in the coexistence region and demonstrated that the hybrid state is the most probable in the tristable dynamics, agreeing with experimental data. Therefore, the present study highlights new mechanisms related to the dynamics of TEM in response to the TGF- $\beta$  molecule in breast epithelial cells and such results help to improve new therapeutic strategies for breast cancer.

**Keywords:** TEM. Logical modeling. MicroRNA.



## LISTA DE FIGURAS

Figura 1.1 – Representação esquemática do processo da transição epitélio-mesenquimal (TEM) e seus respectivos tipos .....	16
Figura 1.2 – Exemplo de uma rede de regulação genética .....	18
Figura 1.3 – Tipos de circuitos celulares .....	19
Figura 1.4 – Exemplo do circuito negativo p53/Mdm2 e sua respectiva atividade funcional observada experimentalmente .....	20
Figura 1.5 – Exemplo dos circuitos positivos SNAIL1/miR-34 e ZEB1/miR-200 aplicados a uma rede da transição epitélio-mesenquimal e suas respectivas atividades funcionais observadas experimentalmente .....	20
Figura 1.7 – Exemplo de aplicação do formalismo lógico em uma rede de regulação genética .....	24
Figura 1.8 – Exemplo da aplicação do método de Avatar .....	27
Figura 1.10 – Aplicação da modelagem lógica no circuito negativo p53/Mdm2 ...	30
Figura 1.11 – Aplicação da modelagem lógica na rede proposta por Zhang et al. (2014) relacionada à transição epitélio-mesenquimal. ....	31
Figura 5.1 – Circuitos entre a componente da via Wnt $\beta$ -catenina e suas moléculas alvo .....	90
Figura A.1 – Sequência temporal de estados de um atrator cíclico .....	101
Figura C.1 – Exemplo de relação de regras lógicas e contexto de funcionalidade de circuitos através do uso do operador AND. ....	105
Figura C.2 – Exemplo de relação de regras lógicas e contexto de funcionalidade de circuitos através do uso do operador OR. ....	106
Figura C.3 – Exemplo de relação de regras lógicas e contexto de funcionalidade de circuitos para uma rede complexa .....	107

## LISTA DE TABELAS

Tabela A.1 – Relação lógica dos valores das variáveis $x$ e $y$ , e suas respectivas funções $X$ e $Y$ . .....	100
Tabela A.2 – Notação em termos da evolução dos estados da Tabela A.1.....	101

## LISTA DE ABREVIATURAS E SIGLAS

<i>AKT1</i>	<i>AKT serine/threonine kinase 1</i>
<i>CD24</i>	Molécula CD24
<i>CDH1</i>	Caderina 1
<i>CTC</i>	Célula tumoral circulante
<i>E2F1</i>	Fator de transcrição E2F1
<i>FCE</i>	Fator de crescimento epidérmico
<i>ESRP1</i>	Proteína reguladora de splicing epitelial 1
<i>FEF</i>	Fator de estabilidade fenotípica
<i>FT</i>	Fator de transcrição
<i>GINsim</i>	Gene Interaction Network simulation suite
<i>GRHL2</i>	<i>Grainyhead like transcription factor 2</i>
<i>GSK-3</i>	<i>Glicogênio sintase quinase 3</i>
<i>GF</i>	Ganho de função ( <i>Gain of function</i> )
<i>GTE</i>	Gráfico de transição de estado
<i>HA</i>	Ácido hialurônico sintetizado por HAS2
<i>HAS2</i>	<i>Ácido hialurônico sintase-2</i>
<i>hTGF-<math>\beta</math>1</i>	<i>Recombinant transforming growth factor beta-1</i>
<i>LEF/TCF</i>	<i>Fatores de transcrição da família TCF/LEF</i>
<i>LncRNAs</i>	RNAs longos não-codificantes
<i>PF</i>	Perda de função
<i>Mdm2</i>	<i>E3 ubiquitin protein ligase homolog protein</i>
<i>miRNA</i>	MicroRNA
<i>miR-129</i>	MicroRNA 129
<i>miR-200</i>	Família miR-200
<i>miR-34</i>	Família miR-34
<i>miRNA</i>	MicroRNA
<i>mRNA</i>	RNA mensageiro ( <i>Messenger RNA</i> )

<i>N-caderina</i>	Caderina N
<i>OVOL</i>	Família Ovol
<i>p53</i>	Proteína 53 supressora de tumor
<i>pre-miRNA</i>	MicroRNA precursor
<i>pri-miRNA</i>	MicroRNA primário
<i>RRG</i>	Rede de regulação genética
<i>S</i>	Espaço de estados
<i>Siah1</i>	<i>Siah E3 ubiquitin protein ligase 1</i>
<i>SNAIL1</i>	Repressor transcricional 1 da família SNAIL
<i>SNAIL2</i>	Repressor transcricional 2 da família SNAIL
<i>STAT3</i>	<i>Transdutor de sinal e ativador de transcrição 3</i>
<i>TEM</i>	Transição epitélio-mesenquimal
<i>TGF-<math>\beta</math></i>	Fator de transformação do crescimento- $\beta$
<i>TWIST1</i>	<i>Twist family bHLH transcription factor 1</i>
<i>VIM</i>	Vimentina
<i>WNT</i>	Família Wnt
<i>ZEB1</i>	<i>Zinc finger E-box binding homeobox 1</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO GERAL</b> .....	<b>13</b>
1.1	TRANSIÇÃO EPITÉLIO-MESENQUIMAL .....	13
1.2	REDES DE REGULAÇÃO GENÉTICA .....	16
1.3	MODELAGEM DE REDES REGULATÓRIAS: FORMALISMO LÓGICO ...	21
<b>1.3.1</b>	<b>Algoritmo Avatar</b> .....	<b>25</b>
<b>1.3.2</b>	<b>Software GINsim</b> .....	<b>28</b>
1.3.2.1	<i>Tradução de uma rede de regulação genética em um modelo lógico no GINsim</i> .....	28
1.3.2.2	<i>Funcionalidade dos circuitos</i> .....	31
<b>1.3.3</b>	<b>Banco de dados em biologia sistêmica</b> .....	<b>32</b>
1.4	OBJETIVOS .....	33
<b>2</b>	<b>ARTIGO: DYNAMICS OF THE FEEDBACK LOOPS REQUIRED FOR THE PHENOTYPIC STABILIZATION IN THE EPITHELIAL-MESENCHYMAL TRANSITION</b> .....	<b>35</b>
<b>3</b>	<b>ARTIGO: SYSTEMS BIOLOGY APPROACH SUGGESTS NEW MIRNAS AS PHENOTYPIC STABILITY FACTORS IN THE EPITHELIAL-MESENCHYMAL TRANSITION</b> .....	<b>56</b>
<b>4</b>	<b>DISCUSSÃO GERAL</b> .....	<b>79</b>
<b>5</b>	<b>CONCLUSÃO</b> .....	<b>88</b>
5.1	PERSPECTIVAS FUTURAS .....	89
	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	<b>91</b>
	<b>APÊNDICE A – DINÂMICA DE UM ATRATOR CÍCLICO</b> .....	<b>100</b>
	<b>APÊNDICE B – FUNDAMENTOS TEÓRICOS DO MÉTODO DE CADEIAS DE MARKOV</b> .....	<b>102</b>
	<b>APÊNDICE C – RELAÇÃO ENTRE REGRAS LÓGICAS E CONTEXTO DE FUNCIONALIDADE DOS CIRCUITOS</b> .....	<b>104</b>
	<b>APÊNDICE D – OUTROS ARTIGOS PUBLICADOS</b> .....	<b>108</b>

# 1 INTRODUÇÃO GERAL

Este capítulo descreve assuntos introdutórios associados à presente pesquisa, abordando os principais tópicos relacionados aos fundamentos biológicos e computacionais da área. Estão apresentados os conceitos básicos sobre a transição epitélio-mesenquimal, redes de regulação genética, modelagem lógica e os objetivos da presente pesquisa.

## 1.1 TRANSIÇÃO EPITÉLIO-MESENQUIMAL

Historicamente, células epiteliais e mesenquimais foram identificadas com base em sua aparência visual e na morfologia de suas estruturas multicelulares (SHOOK; KELLER, 2003). Um epitélio é caracterizado por células unidas por estruturas de adesão, as quais auxiliam na sustentação do citoesqueleto de cada célula epitelial aos seus vizinhos, assim como sobre os materiais da matriz extracelular subjacente ou circundante. Dessa forma, as junções célula-célula regularmente espaçadas e as aderências entre as células epiteliais vizinhas mantêm-nas unidas e inibem o movimento de células individuais para distante da camada epitelial. Esta estrutura é caracterizada por junções adesivas especializadas com base em moléculas de adesão celular dependentes de cálcio denotadas de caderinas (RODRIGUEZ-BOULAN; MACARA, 2014). Em contrapartida, as células mesenquimais geralmente não exibem estrutura e adesão intracelular rígidas. Assim, a migração mesenquimal é mecanicamente diferente do movimento epitelial. As células epiteliais se movem em conjunto em termos teciduais, enquanto a migração mesenquimal é consideravelmente mais dinâmica e pode ocorrer de forma individual (LEE et al., 2006).

No início da década de 1980, a partir de observações feitas por Elizabeth Hay em embriões de aves (HAY, 1995), surgiram evidências de que células epiteliais podem transitar para um fenótipo mesenquimal. Inicialmente descrita como transformação epitélio-mesenquimal, esse processo de diferenciação é agora comumente conhecido como transição epitélio-mesenquimal (TEM) para enfatizar sua natureza transitória. Biologicamente, a TEM é definida como um mecanismo de transição de células epiteliais para mesenquimais através de múltiplas alterações bioquímicas que permitem a estas células assumir tal fenótipo. Estas alterações incluem maior capacidade migratória, invasividade e elevada resistência à morte celular (ISER et al., 2017; KALLURI; WEINBERG, 2009). A conclusão do processo da TEM é caracterizada pela degradação da membrana basal, mecanismo no qual facilita a migração da célula mesenquimal para sítios distantes da camada epitelial em que tal célula se originou.

A TEM pode ser dividida em três tipos de acordo com diferentes processos biológicos: formação de embriões, desenvolvimento de órgãos e a geração de vários tipos de células (tipo 1); cicatrização, regeneração tecidual e fibrose (tipo 2); progressão do câncer e metástase (tipo 3).

O tipo 3 da TEM tem sido muito estudado devido a sua direta influência na progressão tumoral (BRABLETZ et al., 2018; BRABLETZ; BRABLETZ, 2010; SIEMENS et al., 2011). O câncer, por definição, é caracterizado pelo aumento populacional de células de forma desordenada. Através do processo de metástase, tais células podem invadir outros tecidos e, assim, originar outro tumor no organismo. No contexto da TEM, o ganho da capacidade migratória de células epiteliais do câncer favorece a ativação da metástase.

Em um nível molecular, a TEM é governada pela ativação de fatores de transcrição, expressão de proteínas específicas da superfície celular e do citoesqueleto, produção de enzimas que degradam a matriz extracelular e mudanças na expressão de microRNAs (miRNAs) específicos (KALLURI; WEINBERG, 2009). Esse tipo de molécula foi descoberta em 1993 (LEE; FEINBAUM; AMBROS, 1993) e é atualmente descrita como fator determinante no controle da dinâmica molecular da célula. Biologicamente, tais moléculas são pequenos RNAs não-codificadores com média de 22 nucleotídeos (LING; FABBRI; CALIN, 2013). Os miRNAs funcionam como reguladores gênicos pós-transcricionais que, na maioria dos casos, diminuem a estabilidade ou inibem a tradução de RNAs mensageiros (mRNAs) (CAI et al., 2009). Grande parte dessas moléculas são transcritas de sequências de DNA em miRNAs primários (pri-miRNAs) e processados em miRNAs precursores (pré-miRNAs) e miRNAs maduros (O'BRIEN et al., 2018).

As funções celulares são controladas pela informação genética presente nos genes do DNA de cada célula, os quais são codificados em moléculas de proteína que executam de maneira sistemática tais funções (LODISH et al., 2013). O processo de codificação de um gene em uma proteína é realizado através de dois mecanismos fundamentais na dinâmica celular: transcrição e tradução. Na transcrição, a informação no DNA é transferida para uma molécula de RNA mensageiro (mRNA). A partir disso, o mRNA é codificado em uma sequência de aminoácidos, na qual constitui uma proteína, definindo, assim, o processo de tradução. Dessa forma, a expressão de um gene é definida a partir da síntese de sua proteína correspondente. No contexto da TEM, as proteínas podem ser interpretadas como marcadores para cada fenótipo da transição. No caso epitelial, moléculas como ESRP1, GRHL2, CD24 e CDH1 mantêm e auxiliam o funcionamento de tal tecido, enquanto que proteínas como SNAIL1, ZEB1, VIM, SNAIL2 e N-caderina são responsáveis pela estabilização do fenótipo mesenquimal.

Durante a TEM, essas moléculas tem apresentado grande importância através

do controle da dinâmica de transição de cada fenótipo (ABBA et al., 2016). Especificamente, duas famílias de miRNAs, microRNA-200 (miR-200) e 34 (miR-34), foram reportadas como reguladoras da TEM (KORPAL; KANG, 2008; SIEMENS et al., 2011). Dados experimentais sugeriram que essas moléculas podem inibir transcricionalmente proteínas relacionadas ao fenótipo mesenquimal, tais como SNAIL1 e ZEB1 (KORPAL; KANG, 2008; SIEMENS et al., 2011). Entretanto, estudos recentes têm sugerido que outros miRNAs podem afetar tal dinâmica, e, assim, revelando um cenário mais complexo associado à transição (DIEPENBRUCK et al., 2017; HOU et al., 2016).

A sinalização inicial de ativação relacionada à TEM pode ser desencadeada por uma variedade de moléculas, tais como o fator de transformação do crescimento- $\beta$  (TGF- $\beta$ ), fator de crescimento epidérmico (FCE) e a WNT (CRAENE; BERX, 2013). Destas, a molécula de TGF- $\beta$  é estudada mais extensivamente no contexto da TEM (BURGER; DANEN; BELTMAN, 2017). Tal fator é desregulado em muitos tipos de câncer e atua tanto como um supressor tumoral em estágios iniciais (inibindo a proliferação e induzindo a apoptose) quanto como um promotor em estágios finais do tumor (induzindo a TEM ou suprimindo respostas imunes) (MARCUCCI; STASSI; MARIA, 2016). Fenômeno esse conhecido como o "paradoxo do TGF- $\beta$ " (ROSSATO et al., 2019). Como resultado deste comportamento multifuncional, o interesse na compreensão na sinalização de TGF- $\beta$  durante a TEM tem crescido consideravelmente no meio científico.

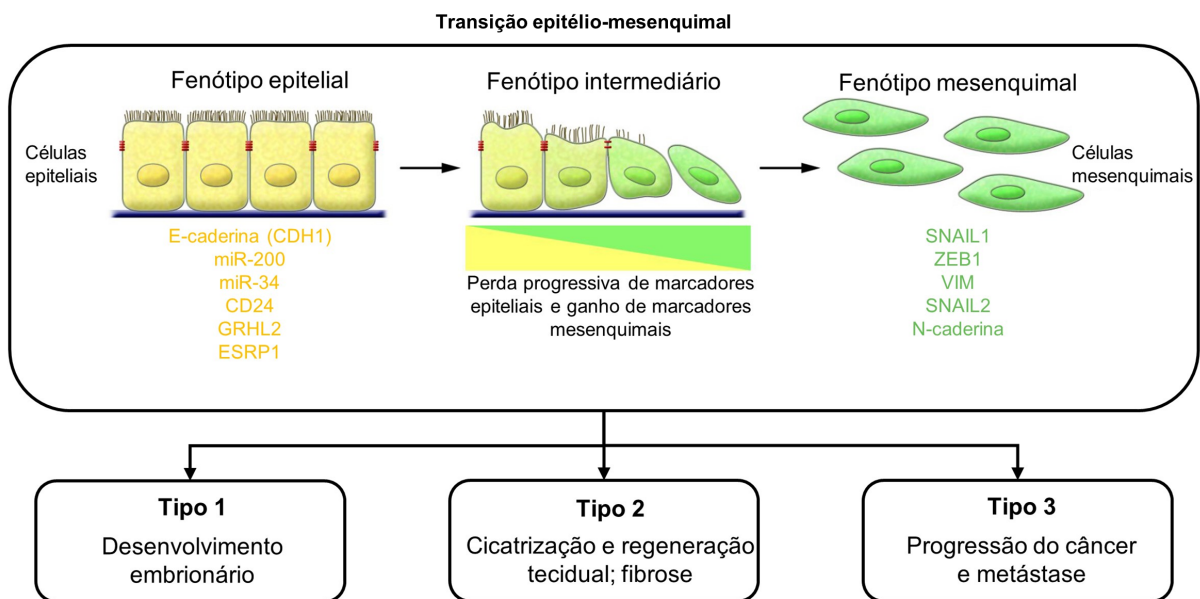
A figura 1.1 resume o processo da TEM através de seus mecanismos associados aos aspectos celulares e moleculares, assim como seus tipos específicos. Além dos fenótipos epitelial e mesenquimal apresentados, um estado intermediário entre tais tipos celulares chamado híbrido também pode ser temporariamente estável. Tal estado é uma combinação de características epiteliais e mesenquimais que podem promover células cancerosas a moverem-se coletivamente, por exemplo (MICALIZZI; FARABAUGH; FORD, 2010; REVENU; GILMOUR, 2009).

De acordo com a literatura, o estado híbrido de células cancerosas foi considerado metaestável, ou seja, tal fenótipo seria transitório durante o processo de TEM (ARNOUX et al., 2005; SAVAGNER, 2015). No entanto, verificou-se que algumas proteínas como OVOL e GRHL2 (denotadas como fatores de estabilidade fenotípica), podem fazer com que o fenótipo híbrido permaneça estável (JOLLY et al., 2016). De fato, dados experimentais recentes demonstraram que o estado híbrido foi observado no processo da TEM (LI et al., 2010; SAMAVARCHI-TEHRANI et al., 2010). Como resultado dessa estabilização, essas células tumorais invadem a membrana basal, entram nos vasos sanguíneos periféricos e se tornam células tumorais circulantes (CTCs) (GARG, 2017), as quais podem apresentar até 50 vezes maior potencial metastático do que células migratórias individuais que apresentam um completo fenótipo mesenquimal (ACETO et al., 2014). Isso ocorre, pois as células híbridas podem se mover



em conjunto através de aglomerados, os quais facilitam a aderência de tais células em outros tecidos.

Figura 1.1 – Representação esquemática do processo da transição epitélio-mesenquimal e seus respectivos tipos. Tal transição é o mecanismo no qual células epiteliais transitam para um fenótipo mesenquimal, ocorrendo em três diferentes tipos de processos biológicos: desenvolvimento embrionário (tipo 1); Cicatrização, regeneração tecidual e fibrose (tipo 2); progressão do câncer e metástase (tipo 3). Um fenótipo intermediário apresentando ambas características (epitelial e mesenquimal) também é observado. Os aspectos moleculares que governam a transição estão associados à perda de marcadores epiteliais (E-caderina, miR-200, miR-34, CD24, GRHL2 e ESRP1) e o ganho de moléculas características do fenótipo mesenquimal (SNAIL1, ZEB1, VIM, SNAIL2 e N-caderina).



Fonte: Adaptado de Kalluri e Weinberg (2009).

Geralmente, as vias de sinalização que desencadeiam o processo da TEM convergem em uma rede de regulação genética (RRG) comum, na qual contempla proteínas que mantêm as características epiteliais, como a expressão de E-caderina, cuja perda é uma característica importante da TEM (YANG; WEINBERG, 2008), e induzem características mesenquimais, tais como a expressão de VIM, ZEB1 e SNAIL1 (MENDEZ; KOJIMA; GOLDMAN, 2010). Tais redes de regulação são importantes para o entendimento sistemático das reações que governam os processos celulares.

## 1.2 REDES DE REGULAÇÃO GENÉTICA

O entendimento dos efeitos da dinâmica dos mecanismos genéticos que governam as características fenotípicas dos organismos é um dos objetivos da ciência

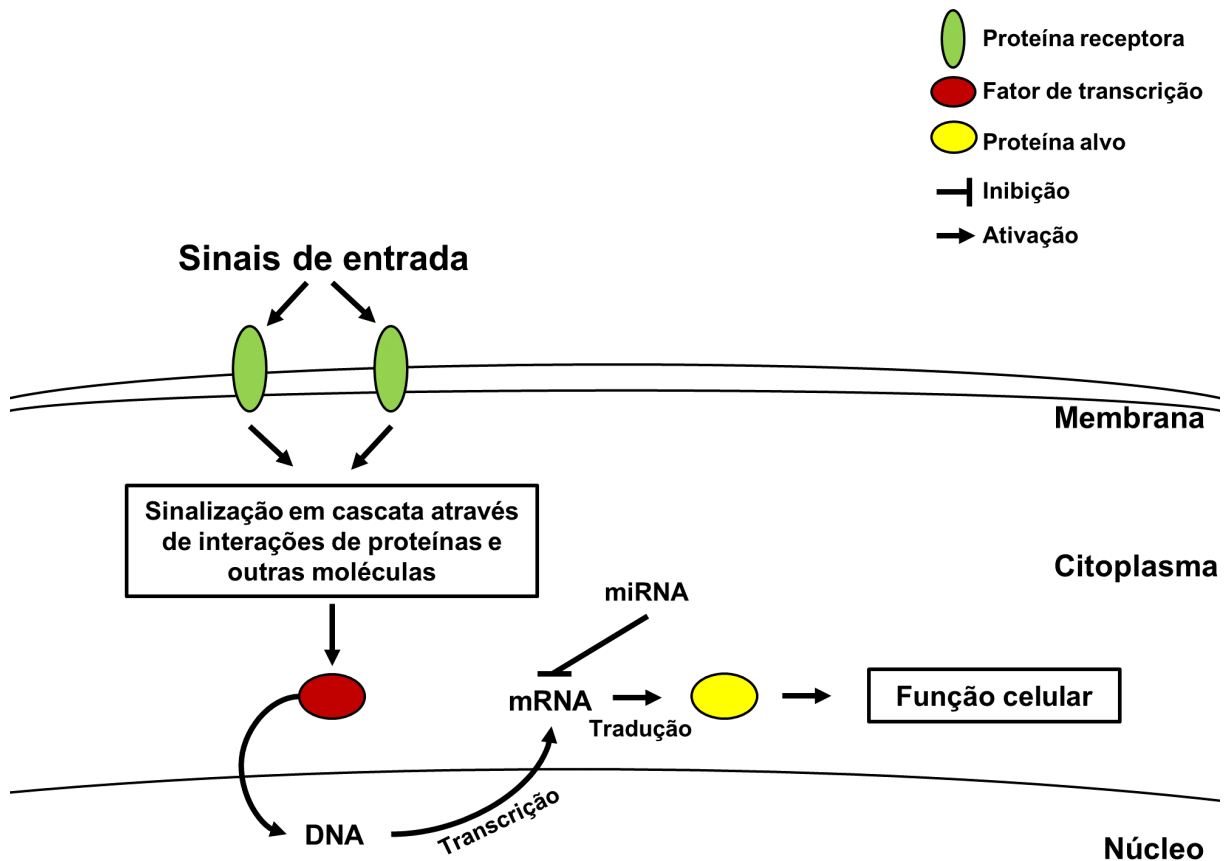
atual. Tal entendimento objetiva especificar as componentes genéticas, como genes e proteínas, a partir de tecnologias que analisam de expressão gênica e identificam as interações entre essas moléculas caracterizando as redes de regulação genética (DAVIDSON; LEVIN, 2005). Adequadas para modelagem e simulação de forma qualitativa e quantitativa e, portanto, utilizadas por biólogos e cientistas computacionais, as redes genéticas têm o potencial de elucidar o efeito da natureza e da topologia das interações nas propriedades sistêmicas dos organismos.

As redes de regulação genética são compostas por nós e arestas. Os nós são os genes e seus reguladores, e as arestas são as relações físicas entre os nós. Os reguladores consistem principalmente em dois tipos: fatores de transcrição (FTs) e miRNAs (SHALGI et al., 2007). Os FTs são proteínas que se ligam a sequências específicas do DNA para regular expressões de genes alvo ao nível transcricional (LATCHMAN, 1997). Diferentemente das redes de interação proteína-proteína, as RRGs são bipartidas devido à existência de seus dois tipos de nós (WALHOUT, 2011), e são direcionais porque os reguladores controlam a expressão dos genes. A figura 1.2 apresenta um exemplo de RRG, cuja sinalização de entrada extracelular é mediada por proteínas receptoras da membrana e tal sinal é conduzido através de interações de moléculas no citoplasma de maneira a ativar um FT (GITTER et al., 2013). Assim, esse fator migra para o núcleo celular com o objetivo de se ligar a uma sequência específica do DNA de para induzir a expressão de um gene alvo (LI; VARALA; CORUZZI, 2015). A partir disso, tal gene é transcrito em um mRNA que, por sua vez, é traduzido em uma proteína alvo, resultando na ativação de determinada função celular. Além disso, o processo de tradução pode ser interrompido pela presença de miRNAs inibindo a tradução da molécula de mRNA (LING; FABBRI; CALIN, 2013).

Além dos processos de transcrição e tradução, outras interações podem ser observadas em uma RRG, incluindo os mecanismos de oxidação, fosforilação, desfosforilação, ligação molecular e transporte. A oxidação está associada à uma interação de perda de elétrons de uma molécula causada normalmente por espécies reativas de oxigênio (moléculas instáveis presentes no organismo e extremamente reativas) (RAY; HUANG; TSUJI, 2012); fosforilação envolve a adição de um grupo fosfato a uma proteína ou outra molécula (RUBIN; ROSEN, 1975); o processo inverso a fosforilação é a desfosforilação, na qual consiste no ato de remover um grupo fosfato de uma molécula (DENU et al., 1996); ligação molecular é uma associação estável entre duas moléculas através de uma interação atrativa (SMITH et al., 2009); por fim, transporte refere-se ao ato de uma proteína transportar uma molécula para outras regiões da célula (XIAO et al., 2007).

Dependendo do processo biológico em análise, uma RRG pode apresentar uma complexa diversidade de variáveis em termos dos nós e interações que integram tal rede. A identificação e caracterização dos motivos funcionais, que fazem parte da

Figura 1.2 – Exemplo de uma rede de regulação genética. Os sinais de entrada da rede produzem uma sinalização em cascata através de proteínas e outras moléculas ativando um FT que, por sua vez, liga-se ao DNA de maneira a induzir a expressão de um mRNA. Dessa forma, o mRNA pode ser traduzido em uma proteína alvo para executar determinada função celular. Adicionalmente, o processo de tradução do mRNA pode ser inibido pela presença de um miRNA correspondente.



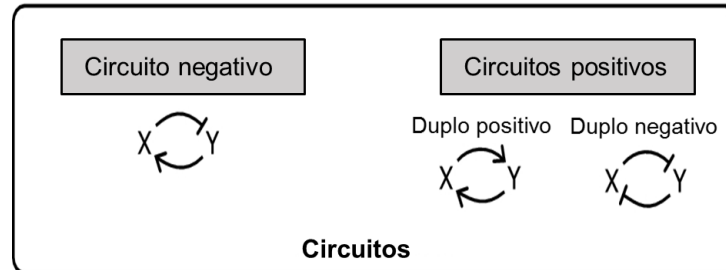
Fonte: Próprio autor.

RRG, tem sido recomendada como uma técnica eficaz para analisar o comportamento dinâmico da rede (HALLINAN; JACKWAY, 2005). Os motivos funcionais são pequenos subconjuntos de nós com um dado padrão de conectividade (BANZHAF; KUO, 2004; MILO et al., 2002). Uma das classes mais estudadas são chamadas de *feedback loops* ou circuitos (CHEN; HUANG; LIN, 2017; LIU; ALBERGANTE; NEWMAN, 2017).

Um *feedback loop* define um circuito fechado de interações de um determinado número de componentes, em que a conexão que gera tal trajetória fechada se origina e termina no mesmo nó. Além disso, os circuitos podem ser definidos como positivos ou negativos, dependendo da paridade do número de interações negativas no *loop* (HALLINAN; JACKWAY, 2005). Assim, um número ímpar de interações negativas no *loop* define um circuito negativo, enquanto um número par estabelece um circuito positivo (figura 1.3).

A atividade funcional do circuito na dinâmica da RRG depende diretamente de

Figura 1.3 – Tipos de circuitos celulares. Os circuitos negativos são representados por um número ímpar de interações negativas (⊖), enquanto que os positivos são caracterizados por um número par dessas interações e, também, pela presença somente de interações ativadoras (→) como, por exemplo, no caso duplo positivo.



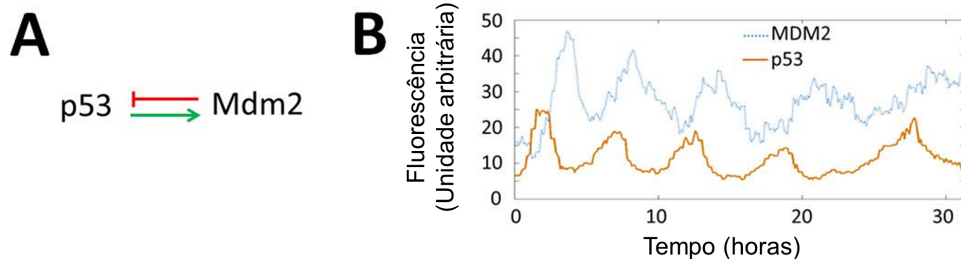
Fonte: Adaptado de Dalchau et al. (2018).

seu sinal. Circuitos negativos desencadeiam uma dinâmica oscilatória (PIGOLOTTI; KRISHNA; JENSEN, 2007), enquanto positivos produzem multiestabilidade celular (KAUFMAN; SOULE; THOMAS, 2007).

Um exemplo de circuito negativo muito estudado no meio científico é o p53/Mdm2 (figura 1.4A), no qual consiste na proteína p53 transcricionalmente ativando Mdm2, que, por sua vez a degrada (MOLL; PETRENKO, 2003). Denotada de proteína "guardiã do genoma", p53 é uma das moléculas supressoras de tumor mais estudadas (CHOI et al., 2019; HAAPANIEMI et al., 2018; HOLLSTEIN et al., 1991), e desempenha um papel fundamental na reparação de danos no DNA ou na indução da apoptose (morte celular programada). Em resposta à estresse celular, tal circuito pode ser ativado resultando em pulsos ou oscilações sustentadas de p53 (figura 1.4B), de acordo com dados quantificados de fluorescência relacionados à expressão dessas moléculas (SHIN et al., 2013). Tais dados foram obtidos a partir de células de câncer de mama (MCF7) com marcadores fluorescentes para p53 e Mdm2 para detectar a atividade dessas moléculas. As concentrações e a dinâmica de ambas as proteínas marcadas com fluorescência foram estudados através de Western blot, que é uma técnica para detectar proteínas específicas em uma determinada amostra.

Em relação ao circuitos positivos, um exemplo biológico que representa as suas funcionalidades é associado à TEM. De acordo com Zhang et al. (2014), a dinâmica da TEM pode ser descrita a partir da rede regulatória apresentada na figura 1.5A, na qual tem como sinais de entrada TGF- $\beta$  de forma exógena. Tal sinalização pode gerar a ativação tanto de marcadores mesenquimais quanto epiteliais, a partir da sub-rede composta por SNAIL1, miR-34, ZEB1, miR-200 e TGF- $\beta$  de forma endógena; sub-rede na qual apresenta, em sua maioria, circuitos positivos (ZEB1/miR-200; SNAIL1/miR-34). A funcionalidade desses circuitos foi observada experimentalmente a partir da linhagem de células epiteliais da mama humana MCF10A tratadas com a proteína hTGF- $\beta$ 1 (ZHANG et al., 2014). Células tratadas com concentrações como 0,5 ng/ml

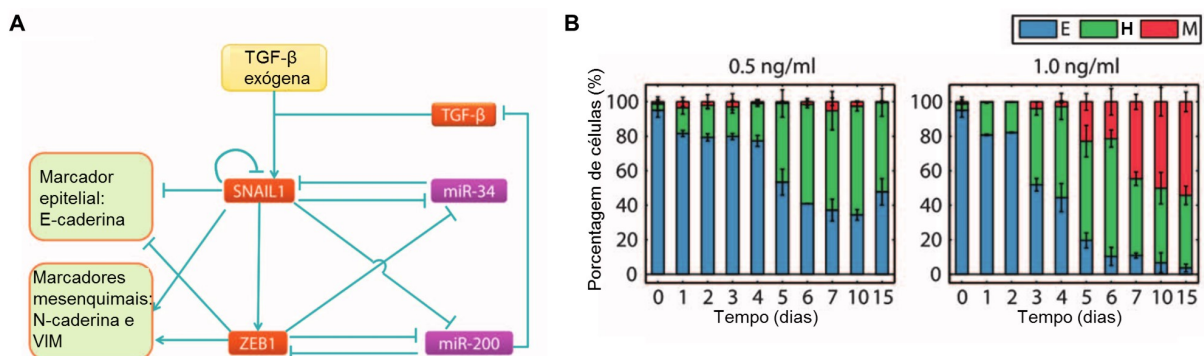
Figura 1.4 – Exemplo do circuito negativo p53/Mdm2 e sua respectiva atividade funcional observada experimentalmente. (A) O circuito é formado através da p53 transcricionalmente ativando Mdm2, que, por sua vez, a degrada. (B) Conforme dados quantificados de fluorescência desse circuito, observa-se um comportamento oscilatório relacionado à expressão dessas moléculas.



Fonte: Adaptado de Shin et al. (2013).

e 1 ng/ml de hTGF- $\beta$ 1 apresentaram uma dinâmica tri-estável, a qual é representada pela estabilização nos diferentes estados da TEM: epitelial (E), híbrido (H) e mesenquimal (M) (figura 1.5B). Dessa forma, esses dados sugerem que o estado final das células epiteliais não é completamente determinado pelas condições iniciais impostas e, assim, indicando que a dinâmica de transição entre os fenótipos é estocástica.

Figura 1.5 – Exemplo dos circuitos positivos SNAIL1/miR-34 e ZEB1/miR-200 aplicados a uma rede da transição epitélio-mesenquimal e suas respectivas atividades funcionais observadas experimentalmente. (A) A atividade de tais circuitos é observada na rede proposta por Zhang et al. (2014), que contém, além dessas moléculas, TGF- $\beta$  de forma exógena (sinal de entrada) e endógena, e respectivos marcadores epiteliais (E-caderina) e mesenquimais (N-caderina e VIM). (B) Experimentalmente, a funcionalidade desses circuitos foi observada a partir linhagem celular MCF10A tratada com concentrações de 0,5 e 1,0 ng/ml de hTGF- $\beta$ 1. Dessa forma, evidenciou-se a presença de porcentagens de células nos estados epitelial (E), híbrido (H) e mesenquimal (M) para diferentes tempos (dias).



Fonte: Adaptado de Zhang et al. (2014)

Devido a complexidade das RRGs em termos de suas interações, nós e mecanismos funcionais, uma análise experimental sobre tal rede torna-se uma abordagem

com um alto padrão de dificuldade. De maneira a favorecer a análise de dados gerados por esses tipos de redes, a modelagem computacional tem sido um método muito utilizado no meio científico atual (GUPTA; SILVEIRA; MOMBACH, 2018; JIA et al., 2017; SILVEIRA; MOMBACH, 2020; SILVEIRA; GUPTA; MOMBACH, 2020b; TIAN; ZHANG; XING, 2013).

### 1.3 MODELAGEM DE REDES REGULATÓRIAS: FORMALISMO LÓGICO

A compreensão da bioquímica de doenças complexas, como um estudo de processos e princípios químicos em organismos vivos, é fundamental para o desenvolvimento de tratamentos terapêuticos. O reducionismo e reconstrução gradual dos sistemas de moléculas que governam as funções de tais doenças é considerada uma das abordagens mais eficientes na bioquímica através do acúmulo de conhecimento de fatos elementares, incluindo transformações químicas (CALZONE; BARILLOT; ZINOVYEV, 2018). A modelagem de tais sistemas pode facilitar ao pesquisador a organização e, conseqüentemente, o entendimento dessas informações de forma sistêmica.

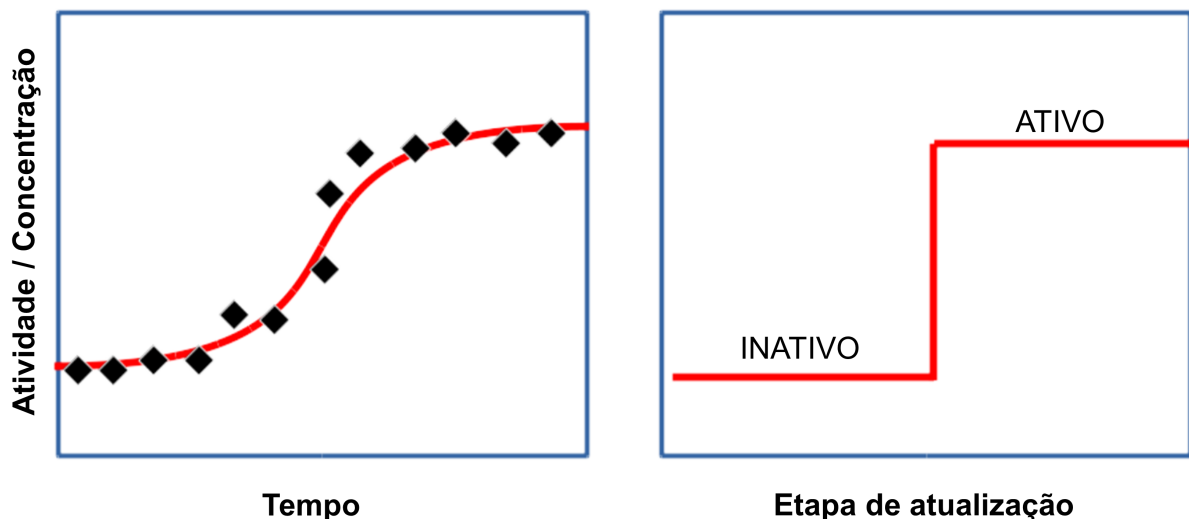
Representações dinâmicas da expressão gênica podem ser utilizadas através de equações diferenciais, em que variáveis específicas permitem representar mudanças nos padrões de expressão gênica em termos espaciais e temporais (ARNOSTI; AY, 2012). Entretanto, essa abordagem requer o conhecimento de todos os parâmetros cinéticos da dinâmica, representando uma restrição para a modelagem devido a dificuldade na determinação de tais parâmetros (JONOSKA; SAITO, 2013). Assim, modelagens relacionadas à abordagens mais simples têm crescido no cenário científico (NOVERE, 2015). Atualmente, a modelagem discreta associada ao formalismo lógico tem sido aplicada para uma variedade de problemas biológicos e, assim, tornou-se um método bem sucedido para modelar e analisar redes reguladoras (GUPTA; SILVEIRA; MOMBACH, 2018; GUPTA; SILVEIRA; MOMBACH, 2020a; GUPTA et al., 2020; GUPTA; SILVEIRA; MOMBACH, 2020b; ISSLER; MOMBACH, 2017; REMY et al., 2015; ROSSATO et al., 2019; SÁNCHEZ; CHAOUIYA, 2016; SILVEIRA; MOMBACH, 2020; SILVEIRA; GUPTA; MOMBACH, 2020b; SILVEIRA; GUPTA; MOMBACH, 2020a).

Os estudos de Kauffman (1969) e Thomas (1973) foram os pioneiros no campo de pesquisa associado à modelagem lógica de redes de regulação genéticas. Ambos convergiram em mostrar que os estados estáveis gerados por tal modelagem podem ser associados a diferentes tipos de células, e que as transições entre os estados lógicos podem ser relacionadas à mudanças na expressão de genes. Além disso, os conceitos formulados por Thomas (1973), em termos da dinâmica de circuitos, são

válidos para outros tipos de modelos como os contínuos (SOULÉ, 2006). Assim, esses estudos foram a base para uma série de pesquisas que abrangem desde estudos teóricos até modelos de redes que controlam uma variedade de processos celulares (TONELLO; FARCOT; CHAOUIYA, 2019), demonstrando a versatilidade e a importância da modelagem lógica na biologia molecular.

A modelagem de redes de regulação genética incorpora o conhecimento experimental sobre os elementos moleculares (chamados de nós) e suas interações; a condição inicial do sistema que leva ao estado final, além do comportamento dinâmico da rede. Modelos validados experimentalmente são capazes de investigar casos complexos e podem levar a previsões sobre o funcionamento do sistema (VIJESH; CHAKRABARTI; SREEKUMAR, 2013). No contexto da modelagem lógica, aos genes (ou componentes reguladores) são atribuídos a valores discretos que respondem por seus níveis funcionais de expressão (ou atividade). Uma função reguladora estabelece a evolução do nível do gene, dependendo dos estados dos nós que controlam a sua atividade. Tal modelo define um sistema dinâmico em que o estado da rede (os níveis de cada componente) evolui de uma forma discreta. Essa representação abstrata dos mecanismos moleculares é muito conveniente para lidar com redes que carecem de dados cinéticos precisos (CHAOUIYA; REMY, 2013). A figura 1.6 apresenta a associação de concentração ou atividade de uma determinada molécula relativa ao tempo e sua abstração discreta.

Figura 1.6 – Relação entre dados experimentais de atividade ou concentração de uma determinada molécula e a sua correspondente abstração discreta.



Fonte: Adaptado de Pais (2018).

Matematicamente, a construção do modelo lógico de uma rede regulatória  $(G, K)$  é conduzida através de informações bioquímicas publicadas na literatura, as quais são representadas por componentes regulatórias discretas  $G = (g_1, g_2, \dots, g_n)$  e interações

entre elas (ativadoras ou inibidoras), caracterizadas por arcos dirigidos. Uma função lógica  $K_i$  define os valores de cada  $g_i$  em um intervalo de  $\{0, \dots, \max_i\}$ , em que  $\max_i$  é o valor máximo definido para  $g_i$ . Essas funções são baseadas na influência dos nós reguladores de  $g_i$ , e são construídas pelos operadores lógicos AND, OR e NOT para representar as interações bioquímicas encontradas na literatura. O conjunto das funções lógicas do modelo é definido pela função de transição  $K(g) = (K_1(g), \dots, K_n(g))$ .

No formalismo lógico, a dinâmica do modelo é definida por seu espaço de estados  $S$ , o qual é representado por um gráfico de transições de estados (GTE), cujos nós denotam vetores  $g = [g_1, g_2, \dots, g_n]$  (representando estados do modelo) e cujas arestas representam transições entre esses vetores. Além disso, cada aresta é caracterizada por uma componente do modelo que é atualizada durante a transição. Resumidamente, a rede de transição de estados representa todas as trajetórias possíveis que um estado inicial pode levar a um estado final. Os vetores terminais que não apresentam estados sucessores são chamados de atratores estáveis, enquanto um conjunto de vetores terminais que conduzem uma trajetória cíclica entre eles compõe um atrator cíclico (ABOU-JAOUDE et al., 2016). O aprisionamento do sistema em um estado cíclico em termos matemáticos é apresentado no apêndice A. Adicionalmente, a transição entre os vetores é definida de forma assíncrona ou sincronizada. Na atualização assíncrona, nós aleatórios são selecionados a cada momento e, em seguida, atualizados, enquanto que os nós evoluindo de forma síncrona são atualizados ao mesmo tempo. Em geral, os sistemas biológicos apresentam um comportamento estocástico em termos de suas evoluções dinâmicas. Assim, o esquema assíncrono é considerado mais realístico através de sua dinâmica estocástica (SÁNCHEZ; CHAOUIYA, 2016; MENDES et al., 2018).

De acordo com a atualização assíncrona, cada estado  $g$ , da componente do modelo  $i$ , dependendo do tempo  $t$ , evolui ao longo de cada etapa de tempo  $t + 1$  como:

$$g_i(t+1) = g_i(t) + \text{sign}[K_i(g_i(t)) - g_i(t)], \quad (1.1)$$

$$g_j(t+1) = g_j(t), \quad (1.2)$$

em que  $\text{sign}[K_i(g_i(t)) - g_i(t)]$  é a função sinal dependente da função lógica  $K_i(g_i(t))$  e estado inicial  $g_i(t)$ . A função sinal  $\text{sign}(p)$ <sup>1</sup> é igual a 1 se  $p > 0$ , -1 se  $p < 0$  e  $\text{sign}(0) = 0$ , para  $p = 0$ . O índice  $j$  representa os componentes que não foram selecionados para o esquema de atualização.

Um exemplo de aplicação do formalismo lógico em uma rede regulatória é apresentado na figura 1.7. A rede é representada por um grafo, no qual os nós são as

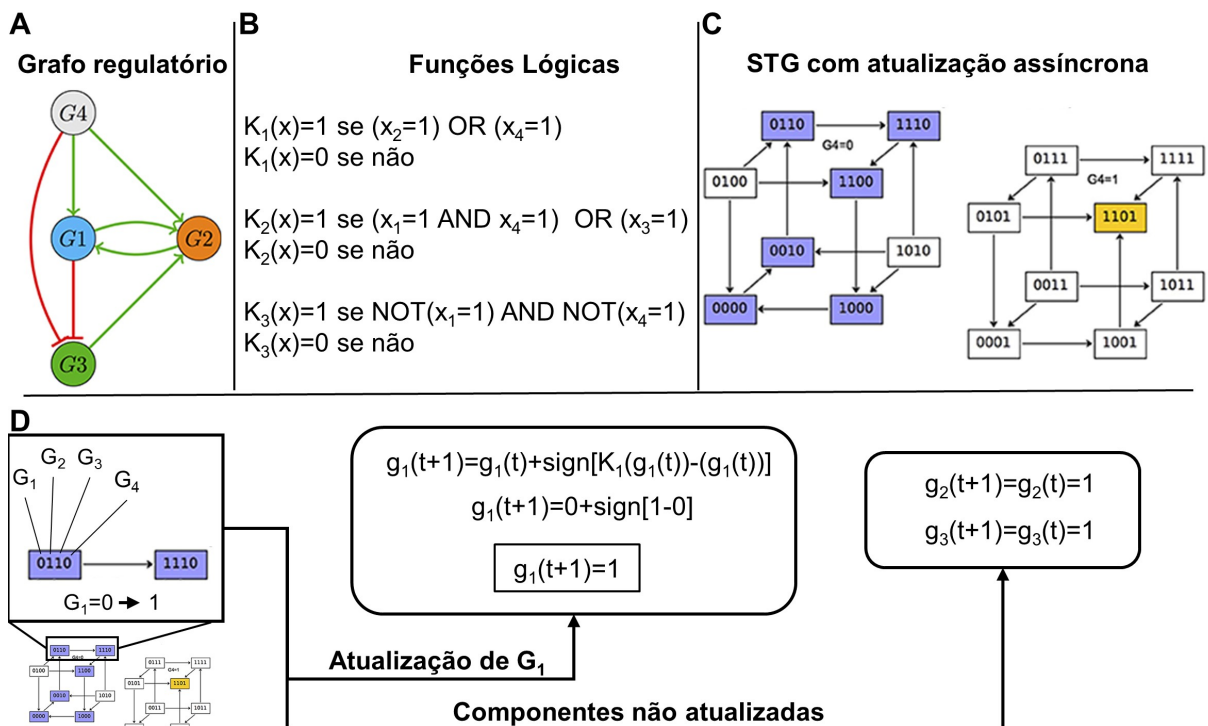
---

<sup>1</sup> $p$  é o argumento da função.



componentes e as arestas são interações; as cores vermelha e verde denotam interações inibidoras e ativadoras, respectivamente (figura 1.7A). As figuras 1.7B e 1.7C apresentam as funções lógicas que controlam cada componente e o GTE relacionado ao método assíncrono do correspondente grafo. Um atrator cíclico (roxo) e estável (amarelo) podem ser observados de acordo com o valor da componente de entrada  $G_4$ . Além disso, as aplicações das eqs. 1.1 e 1.2 são apresentadas para uma das transições do GTE (figura 1.7D).

Figura 1.7 – Exemplo de aplicação do formalismo lógico em uma rede de regulação genética. (A) O grafo regulatório é representado por interações entre as componentes  $G_1$ ,  $G_2$ ,  $G_3$  e  $G_4$  (sinal de entrada). (B) As respectivas funções lógicas controlam os estados dos nós da rede em A. (C) A partir das funções em B, o correspondente gráfico de transição de estados com atualização assíncrona pode ser gerado. De acordo com o estado do sinal de entrada ( $G_4$ ), dois atratores são observados: cíclico (roxo) e estável (amarelo). (D) A aplicação das equações assíncronas podem ser observadas em C para uma das transições.



Fonte: Adaptado de Abou-Jaoudé et al. (2016)

A dinâmica de uma rede regulatória pode ser controlada pelos circuitos que a compõem. Para testar o efeito deles na dinâmica, a abordagem lógica permite perturbações *in silico* nos valores das componentes do modelo, as quais restringem o valor da componente do modelo para permanecer fixa. Através de uma perda de função (PF), a componente é restringida a permanecer no valor zero (desligada), enquanto um ganho de função (GF) estabelece a permanência em um valor maior que zero, de acordo com valores definidos para tal componente.

Embora a atualização assíncrona descreva um comportamento mais realístico da dinâmica dos sistemas celulares, tal método pode ser uma abordagem complexa para redes com um grande número de componentes. Assim, definir a acessibilidade dos atratores de um GTE com inúmeros vetores pode determinar uma análise complexa sobre o sistema (ver abaixo). De maneira a facilitar o acesso a essa informação, um algoritmo chamado Avatar foi proposto por Mendes et al. (2018), que fornece estimativas de probabilidades de acessibilidade de atratores em dinâmicas assíncronas. Tal método explora iterativamente o espaço de estados selecionando aleatoriamente trajetórias para estimar a probabilidade de alcançar um atrator a partir dos estados iniciais. Os princípios operacionais do Avatar baseiam-se no cálculo das probabilidades a partir do método de Monte Carlo e na identificação do armadilhamento em ciclos transientes através da abordagem de cadeias absorventes de Markov (MENDES et al., 2018).

A simulação de Monte Carlo é classicamente usada para estimar a probabilidade de um resultado através de amostragens aleatórias massivas para obter resultados numéricos utilizando sucessivas simulações em um elevado número de vezes (KROESE et al., 2014). No contexto da modelagem lógica, isso significa seguir caminhos aleatórios ao longo da dinâmica assíncrona no GTE finalizando nos estados terminais. Ao realizar um grande número de simulações, é possível avaliar a probabilidade de alcance desses estados. No entanto, na presença de um ciclo transitório, a simulação pode voltar a visitar os mesmos estados um número ilimitado de vezes antes de escapar do aprisionamento (MENDES et al., 2018). Assim, um aprimoramento do método foi proposto associado às cadeias absorventes de Markov. Um GTE pode ser interpretado como uma matriz de transição de uma Cadeia de Markov finita (MENDES et al., 2018). Além disso, essa representação pode abranger atratores distintos e, portanto, definir cadeias absorventes (GRINSTEAD; SNELL, 2012). Assim, o método de cadeias de Markov pode ser aplicado nesse tipo de contexto.

### 1.3.1 Algoritmo Avatar

As probabilidades de transição entre os estados são assumidas serem uniformemente distribuídas. Assim, considerando uma transição de um vetor  $v$  para um  $v'$  em um  $S$  qualquer, a probabilidade será:

$$\Pi(v, v') = \frac{1}{|Succ(v)|} \quad (1.3)$$

em que  $Succ(v)$  é o conjunto de estados sucessores de  $v$ . Caso  $v$  seja um estado absorvente, isto é  $Succ(v) = \emptyset$ <sup>2</sup>, então  $\Pi(v, v) = 1$  (MENDES et al., 2018).

Em termos descritivos, o método de cadeias de Markov baseia-se em uma cadeia de estados que pode atingir um estado absorvente, que, uma vez alcançado, não pode transitar para outro. A distribuição de probabilidade de transições de estado é tipicamente representada como a matriz de transição da cadeia. Para  $N$  estados acessíveis, a matriz será definida como  $N \times N$ , tal que cada elemento  $(i, j)$  é a probabilidade de transição do estado  $i$  para o estado  $j$  (GAGNIUC, 2017). Mais detalhes teóricos sobre cadeias de Markov são apresentados no apêndice B.

A partir da identificação de um ciclo transiente, o Avatar remodela o GTE (MENDES et al., 2018). Assim, a remodelagem do GTE pode ser descrita em termos de dois conjuntos: os caminhos percorridos ao longo do ciclo ( $C = c_1, c_2, \dots, c_k$ ) e os estados diretamente alcançados por  $C$  através da definição de um conjunto  $B$ . Assim, a matriz de transição  $\pi$  pode ser escrita em termos de duas sub-matrizes  $q$  e  $r$ . A matriz  $q$  define as transições entre as componentes do ciclo, enquanto que  $r$  representa as transições de  $C$  para  $B$ . Dessa forma, a matriz de transição é modificada da seguinte maneira:

- (a) Remove as transições entre os estados de  $C$ , substituindo a sub-matriz  $q$  por

$$q^1 = 0, \quad (1.4)$$

a qual representa a matriz nula;

- (b) Adiciona um arco de cada estado de  $C$  para cada estado de  $B$ , substituindo a sub-matriz  $r$  por

$$r^1 = [(I - q)^{-1}r], \quad (1.5)$$

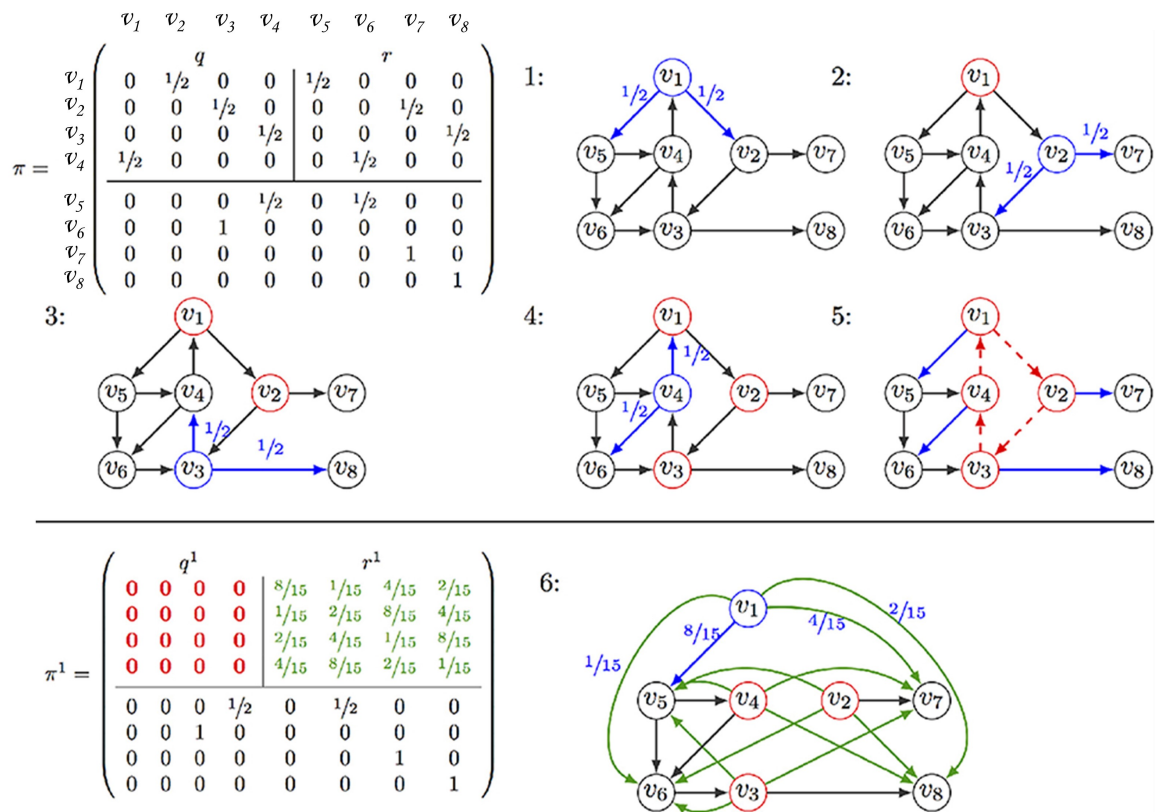
em que  $I$  é a matriz identidade.

A figura 1.8 representa um exemplo da aplicação do método de Avatar. Nessa ilustração, um GTE composto por um conjunto de oito estados ( $v = \{v_1, v_2, \dots, v_8\}$ ) apresenta um ciclo transiente entre os vetores  $v_1, v_2, v_3$  e  $v_4$ . Durante a simulação, a matriz  $q$  é definida a partir das componentes do ciclo, enquanto que  $r$  denota as transições de  $v_1, v_2, v_3$  e  $v_4$  para os seus respectivos vetores alvo ( $v_5, v_6, v_7$  e  $v_8$ ). Na etapa 5 da simulação, o algoritmo detecta o ciclo e remodela o GTE a partir das eqs. (1.4) e (1.5). Na etapa 6, um vetor sucessor a  $v_1$  é selecionado aleatoriamente a partir das novas probabilidades calculadas para cada trajetória. Essa remodelagem não altera as propriedades assintóticas de saída da simulação (MENDES et al., 2018).

---

<sup>2</sup> $\emptyset$  é o conjunto vazio.

Figura 1.8 – Exemplo da aplicação do método de Avatar. De acordo com o respectivo gráfico de transição de estados (GTE). A matriz de transição  $\pi$  é particionada nas sub-matrizes  $q$  (transições entre  $v_1, v_2, v_3$  e  $v_4$ ) e  $r$  (transições levando aos vetores do ciclo). A simulação inicia (etapa 1) de  $v_1$  (simbolizando com contorno azul o vetor selecionado), que pode transitar para  $v_2$  ou  $v_5$  com probabilidade  $1/2$  para ambos. Assim,  $v_2$  é selecionado (etapa 2), indicando  $v_1$  com contorno vermelho como sendo um estado já visitado. A simulação prossegue dessa mesma forma (etapas 3 e 4) até revisitar  $v_1$  (etapa 5). Identificando um ciclo, o procedimento de remodelagem do GTE é iniciado, removendo as transições do ciclo (vermelho pontilhado) e adicionando transições em direção aos vetores de saída do ciclo (verde). As probabilidades são calculadas a partir das eqs. (1.4) e (1.5), resultando em uma nova matriz de transição  $\pi^1$ . A partir de  $v_1$ , um vetor de saída do ciclo é escolhida de acordo com essas probabilidades.



Fonte: Adaptado de Mendes et al. (2018)

Devido a importância da modelagem lógica sobre a biologia sistêmica, softwares foram desenvolvidos para aplicação de tal formalismo em redes de regulação genética (HELIKAR et al., 2012; MÜSSEL; HOPFENSITZ; KESTLER, 2010; NALDI et al., 2018). Um desses algoritmos chamado *Gene Interaction Network simulation suite* (GINsim) tem sido muito utilizado no meio científico por sua interface simples para editar, simular e analisar a dinâmica resultante da rede (NALDI et al., 2018). Além disso, esse software foi recentemente implementado com novas técnicas para estimar a acessibilidade de atratores em modelos lógicos assíncronos como, por exemplo, o

método do Avatar (MENDES et al., 2018).

### 1.3.2 Software GINsim

Experimentos genômicos em escala descobriram progressivamente muitas das numerosas interações que controlam vários processos biológicos cruciais. A complexidade das redes muitas vezes desafia o raciocínio intuitivo, conseqüentemente exigindo o desenvolvimento de ferramentas computacionais adequadas (GONZALEZ et al., 2006).

O software GINsim foi desenvolvido para modelagem qualitativa, análise e simulação de redes de regulação genética desde simples até redes com um alto padrão de complexidade (GONZALEZ et al., 2006). O algoritmo do GINsim é implementado através do formalismo lógico com base em 3 pontos principais:

- (a) uma interface gráfica para construção do grafo regulatório;
- (b) um núcleo de simulação (construção do GTE);
- (c) ferramentas para análise da dinâmica da rede.

Com base nesses itens, dois grafos regulatórios foram construídos para representar os procedimentos metodológicos associado ao software e comparar os resultados obtidos com as observações experimentais apresentadas nas figuras 1.4 e 1.5.

#### 1.3.2.1 *Tradução de uma rede de regulação genética em um modelo lógico no GINsim*

Uma rede de regulação biológica descreve as interações entre as moléculas de um determinado sistema, as quais são representadas por um gráfico de interações cujos vértices denotam moléculas e os arcos entre elas são suas interações. A modelagem de tais sistemas baseia-se na construção do mapa de interações (redes) através de informações de moléculas reguladoras e suas interações relacionadas à via de sinalização de interesse. A ligação entre essas moléculas podem ser interpretadas como diferentes tipos:

- I. transcrição;
- II. tradução;
- III. oxidação;

- IV. fosforilação;
- V. desfosforilação;
- VI. ligação molecular;
- VII. transporte.

Após a construção da rede regulatória, os parâmetros dinâmicos de cada molécula são aplicados através dos operadores lógicos. A figura 1.9 apresenta um exemplo da associação do mecanismo biológico (interação), abstração regulatória no modelo e a regra lógica correspondente. Diferentes funções biológicas como os processos de tradução, fosforilação, desfosforilação e ligação molecular podem ser traduzidos em uma ativação (arco verde) e inativação (arco vermelho) de uma molécula (X). Baseando-se nessa abstração, o uso dos operadores lógicos *AND*, *NOT* e *OR* estabelece os parâmetros dinâmicos da evolução discreta dos estados das moléculas do modelo. Em casos que duas ou mais moléculas são indispensáveis para ativar ou inativar uma determinada proteína ou gene, usa-se o operador *AND*. Em contrapartida, o operador *OR* estabelece um parâmetro independente para cada regulador. Em casos de inibidores, o uso desse operador estabelece que a molécula alvo será ativa se um dos reguladores permanecer inativo. Para regulação de ativação, a molécula alvo será ativa se apenas um dos reguladores permanecer ativo.

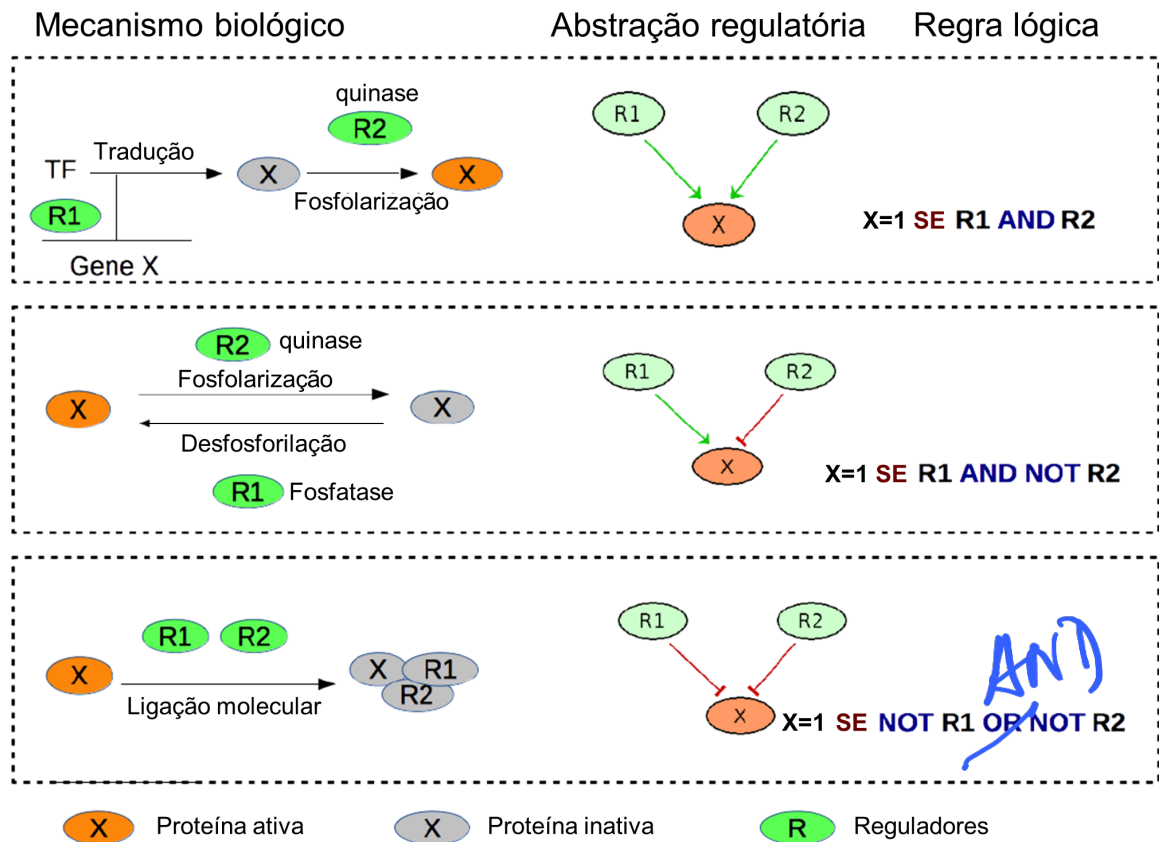
A figura 1.10 apresenta uma rede simplificada relacionada ao mecanismo de ativação do circuito negativo de p53/Mdm2. O dano ao DNA gera estímulos para a ativação de p53 que, por sua vez, induz a ativação de Mdm2, enquanto a mesma inibe p53 (figura 1.10A). As regras e suas respectivas funções lógicas que controlam os estados dos nós são construídas a partir das informações bioquímicas presentes na literatura (figura 1.10B). O GTE do respectivo grafo regulatório apresenta um ciclo terminal (figura 1.10C), que concorda com as observações experimentais observadas na figura 1.4.

A funcionalidade de circuitos positivos na natureza pode ser representada pelo processo da TEM (figura 1.11). Um exemplo é o grafo regulatório proposto por Zhang et al. (2014), que é composto por dois circuitos positivos funcionais: SNAIL1/miR-34 e ZEB1/miR-200 (figura 1.11A). Através da utilização do GINsim, os níveis atribuídos a cada nó devem considerar a concentração de cada componente no processo de acordo com dados experimentais do estudo de Zhang et al. (2014) (figura 1.11B)<sup>3</sup>. Assim, através da método assíncrono, o GTE do correspondente grafo apresentou uma dinâmica tri-estável (figura 1.11C), a qual é representada pelos estados epitelial, híbrido e mesenquimal. Conforme a figura 1.5B, a dinâmica do modelo é consistente com as observações experimentais.

---

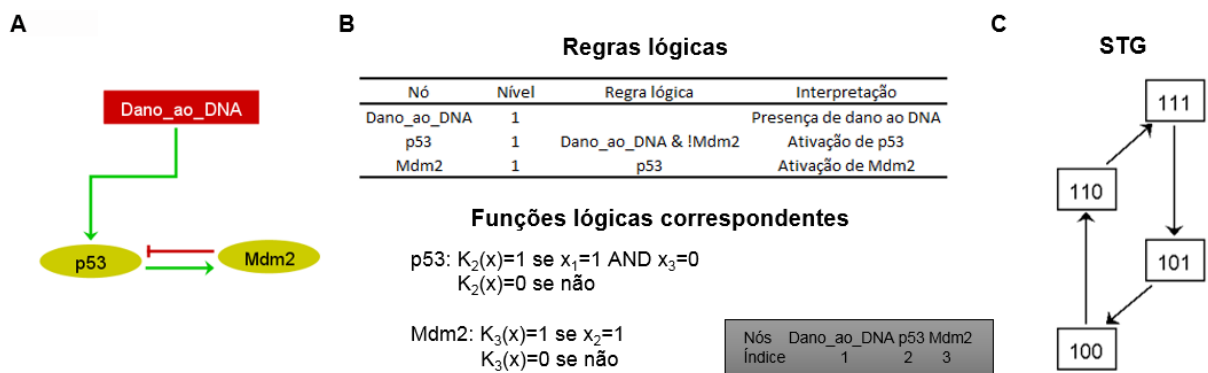
<sup>3</sup>Devido ao maior número, as funções lógicas não foram apresentadas nessa figura.

Figura 1.9 – Relação entre os mecanismos biológicos, abstração regulatória e as regras lógicas correspondentes.



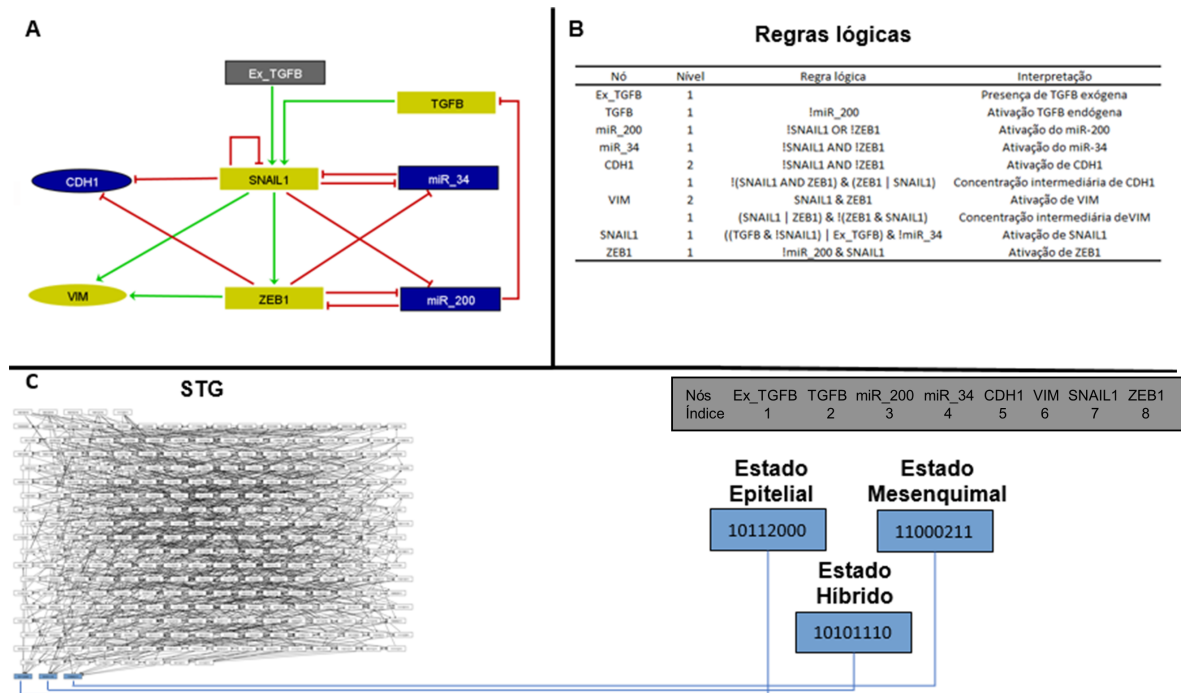
Fonte: Adaptado de Pais (2018).

Figura 1.10 – Exemplo de aplicação da modelagem lógica do circuito negativo p53/Mdm2 no GINsim. (A) Grafo regulatório representado pelo dano ao DNA (sinal de entrada) induzindo a ativação de p53 e, assim, desencadeando a atividade funcional do circuito. (B) As respectivas regras lógicas controlando os estados dos nós, juntamente com suas funções lógicas correspondentes. (C) A partir de B, o gráfico de transição de estados (GTE) de atualização assíncrona é gerado (o índice de cada nó da rede representa a ordem de aparição do seu respectivo estado no vetor do GTE).



Fonte: Próprio autor.

Figura 1.11 – Exemplo de aplicação da modelagem lógica na rede proposta por Zhang et al. (2014) (figura 1.5A) relacionada à transição epitélio-mesenquimal. (A) Grafo regulatório representado pelas moléculas de TGF- $\beta$  de forma exógena (Ex\_TGFB; sinal de entrada) e endógena (TGFB), CDH1 e VIM, além dos circuitos positivos SNAIL1/miR-34 e ZEB1/miR-200. (B) As regras lógicas controlando os estados dos respectivos nós. (C) O gráfico de transição de estados correspondente (assíncrono).



Fonte: Próprio autor.

Além do núcleo de simulação, o GINsim permite definições de perturbações, identificação de circuitos funcionais e redução do modelo lógico (para redes complexas). Essas ferramentas adicionais para a análise da rede possibilitam uma eficiente caracterização das suas propriedades dinâmicas (NALDI et al., 2018).

### 1.3.2.2 Funcionalidade dos circuitos

O comportamento dinâmico das RRGs pode ser definido a partir suas estruturas topológicas através dos circuitos regulatórios. Entretanto, a presença de um deles em uma rede não garante a efetividade de seu comportamento dinâmico correspondente. Para isso ocorrer, o circuito deve ser funcional (THOMAS; D'ARI, 1990). Dessa forma, Naldi, Thieffry e Chaouiya (2007) desenvolveram um método computacional aplicado ao GINsim que determina o contexto de funcionalidade de cada circuito em uma RRG.

De acordo com esse método, um circuito será funcional se as interações que o compõem sejam funcionais, ou seja, elas afetam o valor lógico das suas respectivas



componentes alvo. Essa funcionalidade é diretamente dependente dos reguladores externos do circuito na RRG (NALDI; THIEFFRY; CHAOUIYA, 2007). Assim, o contexto de funcionalidade é um conjunto de restrições sobre os valores das componentes externas que atuam nesse circuito. Tal conjunto é composto pela interseção entre os contextos funcionais de cada interação que constitui o circuito (NALDI; THIEFFRY; CHAOUIYA, 2007). A ausência dessa interseção implica que o circuito não seja funcional. A relação entre as regras lógicas e contexto de funcionalidade dos circuitos é apresentada no apêndice C.

A modelagem discreta de redes biológicas objetiva uma compreensão quantitativa do comportamento dinâmico em função de suas componentes e suas interações. Os modelos computacionais desempenham um papel importante integrando o comportamento das componentes e analisando conjuntos de dados experimentais. Com a grande quantidade de dados produzidos na literatura, banco de dados foram criados para facilitar o acesso dessas informações pelo pesquisador.

### 1.3.3 Banco de dados em biologia sistêmica

A biologia sistêmica objetiva estudar as respostas biológicas dos elementos de uma rede de interações através de perturbações genéticas ou ambientais. Dessa forma, modelos computacionais são criados permitindo a simulação *in silico* desses elementos. A natureza dessa abordagem sistêmica requer a integração de dados estruturais da rede como nós e interações, e os parâmetros dinâmicos entre tais moléculas.

Os dados estruturais da rede podem ser estudados através de banco de dados específicos que fornecem informações de cadeias de reações químicas de uma via genética. Entre os mais utilizados estão o *Kegg pathway database* e o *Reactome* (FABREGAT et al., 2018; KANEHISA et al., 2021). Ambos apresentam informações bioquímicas recentes de interações entre moléculas e são repositórios de acesso aberto ao pesquisador.

O KEGG permite a reconstrução computacional de sistemas biológicos a partir de informações do genoma (KANEHISA et al., 2021). Tal repositório de dados é atualizado a partir de dados experimentais publicados na literatura, representados em termos de interações molecular e redes. O Kegg também fornece ferramentas de mapeamento, as quais permitem a compreensão das funções celulares a partir de sequências do genoma e outros conjuntos de dados moleculares. Além disso, esse banco de dados abrange uma grande variedade de espécies.

O Reactome, por sua vez, é caracterizado por focar seu repositório de dados, na grande maioria, em uma única espécie, *Homo sapiens* (FABREGAT et al., 2018).

Tal banco de dados tem uma grande variedade de vias genéticas que podem ser acessadas para pesquisa básica, análise de genoma, modelagem de vias, biologia de sistemas e educação. Além disso, o Reactome pode ligar sistematicamente proteínas humanas às suas funções moleculares, fornecendo um recurso que pode facilitar o descobrimento de novas relações funcionais em dados como estudos de expressão gênica, por exemplo.

Além dos supramencionados, outro banco de dados chamado *Biological General Repository for Interaction Datasets* (BioGRID) é útil para a construção de modelos computacionais baseados em redes biológicas (OUGHTRED et al., 2021). Tal repositório de acesso aberto contempla proteínas e interações genéticas de várias espécies, incluindo *Homo sapiens*. A investigação de interações entre moléculas é de fácil acesso e facilita na construção das redes regulatórias. Além disso, BioGRID armazena interações de proteínas ou genes com pequenas moléculas bioativas, incluindo muitos medicamentos conhecidos.

Para auxiliar a busca por moléculas nos bancos de dados, o nome e símbolo atualizados de tais genes são importantes. Nesse contexto, o *HUGO Gene Nomenclature Committee* (HGNC) é um comitê do *Human Genome Organization* (HUGO) e tem como objetivo atribuir um nome e um símbolo únicos e idealmente significativos a cada gene humano (POVEY et al., 2001; TWEEDIE et al., 2021). Todos os símbolos aprovados são armazenados no banco de dados de nomenclatura do gene humano (<https://www.genenames.org/>). Além disso, tal repositório fornece informações sobre outros símbolos e nomes que são utilizados na literatura para uma determinada molécula de outros banco de dados.

## 1.4 OBJETIVOS

Este estudo tem como objetivo geral analisar a dinâmica da TEM através da modelagem lógica utilizando informações bioquímicas publicadas recentemente na literatura para contribuir na evolução do conhecimento científico dessa transição. De acordo com o objetivo geral, os objetivos específicos estabelecidos neste estudo foram analisar:

- I. a ativação da TEM em resposta às vias genéticas indutoras de câncer, incluindo TGF- $\beta$ ;
- II. o comportamento dinâmico de miRNAs durante a transição;
- III. a funcionalidade dos circuitos ativados durante a TEM;
- IV. as probabilidades de acessibilidade dos atratores.

Baseados nos objetivos estabelecidos no presente estudo, os dois capítulos subsequentes abordarão os artigos científicos publicados durante a realização do projeto.

## 2 ARTIGO: DYNAMICS OF THE FEEDBACK LOOPS REQUIRED FOR THE PHENOTYPIC STABILIZATION IN THE EPITHELIAL-MESENCHYMAL TRANSITION

Artigo publicado no periódico científico "*The FEBS Journal*"<sup>1</sup> e disponível, na sua versão final, em: <https://doi.org/10.1111/febs.15062>. Tal publicação objetiva descrever a dinâmica relacionada à transição epitélio-mesenquimal (TEM) de acordo com a funcionalidade dos circuitos necessários para a estabilização fenotípica em cada um dos estados estáveis da transição. Os seguintes resultados foram observados nesse estudo:

- I. A simulação de tipo selvagem mostrou uma tri-estabilidade para *input* ativo, a qual corresponde a presença dos estados E, H e M;
- II. Oito circuitos funcionais, dos quais sete são positivos e um é negativo.
- III. O circuito positivo SNAIL1/miR-34 foi responsável pela estabilização do estado E para o H;
- IV. Os circuitos mediados por ZEB1 com GRHL2 e os miRNAs 200, 340 e 1199 foram associados à estabilização do estado H;
- V. ZEB1/HA/HAS2 e ZEB1/ESRP1/CD44s foram responsáveis pela estabilização do estado M;
- VI. O modelo apresentou uma excelente concordância quantitativa com dados experimentais para células de mama MCF10A;
- VII. Os miRNAs 340 e 1199 podem ser SPFs do estado H;
- VIII. As probabilidades de acessibilidade dos atratores para o caso tri-estável concordaram com dados experimentais para doses de 0,5 ng/ml de TGF- $\beta$  aplicadas em células MCF10A.

---

<sup>1</sup>Fator de impacto (2020): 5.542

**This is the accepted version of the following article: SILVEIRA, Daner A.; MOMBACH, José Carlos M. Dynamics of the feedback loops required for the phenotypic stabilization in the epithelial-mesenchymal transition. The FEBS journal, v. 287, n. 3, p. 578-588, 2020, which has been published in final form at <https://doi.org/10.1111/febs.15062>. This article may be used for non-commercial purposes in accordance with the Wiley Self-Archiving Policy (<http://www.wileyauthors.com/self-archiving>).**

# Dynamics of the feedback loops required for the phenotypic stabilization in the epithelial-mesenchymal transition

Daner A. Silveira<sup>1\*</sup> | José Carlos M. Mombach<sup>1\*</sup>

<sup>1</sup>Universidade Federal de Santa Maria, Santa Maria, RS, Brazil

\*Equally contributing authors

## Correspondence

José Carlos M. Mombach, Department of Physics, Universidade Federal de Santa Maria, Santa Maria, 97105-900, Brazil  
Email: [jcmombach@ufsm.br](mailto:jcmombach@ufsm.br)

## Present address

Department of Physics, Universidade Federal de Santa Maria, Santa Maria, 97105-900, Brazil

## Funding information

Conselho Nacional de Desenvolvimento Científico e Tecnológico (145708/2017-9)

**Running Title:** Dynamics of the feedback loops required for...

**Keywords**

EMT, breast cancer, logical model, hybrid phenotype, microRNA

**Abbreviations:** EMT, epithelial-mesenchymal transition; E, epithelial; H, hybrid; M, mesenchymal; CBS, cascading bistable switches; PSF, phenotypic stability factor.

**Abstract**

The epithelial-mesenchymal transition (EMT) is a complex mechanism in which cells undergo a transition from epithelial to mesenchymal phenotypes (there is also an intermediary hybrid state) in response to microenvironmental alterations and aberrant stimuli triggered by molecules such as TGF- $\beta$ . Recent studies in breast cancer progression reported new feedback loops and new participant molecules such as microRNAs 340 and 1199. In this work we propose a logical model of EMT contemplating the influence of these new published molecules on the regulatory core of EMT. The model results were compared with theoretical and experimental data for the human breast epithelial cell line MCF10A presenting excellent agreement. We propose that miRNAs 340 and 1199 should be considered phenotypic stability factors of the hybrid state based on the positive feedback loops they form with ZEB1. In addition, the model allows the prediction of phenotype probabilities at the coexistence region. For the tristable dynamics when epithelial, hybrid and mesenchymal phenotypes coexist, we found that the hybrid state is the most probable, agreeing with experiments. Our results highlight new mechanisms related to the EMT dynamics in response to TGF- $\beta$  stimulus in epithelial breast cells and might help the design of therapeutic strategies for breast cancer.

# 1. INTRODUCTION

The epithelial-mesenchymal transition (EMT) is a biological process where epithelial cancer cells are converted into mesenchymal ones during tumoral progression. Epithelial cells undergo this mechanism losing their adhesion properties to acquire mesenchymal features such as the secretion of proteases and migratory abilities [1]. During the process, cells can become temporarily stable in three different phenotypes: epithelial (E), mesenchymal (M) or in a hybrid one with both E and M features (H) [2]. In this way, EMT has been reported as a three-state transition process. Due to the stabilization in the H state, cells present both adhesive and migratory properties and can migrate collectively as a cluster presenting up to 50 times higher metastatic potential than individual migratory cells in the M state [3]. Thus, studying the complex mechanism of EMT is fundamental for prevention and treatment of cancer metastasis.

Epithelial cells adhere through the cadherin 1 (CDH1) protein [4]. This protein has been reported to be directly repressed by the EMT-promoting proteins: the snail family transcriptional repressor 1 (SNAIL1) and 2 (SNAIL2) and the zinc finger E-box binding homeobox 1 (ZEB1) and 2 (ZEB2), while inducing the mesenchymal marker Vimentin (VIM) [5]. This mechanism constitutes the EMT regulatory core that is induced by the transforming growth factor beta (TGF- $\beta$ ) [6]. The downstream control in this process is performed by proteins and microRNAs that regulate SNAIL1/2 and ZEB1/2 expressions. MicroRNAs or miRNAs are small, non-coding RNAs that regulate the degradation of messenger RNAs (mRNAs). The molecular circuitry of EMT has different feedback loops that control its dynamics. The most studied double-negative feedback loops in the EMT process are composed by SNAIL1/2 and microRNA-34a/b/c (collectively referred here as miR-34) [7], and ZEB1/2 and microRNA-200a/b/c (collectively referred here as miR-200) [8].

Recently, new molecules were reported to interact with ZEB1 during TGF- $\beta$ -responsive EMT in breast cancer. In this context, two new positive feedback loops were found involving new autocrine mechanisms related to ZEB1: the first one involving hyaluronan synthase 2 (HAS2) and the HAS2-synthesized hyaluronic acid (HA) and the second one related to the epithelial splicing regulatory protein 1 (ESRP1) and the CD44 standard isoform (CD44s) [9, 10]. Both feedback loops were identified as mechanisms in which ZEB1 accelerates EMT. Another molecule that has been identified to interact with ZEB1 is the grainyhead like transcription factor 2 (GRHL2). This protein is linked to ZEB1 through a double negative feedback loop, which was described to control EMT-driven tumor progression. GRHL2 also prevents the expression patterns of the M state in breast cancer cells associated to high levels of CD44s concentration and low levels of the CD24 molecule (CD24) [11, 12]. Furthermore, new miRNAs were also described to interact with ZEB1, such as miR-1199-5p, miR-340 and miR-190. The first and the second participate in a double-negative



feedback loop with ZEB1 that controls EMT-mediated breast cancer progression [13, 14]. The third one was identified to suppress TGF- $\beta$  signaling during EMT [15]. In addition to this finding, miR-190 was also described to be transcriptionally suppressed by ZEB1 [15]. Thus, these new findings reveal a more complex regulatory scenario for TGF- $\beta$ -responsive EMT and demand updated modeling.

Models are constantly being produced to elucidate the dynamics of complex biological systems [16–20]. In the EMT context, Tian et al. proposed a mathematical model consisting of "Cascading Bistable Switches" (CBS) [18]. The CBS model contemplated the role of SNAIL1/miR-34 and ZEB1/miR-200 double-negative feedback loops in the EMT process describing bistable switches that control the transition between the EMT states. Zhang et al. found experimental support to the bistable dynamics in the CBS model [21]. In addition, they expanded the interactions considered in the model demonstrating a more complex dynamic for EMT.

Here we propose a new model of EMT based on recent identified data on new molecules and feedback loops. Our model agrees with theoretical and experimental data and predicts new roles for miR-340 and miR-1199 related to the stabilization of the H state in EMT.

## 2. RESULTS

### 2.1 Graph of regulatory mechanisms for EMT

**Fig. 1** shows the graph of regulatory interactions reported in the literature used by our logical model (see section Methods). It is well established that exogenous or endogenous TGF- $\beta$  is a major inducer of EMT [6] through activation of SNAIL1 in epithelial cells [22]. MiR-34 controls SNAIL1 expression through inhibition of its translation, whereas the SNAIL1 protein represses miR-34 transcription forming a positive circuit (also known as double-negative feedback) [7]. SNAIL1 can also self-regulate through a negative feedback loop [23]. SNAIL1 activation stimulates expression of VIM and inhibition of CDH1 [21]. SNAIL1 also promotes ZEB1 expression [24], while its transcription is inhibited by multiple other molecules in this process such as miR-200, miR-340, miR-1199 and GRHL2. ZEB1 can also inhibit the expression of these molecules as it forms double-negative circuits with each one of them [10, 12–14]. ZEB1 and SNAIL1 also repress the expression of miR-34 and miR-200, respectively [21]. In addition, ZEB1 is observed to be inhibited by CD24 [12], which is induced by GRHL2 [12]. ZEB1 can also repress the TGF- $\beta$  inhibitor miR-190 [15]. Furthermore, ZEB1 regulation involves two positive circuits: one involving ESRP1 and CD44s [9] and another involving HAS2 which promotes HA production [10]. Finally, both SNAIL1 and ZEB1 control the activation of VIM and CDH1 which regulate the EMT phenotype [21]. Each node and directed arc of the network in **Fig. 1** represent a model component and an (activatory or inhibitory) interaction between two components, respectively. The model has a single input and output that represent exogenous TGF- $\beta$  (Ex\_TGFB) and EMT

activation, respectively. Blue and yellow nodes represent epithelial and mesenchymal markers, respectively. Rectangular nodes represent boolean components while elliptic nodes represent non-boolean components taking the values 0, 1 or 2 in order to represent in more detail the dynamics of activation of EMT. Logical rules controlling the evolution of each component and the corresponding bibliographical references of each molecular interaction required for rule proposal are presented in Table 1.

## 2.2 Dynamics of the wild-type case

We found stable and multistable dynamics for the wild-type case of our model. For  $Ex\_TGFB=OFF$  there is a single stable state where EMT is OFF. However, for  $Ex\_TGFB ON$  there is a tristable state, i.e. the E, H and M states or phenotypes coexist (as shown in Fig. 2), being each phenotype defined in agreement with experimental markers. The activation of ESRP1, miR-190, CD24, GRHL2, miR-200, miR-1199-5p (referred to as miR-1199 thereafter), miR-340, miR-34 and CDH1 characterize the epithelial (E) phenotype [7, 9, 12–14, 21, 25]. The M phenotype corresponds to the activation of VIM, SNAIL1, ZEB1, HAS2, CD44s and HA [7, 9, 10, 21]. Both, SNAIL1 and ZEB1 activations are required for inhibition of CDH1 and activation of VIM (respectively) to stabilize the M state [21]. The stabilization of the hybrid phenotype (H) with both E and M features is characterized by the induction of SNAIL1, that changes to the intermediary level 1 the expression of the epithelial marker CDH1, and of the mesenchymal marker VIM. The H phenotype also requires the inhibition of ZEB1 which blocks the inhibition of ESRP1, GRHL2 and the miRNAs: 190, 200, 1199, 340. Zhang and colleagues [21] verified this stabilization mechanism of the H state for the human breast epithelial cell line MCF10A.

The tristable dynamics is controlled by the functional circuits presented in Table 2 which are discussed in the next section.

## 2.3 Perturbation analysis and model validation

In order to investigate the role of these circuits in the model dynamics, we performed circuit component perturbations to determine which circuits are required to stabilize a given phenotype (see section Methods). From the large number of possible perturbations, we selected only a few capable of reproducing the phenotypes observed in the wild-type state (Fig. 3 (A)). These results show that SNAIL1/miR-34 presents an important role to stabilize the E and H states, whereas the other circuits are related to the transition from the H to the M state, the same behavior is observed for these circuits under perturbations. The negative circuit SNAIL1 was not considered in this analysis since it is involved in triggering oscillatory behavior not observed to play a role

in EMT. Moreover, the circuits involving ZEB1, miR-200, miR-1199, miR-340, and GRHL2 were perturbed altogether since their individual perturbations yielded the same results.

As seen in Fig. 3(A), ZEB1-mediated circuits presented an important role in the stabilization of the M state. However, all these circuits except ZEB1/HAS2/HA presented activated components in the E and the H states. For this reason, we knocked out its components to investigate their specific role (Fig. 3B). We found that miR-200, GRHL2, miR-340, and miR-1199 are required to stabilize the hybrid state since their knockout preserve only the E and the M states. In the literature, proteins with these features are known as phenotypic stability factors (PSFs) and the first two are reported in the literature as having this exact function [12, 21, 26]. The knockout of ESRP1 has little influence on the phenotypes suggesting that it is not essential for stabilizing the H state. We concluded that only the two first circuits in Fig. 3(A) are required to generate the tristable dynamics. A diagram showing how the input states of TGF- $\beta$  affect the functionality of the circuits controlling the EMT states is presented in Fig. 3 (C).

To clarify the functional role of each PSF on the H state, we performed the knockout of each 3 of the 4 PSFs in order to show that each one alone can preserve the stability of the H state (Fig. 4). In Table 3 we present the two PSFs predicted by the model for the hybrid state, the other two PSFs were already determined experimentally.

To expand the perturbation analysis, we performed a systematic assessment of the effects of all single and double perturbations of the model components (see Fig. 5). In this way, the knockout (KO) or ectopic activation (E) of its corresponding levels were performed to compare the effects of different perturbations, facilitating the detection of dominant or synergistic effects on the EMT states. The results were represented by colors corresponding to each observed phenotype under the respective perturbation, i.e. red, yellow, and green represent E, H, and M states, respectively. This analysis complements the results in Fig. 3 in the sense that mutations including the elements of the circuits in Fig. 3 present higher impact on phenotypes. For example, single mutations of ESRP1 do not present a relevant effect on phenotypes, while most of the mutations of miR-340 and miR-1199 tend to destabilize the M state. Fig. 5 also provides sequential stepwise information of the EMT dynamics. For instance, the activity of SNAIL1 is clearly the first step of EMT signaling as its inactivation blocks the input signal preserving only the E state, whereas its active state allows the system to transit to the H and M states. The second step is focused on ZEB1 activity, whose inactivation does not allow the transition to the M state, while its active state induces predominantly this phenotype. These findings are also consistent with experimental data from Zhang and colleagues [21].

Finally, we determined phenotype probabilities for the wild-type case when the model input is ON and the system has a tristable dynamics. We estimated the probability of each attractor by an efficient Monte Carlo sampling algorithm (Avatar) of the state space over 10,000 runs, which is much longer than the convergence transient observed [27]. The results are presented in Table 4 and were compared to the corresponding fraction of cells found in

cultures treated with 0.5 ng/ml concentration of human recombinant transforming growth factor beta-1 (hTGF- $\beta$ 1) for 7 days [21] (see Fig. 6). The experiments and the model seem to agree to find a population of H type cells slightly higher than the other two.

### 3. DISCUSSION

The EMT process is reported as a mechanism required for cancer metastasis [28]. The operating principles of EMT are based on epithelial cancer cells undergoing a phenotypic alteration mediated by an external stimulus. This mechanism converts cancer epithelial cells into motile mesenchymal ones allowing migration and invasion of other sites. Importantly, the TGF- $\beta$ -induced EMT is reported to produce 3 stable and distinct phenotypes: E, H and M. The E and the M states are characterized by their respective phenotypic markers, with the H state having both E and M features. In this context, different models were proposed to investigate the dynamics of the TGF- $\beta$ -induced EMT. However, they ignore the recent new molecules and circuits reported in the literature. Then we decided to propose a model contemplating this recent information to explore their role on the tristable dynamics [21]. The model was proposed based on the established EMT regulatory core with inclusion of the new experimental evidence. Indeed, according to Zhang and colleagues [21] the analysis of an expanded network is an important approach for a better understanding of EMT.

The tristable dynamics of TGF- $\beta$ -induced EMT is observed experimentally in epithelial cells (MCF10A) treated with different concentrations of exogenous TGF- $\beta$  [21]. Our model predicts similar behavior for Ex\_TGFB ON presenting better quantitative agreement with 0.5 ng/ml concentration results. Both, experiments and the model find a population of the H type cells slightly higher than the other two. In addition, it also agrees with the experimental observation that cells remain in the E state for inactive Ex\_TGFB [21].

The circuit perturbation analysis predicted a switch behavior in all functional circuits which seems to be responsible for the transitions between the phenotypes. We observed that the SNAIL1/miR-34 circuit presented a switch behavior controlling the E to H transition. This process was observed in EMT of MCF10A cell line [21]. The other ZEB1-mediated circuits were observed to control the H to the M transition. In particular, the circuit ZEB1/miR-200 was experimentally observed to affect the transition [21]. Another recently published circuit, ZEB1/HAS2/HA, that was implemented in our model, was observed to stabilize the M state [10]. Then, we investigated the knockdown of these components to determine their phenotypic role. We found that ZEB1/miR-200/miR-340/miR-1199/GRHL2 circuits have a functional role in the stabilization of the H state. The circuits involving ZEB1, miR-200 and GRHL2 were already reported to have this feature [12, 21, 26].

The stability of the tristable dynamics of our model occurs via the functionality of the positive circuits. Firstly, the input signaling induces the bistable switch behavior on SNAIL1/miR-34 circuit. The switch presenting active SNAIL1 is able to induce its downstream target ZEB1 affecting the

functionality of ZEB1/miR-1199/miR-340/GRHL2/miR-200 circuits. Each mutual inhibition in these circuits controls the stability of the H and M states. While miR-340, miR-1199, miR-200, and GRHL2 block the SNAIL1 signaling through the inhibition of the activity of ZEB1, the system remains in the H state, otherwise it transits to the M state. Then, the present study extends the knowledge in terms of the transitory mechanism controlled by the circuits related to EMT, since Zhang and colleagues [21] contemplated only two of the circuits in their work.

The conclusion that miR-340 and miR-1199 are PSFs originates mainly from the result that their silencing destabilizes the H state. The H state can be stabilized by only one of the circuits involving miR-340 and miR-1199 (while the other 3 are turned off, see Fig. 4). The activity of these miRNAs is thus an important feature to be explored on the tristable dynamics triggered by TGF- $\beta$  during EMT. The identification of PSFs has been an important strategy to understand the ability of cells to restrict phenotypic changes in a biological process [26, 29]. Moreover, less toxic treatments [29] can be implemented through the combination of low drug doses and the suppression of these factors in order to eradicate cells in the H state.

MiR-340 and miR-1199 were reported as controllers of ZEB1 expression during TGF- $\beta$  signaling in MCF10A cells. Moreover, both microRNAs were described as inhibitors of migration and invasion by inactivation of ZEB1. In our model, these features are present including new ones related to the H state stabilization, state which is reported to present 50 times higher metastatic potential than cells in the M state [3]. According to Zhang et al. [21], breast cancer epithelial cells (MCF10A) treated with 0.5 ng/ml hTGF- $\beta$ 1 undergo a high phenotypic alteration toward the H phenotype. Hou and colleagues demonstrated that this same concentration of TGF- $\beta$ 1 does not significantly reduce the expression of miR-340 in MCF10A cells. In addition, the study suggests that the ZEB1/miR-340 negative circuit becomes functional due to TGF- $\beta$ 1 treatment in a dose and time-dependent manner. These results strongly indicate that miR-340 is required for stabilization of the H phenotype. Regarding miR-1199, Diepenbruck et al. [13] showed that in the presence of TGF $\beta$  this molecule presents similar functions to the miR-200 family members miR-200b-3p and miR-429-3, whose overexpression destabilizes the transition to the mesenchymal state in MCF10A cells. In spite of the role played in EMT of normal breast cells (MCF10A), miR-1199 and miR-340 also present an important role in breast cancer cell lines: e.g. miR-340 overexpression inhibits breast cancer progression in the MDA-MB-231 breast cancer cell line [14], while miR-1199 was verified to inhibit ZEB1 in the Py2T murine breast cancer cell line [13]. Moreover, in the presence of TGF- $\beta$ , miR-1199 delays EMT in Py2T cells [13]. Thus, our predictions about miR-340 and miR-1199 are consistent with experimental data.

Summarizing, the model presents good quantitative agreement with experimental and other theoretical works associated with EMT in breast cancer cell lines. Worth of note is the prediction of a new role for miR-340 and miR-1199 in stabilizing the H state during EMT. Therefore, our approach may contribute to suggest alternative therapeutic strategies for cancer through drugs

targeting these microRNAs that might significantly decrease the metastatic potential of malignant cells.

## 4. METHODS

We used the logical formalism [30] explained briefly below and the simulations were conducted using the software GINsim 2.9.5 available for download from <http://compbio.igc.gulbenkian.pt/nmd/node/82> [31]. The GINsim version of the model can be found in File S1.

The logical formalism is based on the components of the regulatory graph  $G$  defined as  $G = \{g_1, g_2, \dots, g_n\}$  and the transition function  $K$ , which control the state of  $G$  associating a logical function  $K(g_i)$  to each  $g_i$ . These logical functions define the activity or concentration level of each  $g_i$  represented by the values in  $\{0, \dots, \max_i\}$ , where  $\max_i$  is the maximum value defined for the model component  $g_i$ . These functions are based on the influence of the edges that connect the  $g_i$ , which are built by the logical operators AND, OR, and NOT to represent the biochemical interactions found in the literature. A state of the model is defined by a vector  $g = [g_1, g_2, \dots, g_n]$  in the state space  $S$ .

To simulate the model dynamics two approaches can be performed: synchronous or asynchronous updating schemes. In the first scheme all nodes evolve at the same time and the process is completely deterministic, whereas in the second one, each component is updated in a random order and the process may not be deterministic. In this work, we used only the asynchronous scheme since non-deterministic behavior is commonly observed in molecular networks. Negative circuits have an odd number of inhibitions and are required for oscillatory dynamics, whereas positive circuits have an even number of inhibitions and are required for multi-stability [32]. Stable states of the network correspond to attractors that have a single state and are associated with phenotype determination, while oscillatory dynamics have a cyclic attractor and are associated with transient phenotypes [33]. In the case of stochastic dynamics, we estimate the probability of the system being captured by a given attractor of the dynamics. This is performed by the Avatar algorithm provided by GINsim. Observe that for small enough models is possible to identify the attractors, however it is not trivial to evaluate their reachability probabilities [27].

Each model component value is represented by colors, in this way boolean components with 0 and 1 values are represented by red and green colors, respectively. In the case of multi-value components with three states, the intermediary state has yellow color.

In order to analyze the role of the functional circuits of the model, we performed perturbations in the circuit components to investigate their influence on the phenotypes of the wild-type case. These perturbations constrain the

level of the corresponding model component to remain fixed, e.g. a component 'knockdown' forces its value to remain zero (OFF), whereas an 'overexpression' forces it to remain in a value greater than zero.

## REFERENCES

1. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *The Journal of clinical investigation*. 2009;119:1420–1428.
2. Grosse-Wilde A, d'Hérouël AF, McIntosh E, Ertaylan G, Skupin A, Kuestner RE, et al. Stemness of the hybrid epithelial/mesenchymal state in breast cancer and its association with poor survival. *PLoS one*. 2015;10:e0126522.
3. Aceto N, Bardia A, Miyamoto DT, Donaldson MC, Wittner BS, Spencer JA, et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell*. 2014;158:1110–1122.
4. Conacci-Sorrell M, Zhurinsky J, Ben-Ze'ev A. The cadherin-catenin adhesion system in signaling and cancer. *The Journal of clinical investigation*. 2002;109:987–991.
5. Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. *cell*. 2009;139:871–890.
6. Xu J, Lamouille S, Derynck R. TGF- $\beta$ -induced epithelial to mesenchymal transition. *Cell research*. 2009;19:156.
7. Siemens H, Jackstadt R, Hünten S, Kaller M, Menssen A, Götz U, et al. miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. *Cell cycle*. 2011;10:4256–4271.
8. Brabletz S, Brabletz T. The ZEB/miR-200 feedback loop—a motor of cellular plasticity in development and cancer? *EMBO reports*. 2010;11:670–677.
9. Preca B-T, Bajdak K, Mock K, Sundararajan V, Pfannstiel J, Maurer J, et al. A self-enforcing CD44s/ZEB1 feedback loop maintains EMT and stemness properties in cancer cells. *International journal of cancer*. 2015;137:2566–2577.
10. Preca B-T, Bajdak K, Mock K, Lehmann W, Sundararajan V, Bronsert P, et al. A novel ZEB1/HAS2 positive feedback loop promotes EMT in breast cancer. *Oncotarget*. 2017;8:11530–11543.
11. Cieply B, Farris J, Denvir J, Ford HL, Frisch SM. Epithelial–mesenchymal transition and tumor suppression are controlled by a reciprocal feedback loop between ZEB1 and Grainyhead-like-2. *Cancer research*. 2013.

12. Mooney SM, Talebian V, Jolly MK, Jia D, Gromala M, Levine H, et al. The GRHL2/ZEB feedback loop—a key axis in the regulation of EMT in breast cancer. *Journal of cellular biochemistry*. 2017;118:2559–2570.
13. Diepenbruck M, Tiede S, Saxena M, Ivanek R, Kalathur RKR, Lüönd F, et al. miR-1199-5p and Zeb1 function in a double-negative feedback loop potentially coordinating EMT and tumour metastasis. *Nature communications*. 2017;8:1168.
14. Hou L-K, Yu Y, Xie Y-G, Wang J, Mao J-F, Zhang B, et al. miR-340 and ZEB1 negative feedback loop regulates TGF- $\beta$ -mediated breast cancer progression. *Oncotarget*. 2016;7:26016.
15. Yu Y, Luo W, Yang Z-J, Chi J-R, Li Y-R, Ding Y, et al. miR-190 suppresses breast cancer metastasis by regulation of TGF- $\beta$ -induced epithelial–mesenchymal transition. *Molecular cancer*. 2018;17:70.
16. Gupta S, Silveira DA, Mombach JCM. Modeling the role of microRNA-449a in the regulation of the G2/M cell cycle checkpoint in prostate LNCaP cells under ionizing radiation. *PloS one*. 2018;13:e0200768.
17. Issler MVC, Mombach JCM. MicroRNA-16 feedback loop with p53 and Wip1 can regulate cell fate determination between apoptosis and senescence in DNA damage response. *PloS one*. 2017;12:e0185794.
18. Tian X-J, Zhang H, Xing J. Coupled reversible and irreversible bistable switches underlying TGF $\beta$ -induced epithelial to mesenchymal transition. *Biophysical journal*. 2013;105:1079–1089.
19. Mombach JC, Vendrusculo B, Bugs CA. A model for p38MAPK-induced astrocyte senescence. *PloS one*. 2015;10:e0125217.
20. Steinway SN, Zañudo JG, Ding W, Rountree CB, Feith DJ, Loughran TP, et al. Network modeling of TGF $\beta$  signaling in hepatocellular carcinoma epithelial-to-mesenchymal transition reveals joint sonic hedgehog and Wnt pathway activation. *Cancer research*. 2014;74:5963–5977.
21. Zhang J, Tian X-J, Zhang H, Teng Y, Li R, Bai F, et al. TGF- $\beta$ -induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci Signal*. 2014;7:ra91–ra91.
22. Peinado H, Quintanilla M, Cano A. Transforming growth factor  $\beta$ -1 induces Snail transcription factor in epithelial cell lines mechanisms for epithelial mesenchymal transitions. *Journal of Biological Chemistry*. 2003;278:21113–21123.
23. Peiro S, Escriva M, Puig I, Barbera MJ, Dave N, Herranz N, et al. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic acids research*. 2006;34:2077–2084.
24. Dave N, Guaita-Esteruelas S, Gutarra S, Frias À, Beltran M, Peiro S, et al. Functional cooperation between Snail1 and Twist in the regulation of Zeb1 expression



during epithelial-to-mesenchymal transition. *Journal of Biological Chemistry*. 2011;;jbc–M110.

25. Burk U, Schubert J, Wellner U, Schmalhofer O, Vincan E, Spaderna S, et al. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO reports*. 2008;9:582–589.

26. Jolly MK, Tripathi SC, Jia D, Mooney SM, Celiktas M, Hanash SM, et al. Stability of the hybrid epithelial/mesenchymal phenotype. *Oncotarget*. 2016;7:27067.

27. Mendes ND, Henriques R, Remy E, Carneiro J, Monteiro PT, Chaouiya C. Estimating attractor reachability in asynchronous logical models. *Frontiers in physiology*. 2018;9.

28. Heerboth S, Housman G, Leary M, Longacre M, Byler S, Lapinska K, et al. EMT and tumor metastasis. *Clinical and translational medicine*. 2015;4:6.

29. Yaswen P. Reinforcing targeted therapeutics with phenotypic stability factors. *Cell Cycle*. 2014;13:3818–3822.

30. Abou-Jaoudé W, Traynard P, Monteiro PT, Saez-Rodriguez J, Helikar T, Thieffry D, et al. Logical modeling and dynamical analysis of cellular networks. *Frontiers in genetics*. 2016;7:94.

31. Chaouiya C, Naldi A, Thieffry D. Logical modelling of gene regulatory networks with GINsim. In: *Bacterial Molecular Networks*. Springer; 2012. p. 463–479.

32. Kaufman M, Soule C, Thomas R. A new necessary condition on interaction graphs for multistationarity. *Journal of Theoretical Biology*. 2007;248:675–685.

33. Thieffry D. Dynamical roles of biological regulatory circuits. *Briefings in bioinformatics*. 2007;8:220–225.

## TABLES

**TABLE 1** - Logical rules controlling the states of the nodes in Fig. 1. The logical operators AND, OR and NOT are used to define the rules for each node according to the states of its regulators. The model contains a single boolean input, Ex\_TGFB, and a single output, EMT. The corresponding bibliographical references of each molecular interaction required for rule proposal are listed.

Node	Highest level	Rules	Biological interpretation	References
Ex_TGFB	1		Exogenous TGFB present	[6]
TGFB	1	<i>NOT</i> (miR-200 <i>OR</i> miR-190)		[15]
ESRP1	1	<i>NOT</i> ZEB1		[9]
miR-190	1	<i>NOT</i> ZEB1		[15]
CD24	1	GRHL2		[12]
GRHL2	1	<i>NOT</i> ZEB1		[12]
miR-200	1	<i>NOT</i> (ZEB1 <i>OR</i> SNAIL1)		[21, 25]
miR-1199	1	<i>NOT</i> ZEB1		[13]
miR-340	1	<i>NOT</i> ZEB1		[14]
miR-34	1	<i>NOT</i> (ZEB1 <i>OR</i> SNAIL1)		[7, 21]
CDH1	2	<i>NOT</i> ZEB1 <i>AND</i> <i>NOT</i> SNAIL1	Activated CDH1	[21]
	1	<i>NOT</i> (ZEB1 <i>AND</i> SNAIL1) <i>AND</i> (ZEB1 <i>OR</i> SNAIL1)	Intermediate concentration of CDH1	[21]
VIM	2	SNAIL1 <i>AND</i> ZEB1	Activated VIM	[21]
	1	(SNAIL1 <i>OR</i> ZEB1) <i>AND</i> <i>NOT</i> (ZEB1 <i>AND</i> SNAIL1)	Intermediate concentration of VIM	[21]
SNAIL1	1	((TGFB <i>AND</i> <i>NOT</i> SNAIL1) <i>OR</i> Ex_TGFB) <i>AND</i> <i>NOT</i> miR-34		[7, 22, 23]
ZEB1	1	<i>NOT</i> CD24 <i>AND</i> <i>NOT</i> GRHL2 <i>AND</i> <i>NOT</i> miR-340 <i>AND</i> <i>NOT</i> miR-200 <i>AND</i> <i>NOT</i> miR-1199 <i>AND</i> ((CD44s <i>AND</i> HA) <i>OR</i> SNAIL1)		[9, 10, 12–14, 24]
HAS2	1	ZEB1		[10]
CD44s	1	<i>NOT</i> ESRP1		[9]
HA	1	HAS2		[10]
EMT	2	VIM <i>AND</i> <i>NOT</i> CDH1	Mesenchymal phenotype	[21]
	1	(VIM <i>OR</i> <i>NOT</i> CDH1) <i>AND</i> <i>NOT</i> (VIM <i>AND</i> <i>NOT</i> CDH1)	Hybrid phenotype	[21]

**TABLE 2** - Functional circuits in the model and corresponding experimental validation.

<b>Circuits</b>	<b>References</b>
<b>Positive</b>	
ESRP1/CD44s/ZEB1	[9]
ZEB1/HAS2/HA	[10]
ZEB1/miR-200	[25]
ZEB1/miR-1199	[13]
ZEB1/GRHL2	[12]
ZEB1/miR-340	[14]
SNAIL1/miR-34	[7]
<b>Negative</b>	
SNAIL1	[23]

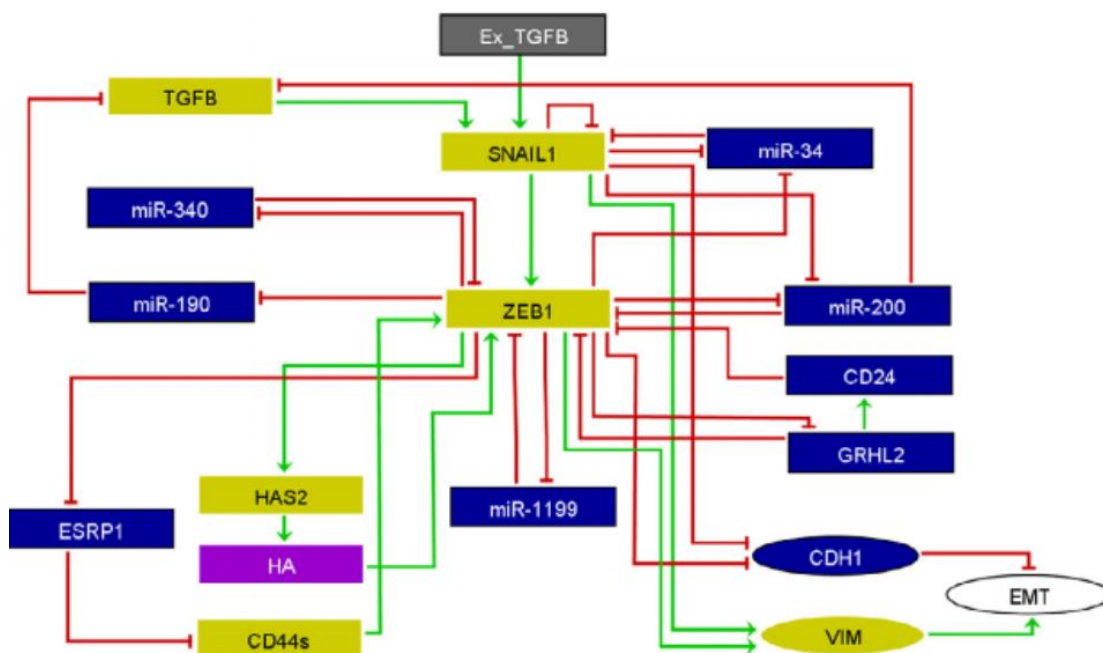
**TABLE 3** - Phenotypic stability factors found in the model and experimental observations. Cases for which no direct evidence were found are considered predictions of the model.

<b>Phenotypic stability factor</b>	<b>Status</b>	<b>References</b>
miR-200	known	[21]
GRHL2	known	[12, 26]
miR-340	not confirmed	-
miR-1199	not confirmed	-

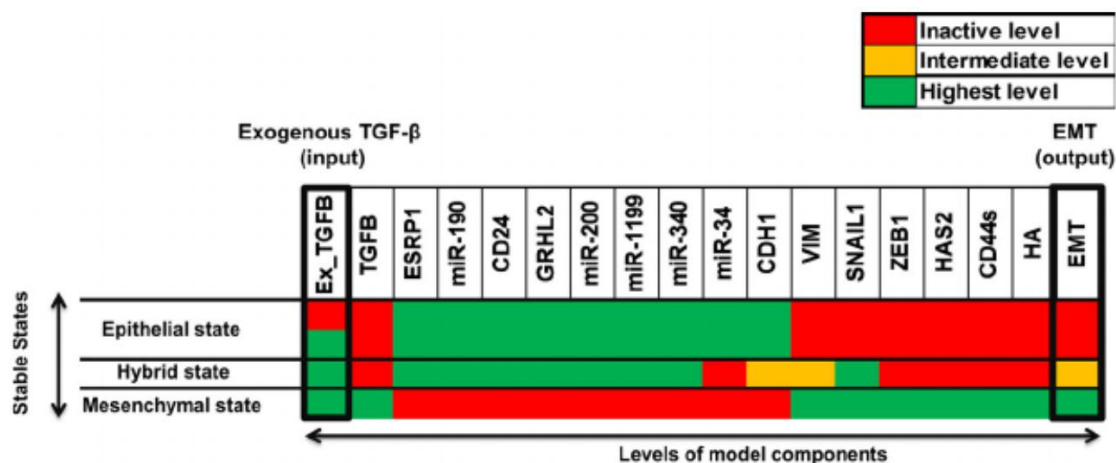
**TABLE 4** - Wild-type state probabilities for the tristable dynamics.

<b>Phenotypes</b>	<b>Probability</b>
E state	40.61%
H state	49.16%
M state	10.23%

## FIGURES

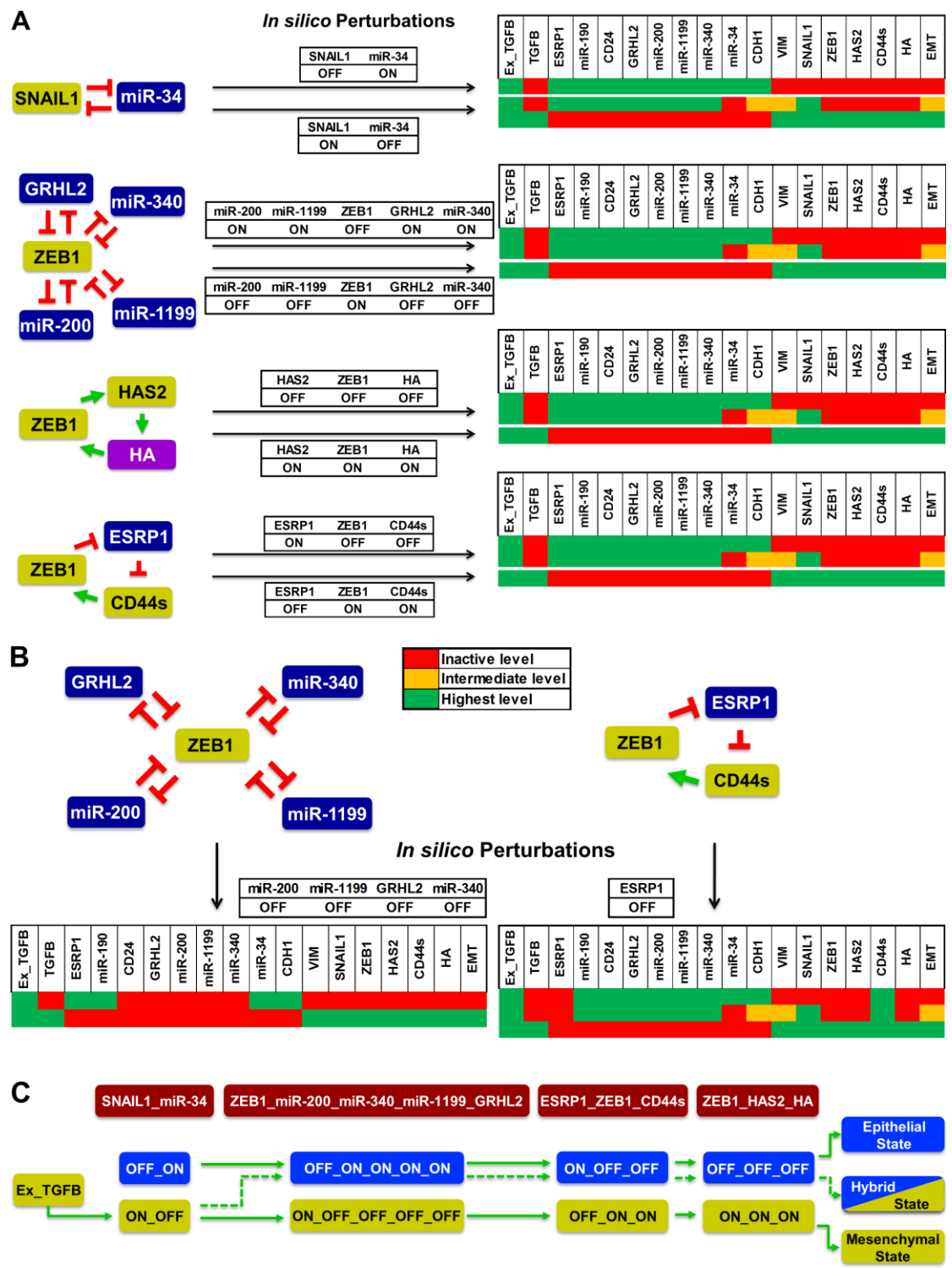


**FIGURE 1** - Regulatory network for EMT in response to exogenous TGF- $\beta$ . Green and red arcs denote activation and inhibition interactions, respectively. Elliptical nodes represent multi-valued components, while rectangular nodes are boolean ones. Nodes in blue and yellow colors represent epithelial and mesenchymal markers, respectively. The input node in gray color denotes Ex\_TGFB (exogenous TGF- $\beta$ ), whereas the elliptical white node represents the model output (EMT). The rectangular node in purple color denotes hyaluronic acid (HA). This figure was generated using the software GINsim 2.9.5 [31].



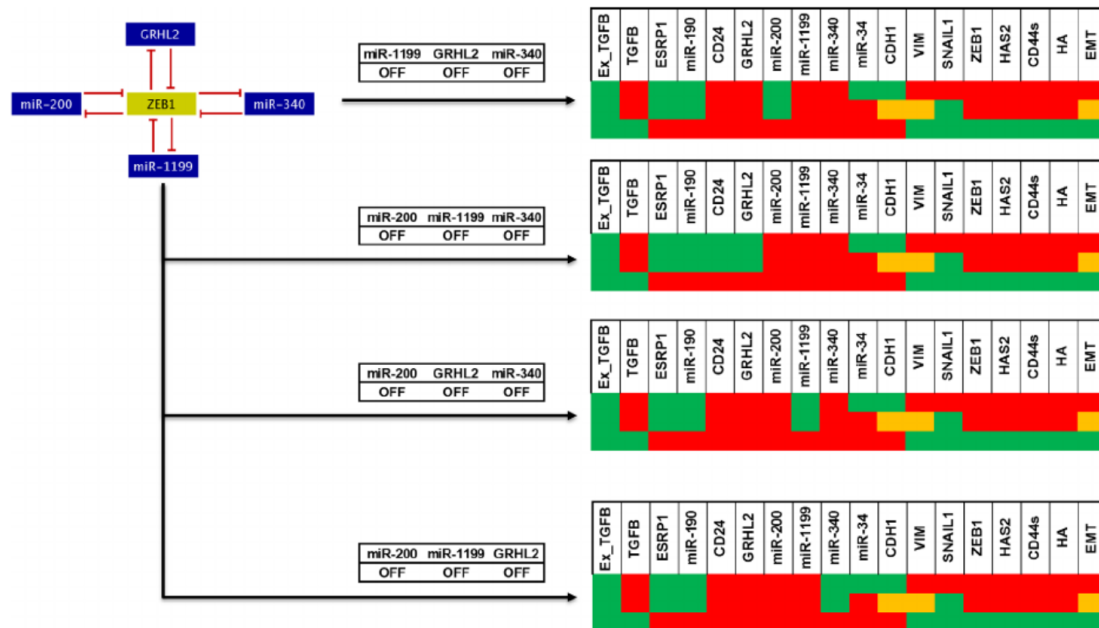
**FIGURE 2** – Stable states of the wild-type case. Each line represents a single stable state of the model. Left-most column presents the levels of the input, Ex\_TGFB. Right-most column presents the model output, EMT. Red, yellow, and green colors correspond to active, intermediate and inactive node states,

respectively. Observe that for Ex\_TGFB ON the dynamics is tristable, i. e. the E, H and M phenotypes coexist.

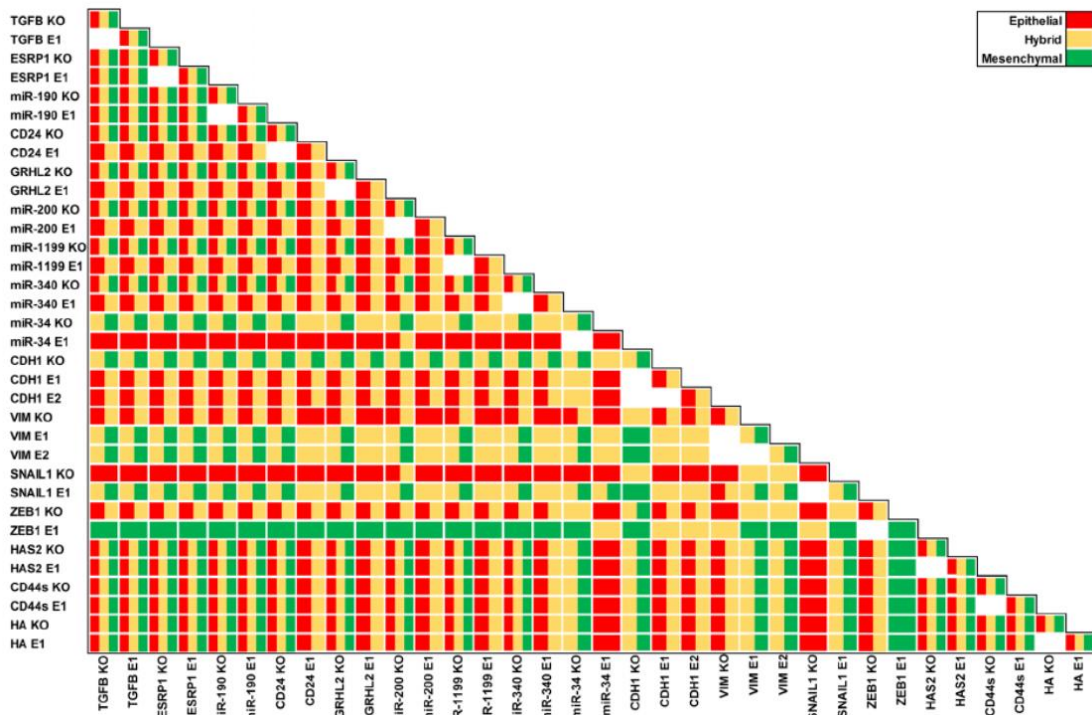


**FIGURE 3** – Phenotype alterations according to perturbations of functional circuits. Red, yellow, and green colors represent the model activation levels corresponding to high, intermediary and inactivation, respectively. (A) Perturbations for Ex\_TGFB ON. The positive circuit SNAIL1/miR-34 switches

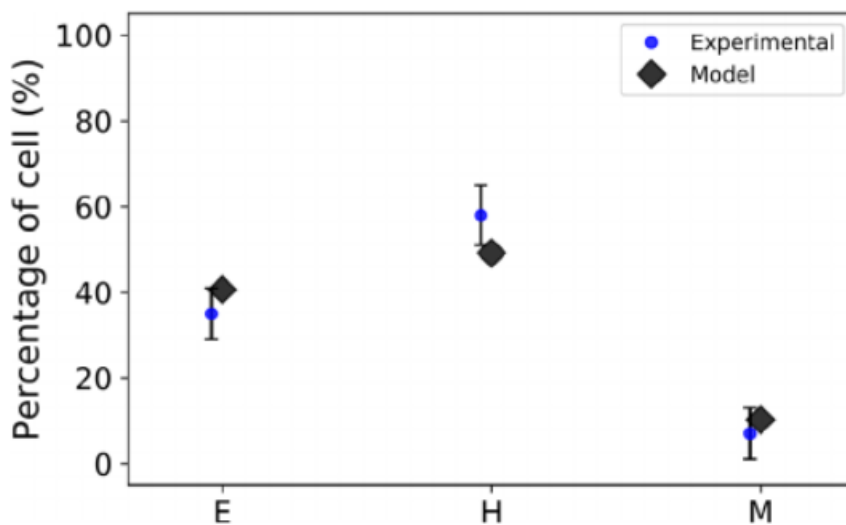
from inactive SNAIL1 and active miR-34 to the inverse state in the E and the H phenotypes, respectively. The other circuits present similar behavior but are related to the stabilization of the H and M state. (B) Knockdown of active components in the E and H states related to ZEB1-mediated circuits according to perturbations presented in (A). We found that miR-200, GHRL2, miR-340, and miR-1199 are required to stabilize the hybrid state since their knockout preserve only the E and M states. (C) Schematics of the network dynamics showing the downstream signaling effect of the input Ex\_TGFB. Top: model elements. Below: states. Dash and solid arrows represent the induction of the H and M states, respectively. For inactive Ex\_TGFB, the solid arrows show the induction of the E state.



**FIGURE 4** – Functional role of each circuit involving ZEB1, miR-1199, miR-340, and GRHL2 in the stabilization of H state. Each circuit alone can preserve the stability of the H state. Red, yellow, and green colors denote the activation levels high, intermediary and inactivation, respectively.



**FIGURE 5** – Phenotypes found for single and double perturbations of the model components. The hypotenuse of the triangle shows the phenotypic results of single perturbations of the model components on the left. The other elements of the triangle correspond to multiple perturbations combining left component perturbations with those at the bottom of the figure. KO denotes that the corresponding component was knocked out, whereas E represents overexpression related to its respective activity level (1 or 2). The resulting phenotypes for each perturbation are identified by colors as in Fig. 2. Red, yellow, and green colors represent epithelial, hybrid, and mesenchymal states, respectively.



**FIGURE 6** – Comparison of the model predicted probabilities (when Ex\_TGFB is ON) for E, H, and M states with experimental observations [21]. Blue dots

represent experimental data and black diamonds denote the model predictions. Experimental results corresponding to cells treated with concentrations of 0.5 ng/ml of hTGF- $\beta$ 1 for 7 days. The experimental data is presented as means plus one standard deviation related to three biological replicates.

## **ACKNOWLEDGEMENTS**

D.A.S. acknowledges financial support from CNPq (145708/2017-9).

## **AUTHORS' CONTRIBUTIONS**

D.A.S. and J.C.M.M. conceived the experiment(s), D.A.S. conducted the experiment(s), D.A.S. and J.C.M.M. analyzed the results. All authors reviewed the manuscript.

## **CONFLICT OF INTEREST**

There are no conflicts to declare.



### 3 ARTIGO:SYSTEMS BIOLOGY APPROACH SUGGESTS NEW MIRNAS AS PHENOTYPIC STABILITY FACTORS IN THE EPITHELIAL-MESENCHYMAL TRANSITION

Artigo publicado no periódico científico "*Journal of The Royal Society Interface*"<sup>1</sup> e disponível, na sua versão final, em: <https://doi.org/10.1098/rsif.2020.0693>. Tal publicação objetiva explorar a dinâmica da transição epitélio-mesenquimal (TEM) para diferentes níveis de atividade de *input* (TGF- $\beta$ ) e analisar a funcionalidade dos circuitos necessários para a estabilização fenotípica em cada um dos estados estáveis relacionados aos diferentes níveis do *input*. Os seguintes achados foram observados nesse estudo:

- I. A transição ocorre a partir do estado E, evoluindo para os estados H e finalizando no estado M;
- II. TGF- $\beta$  pode iniciar a TEM via a ativação de SNAIL1 induzindo os estados H, enquanto que a subsequente indução de ZEB1 mantém a estabilidade do estado M;
- III. O circuito negativo associado à SNAIL1 pode preservar o estado E;
- IV. TWIST1/miR-129 está envolvido na indução de um estado H através de mecanismo de irreversibilidade;
- V. O microRNA (miR) 129 pode ser um fator de estabilidade fenotípica (FEF) do estado H;
- VI. A estabilidade TWIST1 ON e miR-129 OFF do circuito correspondente resulta da funcionalidade dos circuitos mediados por ZEB1 com GRHL2 e os miRNAs 200, 340 e 1199;
- VII. O circuito TGF $\beta$ /SNAIL1/miR-200 pode estar envolvido na manutenção do estado M.

---

<sup>1</sup>Fator de impacto (2020): 4.118



Article submitted to Journal of the  
Royal Society Interface

**Subject Areas:**

Computational biology,  
Bioinformatics, Systems biology

**Keywords:**

modeling, epithelial-mesenchymal  
transition, miRNAs

**Author for correspondence:**

José Carlos Merino Mombach  
e-mail: [jcmombach@ufsm.br](mailto:jcmombach@ufsm.br)

## Systems biology approach suggest new miRNAs as phenotypic stability factors in the epithelial-mesenchymal transition

Daner A. Silveira<sup>1,\*</sup>, Shantanu Gupta<sup>1,\*</sup>  
and José Carlos M. Mombach<sup>1,\*</sup>

<sup>1</sup>Departamento de Física, Universidade Federal de  
Santa Maria, RS, Brazil

\* Equally contributing authors

The epithelial-mesenchymal transition (EMT) is a cellular program on which epithelial cells undergo a phenotypic transition to mesenchymal ones acquiring metastatic properties such as mobility and invasion. TGF- $\beta$  signaling can promote the EMT process. However, the dynamics of the concentration-response of TGF- $\beta$ -induced EMT is not well explained. In this work we propose a logical model of TGF- $\beta$  dose dependence of EMT in MCF10A breast cells. The model outcomes agree with experimentally observed phenotypes for the wild-type and perturbed/mutated cases. As important findings of the model, it predicts the coexistence of more than one hybrid states and that the circuit between TWIST1 and miR-129 is involved in their stabilization. Thus, miR-129 should be considered as a phenotypic stability factor. The circuit TWIST1/miR-129 associates with ZEB1-mediated circuits involving miRNAs 200, 1199, 340, and the protein GRHL2 to stabilize the hybrid state. Additionally, we found a possible new autocrine mechanism composed by the circuit involving TGF- $\beta$ , miR-200 and SNAIL1 that contributes to the stabilization of the mesenchymal state. Therefore, our work can extend our comprehension of TGF- $\beta$ -induced EMT in MCF10A cells to potentially improve the strategies for breast cancer treatment.

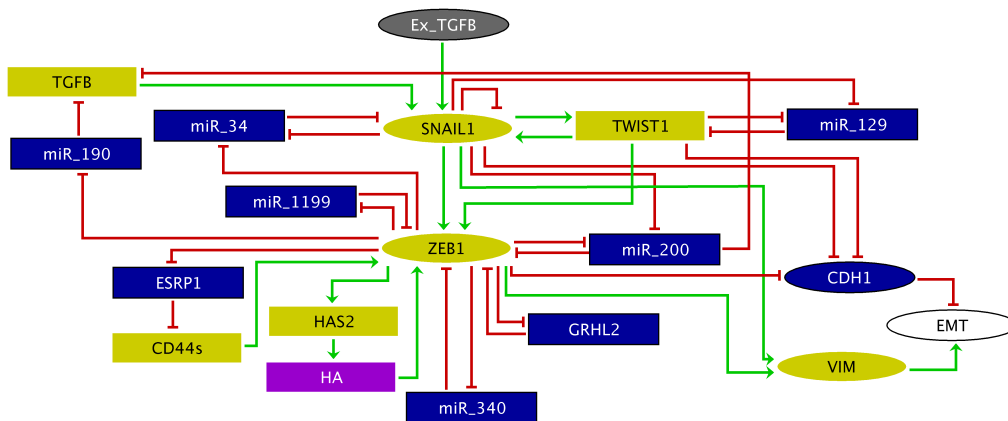
## 1. Introduction

The epithelial-mesenchymal transition (EMT) is a cellular program in which epithelial cells transit to a mesenchymal phenotype occurring in various pathological processes, mainly cancer progression [1]. The loss of adhesive properties and the gain of migratory abilities are the main EMT features characterizing phenotypic changes from epithelial (E) to mesenchymal (M). Additionally, some studies demonstrated the stabilization of a hybrid (H) state having both E and M features [2, 3, 4]. Cells in this state present adhesive and invasive abilities, which confers high-grade malignancy on individual carcinoma cells [4]. This phenomenon directly depends on the signals received by a cell in the tissue environment affecting the intracellular gene circuitry dynamics.

The malignant progression is dependent on the EMT activation and the development of metastasis [5]. This mechanism is a multistep process in which EMT is important to the migration of cancer cells until the metastatic colonization in distant organs [6]. The initiation of EMT is triggered by the tumour microenvironment. The secretion of various cytokines and chemokines by stromal cells can act as initiation factors inducing EMT in epithelial cancer cells [7]. The Transforming growth factor- $\beta$  (TGF- $\beta$ ) is one common secreted molecule in the tumour microenvironment and its corresponding pathway extensively studied in the EMT context. By activation of the corresponding surface receptors of the cell membrane, TGF- $\beta$  triggers the induction of the transcription factors (TFs), the Snail family transcriptional repressor 1 (SNAIL1) and the Zinc finger E-box binding homeobox 1 (ZEB1) which, in turn, repress an important feature of the epithelial phenotype, the calcium-dependent cell-cell adhesion glycoprotein Cadherin-1 encoded by the CDH1 gene [8]. Thus, activation of SNAIL1 and ZEB1 is fundamental for TGF- $\beta$ -induced EMT.

The well-known controllers of SNAIL1 and ZEB1 expression are the family of microRNAs 34 (miR-34) and 200 (miR-200), respectively. Their interactive correlation occurs via reciprocal inhibitions forming double-negative feedback loops (or circuits) between SNAIL1 and miR-34 [9] and ZEB1 and miR-200 [10]. The activation of these circuits is reported as essential for EMT [8]. With the advances in EMT research, other molecules were found to regulate ZEB1 expression affecting the EMT program. For instance, two new microRNAs, miR-340 and miR-1199, were shown to form double negative circuits with ZEB1 [11, 12]. Regarding SNAIL1 regulators, the Twist family bHLH transcription factor 1 (TWIST1), another potent EMT inducer, was reported to form a positive feedback circuit with SNAIL1 during TGFB signaling [13]. The study showed that this positive circuit can down-regulate the miR-129-5p forming a double-negative feedback circuit with TWIST1. In addition, TWIST1 is also a negative regulator of CDH1, favoring the EMT process. Understanding the functional role of the negative regulators of SNAIL1 and ZEB1 during EMT can help the development of new strategies to target these regulators and to control EMT.

Recently, we published a model about the dynamics of the regulatory cores of SNAIL1 and ZEB1 during TGF- $\beta$ -induced EMT using a logical computational approach [14] based on the experimental work by Zhang *et al.* [8]. The logical method is recognized as a valuable tool to study biological regulatory processes [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25]. Our analysis indicated testable predictions that could help improve strategies for breast cancer treatment. In our approach we focused only on the coexistence region of the E, H and M states which corresponds to a specific dose of TGF- $\beta$  used in the work by Zhang *et al.* [8]. However, different concentrations of TGF- $\beta$  can lead to distinct dynamics in the EMT process [8]. So, in order to study a more realistic dose-dependent dynamics of EMT, here we present a more general model (Fig. 1) contemplating different levels of activity of TGF- $\beta$ -induced EMT and also the additional molecules, TWIST1 and miR-129, that play important roles on the regulatory core of EMT.



**Figure 1.** Proposed regulatory network of TGF- $\beta$ -driven EMT. Elliptic and rectangular nodes represent multi-valued and Boolean components, respectively. Nodes in gray and white denote the input (Ex\_TGFB) and the output (EMT) of the network, respectively. Epithelial markers in the network are represented by blue color, while mesenchymal markers are in yellow color. Green arrows represent activatory interactions and hammer-head arcs denote inhibitory ones. The rectangular node in purple color represents hyaluronic acid (HA). This figure was generated using GINsim 3.0.0b.

## 2. Methods

### The logical formalism and GINsim

The logical formalism aims to explain qualitatively the dynamical systems [15, 26]. For molecular systems, the molecules in a regulatory graph are represented as nodes and their interactions as directed edges (representing activatory or inhibitory interactions). In a logical model nodes can have two states (active or inactive, also known as Boolean variables) or more than two (also known as multi-valued variables) which considers intermediate levels between active and inactive states. This definition depends on the activity level of the molecule in the biological process. In other words, if the node states depends on concentration levels that are well characterized by only active or inactive states, then the modeling via Boolean variables is considered, otherwise multi-valued nodes are used. The regulatory interactions reported in the experimental literature are translated into logical functions that control the state of each node in terms of its regulators through combinations of the logical operators AND, OR and NOT. The operator OR represents an individual effect of a regulator on a specific node independent of other regulators, whereas the operator AND denotes a joint requirement to perform an effect. The operator NOT represents inhibitory actions.

The update of the model components can be performed by asynchronous or synchronous methods. In the first one, the components are updated randomly, whereas in the second one all nodes are updated at the same time and the evolution is completely deterministic. In the present study, we performed the simulations using only the asynchronous updating scheme due to its potential to present stochastic behavior [26].

The network evolution in its state space is determined by its attractors and initial states. Each state in this space is represented by a vector encompassing the current levels of activity of all network nodes. The transitions among the network states is determined by the logical functions controlling each network node. In the asynchronous updating scheme, nodes are selected randomly for updating. When the update scheme reaches an unchanging state, we call it a stable attractor (also known as stable state or steady state). Otherwise, it can be trapped in a subset of network states characterizing a cyclic attractor (also known as complex attractor). The

network evolution can be represented in the form of a State Transition Graph (STG), whose nodes are network states and whose edges represent the transitions between these states.

The presence of functional circuits (or feedback loops) characterize in more detail the origin of attractors [27]. The existence of a negative circuit can trigger sustained oscillations, while a positive one is a necessary condition for multistationarity. A closed circuit in the network with an odd/even number of inhibitions is defined as negative/positive.

The logical approach allows *in silico* perturbations of nodes to analyze the effect of such circuits on the dynamics of the network [28]. Logic-based perturbations represent modifications in the dynamics of the model that account for a perturbation affecting the states of the nodes. Knock-out of a gene encoding a specific component is represented by a loss of function (LoF) of the corresponding node by fixing its value at 0. On the other hand, overexpression or ectopic expression of the gene is translated into the model as its gain of function (GoF) by forcing it to remain in a value greater than zero. To identify important circuits we performed node perturbations to evaluate their role on the model dynamics.

In the present study, we used the GINsim 3.0.0b tool for the simulation and logical analysis of the model. This tool is a Java software suite, freely available for download from the GINsim website (<http://ginsim.org/downloads>) [29]. This software can compute the dynamical behavior of the model for any initial state as well as the functional circuits of the network. Moreover, *in silico* perturbations can be tested with GINsim.

It is possible to compute the dynamical behavior of the model for any initial state. The state of each model component is iteratively updated according to the logical formulae. The resulting dynamics is represented in terms of state transition graphs (STG). The number of states in STG, i.e., the graph connecting all network states visited in the evolution, increases exponentially with the number of network nodes. Then, to deal with these complexity, we used the Hierarchical Transition Graph (HTG) algorithm in GINsim. A node in the HTG represents a cluster of states that share the same set of successor states. The clustered nodes represent irreversible sequences of states or cycles. A transient cyclic component denotes a set of states composed by cyclic trajectories that present outgoing transitions to other cluster in the HTG. Whereas the irreversible ones containing outgoing transitions are not composed by cyclic trajectories. The attractors (cyclic or stable) have no outgoing transitions to other components in the HTG. For instance, the presence of cycles in the dynamics elucidate the role of negative circuits in the system. For a deeper description of HTGs see [30]. Cytospace 3.6.1 was used to generate the HTGs in the present study [31].

### 3. Results

#### Proposed signaling network driving the EMT

The proposed signaling network is based on our previously published model for EMT [14], which contemplated the recent biochemical data on the EMT regulatory core composed by the positive circuits SNAIL1/miR-34 and ZEB1/miR-200 and endogenous TGF- $\beta$  [8]. New experimentally validated circuits were considered in the proposed network such as ZEB1-mediated positive circuits with ESRP1 and CD44s [32], miR-1199 [12], miR-340 [11], and GRHL2 [21]. HA and HAS2 were also included in the network as performed by Jolly *et al.* [33]. MiR-190 and CD24 were also considered due to their influence on TGF- $\beta$  signaling during EMT [21, 34]. CDH1 and VIM are considered important epithelial and mesenchymal markers, respectively, [35] then they were also included. However, in the present study, we propose a much more general model considering five different levels for the input representing the concentration of extracellular TGF- $\beta$ : absent (0), low (1), intermediate (2), high (3), and very high (4). This allows us to model the influence on the EMT dynamics of extracellular TGF- $\beta$  concentration, as performed in the detailed experiment by Zhang *et al.* [8]. The experiment analyzed epithelial breast cells (MCF10A) exposed to different concentrations of human recombinant TGF- $\beta$ 1 (hTGF- $\beta$ 1). Cells without treatment with hTGF- $\beta$ 1 remained in the E state, whereas with 0.5 ng/ml of hTGF- $\beta$ 1 close fractions of E and H cells

were produced. For a concentration of 1.0 ng/ml, E, H and M states were populated, but a smaller fraction of cells remained in the E state. The high 2.0 ng/ml produced only H and M cells, whereas the highest one (4.0 ng/ml) induced mostly the M state. In addition, we also contemplate in the model the observation by Yu and colleagues that TWIST1 can cooperate with SNAIL1 to induce EMT in MCF10A cells [13]. Such cooperation occurs via a double-positive circuit between these molecules controlled by miR-129.

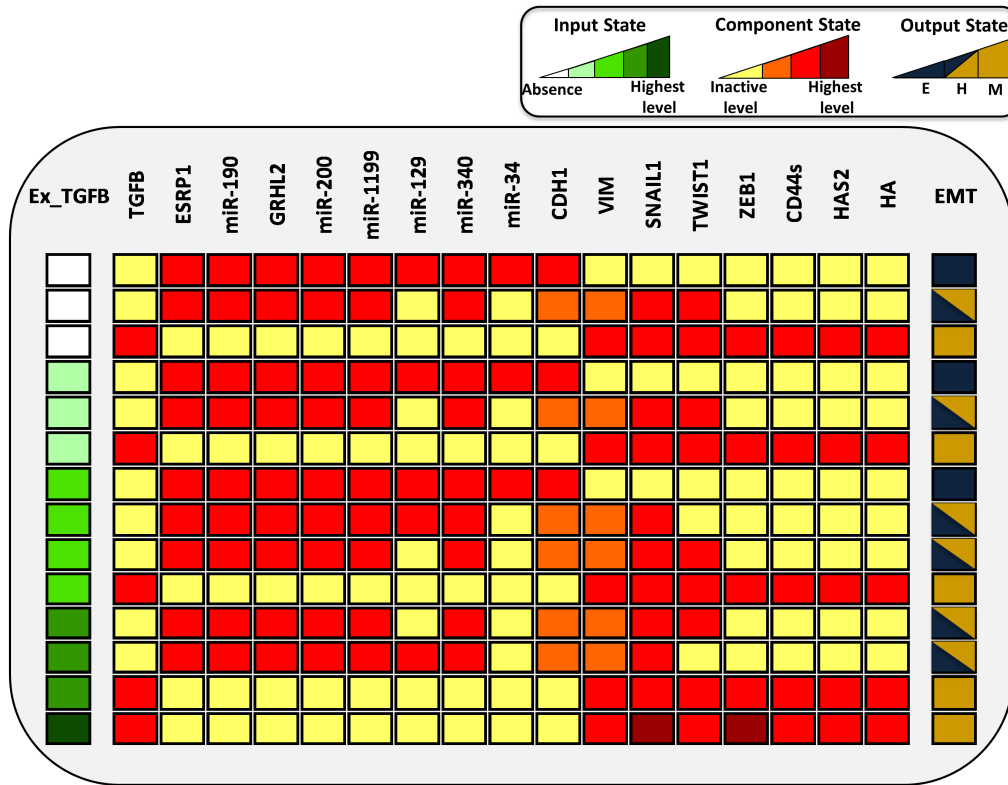
Then, these experimental observations motivated us to propose this new model for the EMT dynamics.

## Translating the proposed network into a logical model

In order to analyze the dynamics of the proposed regulatory network, we translated it into a computational logical model. Each node representing a molecule of the network is defined as Boolean or multi-valued (see section Methods). Based on the work of Zhang *et al.* [8], CDH1 and VIM require three levels of activity to characterize the EMT states [8]. Exposure of cells to different concentrations of hTGF- $\beta$ 1 leads to three expression patterns of CDH1 (E-cadherin) and VIM [8]: high CDH1 and low VIM; intermediate CDH1 and intermediate VIM; low CDH1 and high VIM. This suggested us to consider 3 levels of activity for both molecules in the model. High and low levels of CDH1 and VIM, respectively, correspond to the E state, whereas the inverse pattern is associated with the M state. When both present intermediate concentrations the H state is featured. The same approach was applied to the input, exogenous TGF- $\beta$ , in order to model the different concentrations used in the study by Zhang and colleagues [8]. Concentrations of 0 (absence), 0.5 (low), 1 (intermediate), 2 (high) and 4 (very high) ng/ml of TGF- $\beta$  were assigned to levels 0, 1, 2, 3 and 4, respectively (see previous section). Similarly, the network output (EMT) was defined with three levels: 0, 1 and 2, corresponding to E, H, and M states, respectively. For SNAIL1 and ZEB1 were also assigned 3 levels denoting low, intermediate and high levels. According to Zhang and colleagues [8], SNAIL1 and ZEB1 present mainly two expression patterns: low/high SNAIL1 and low ZEB1 (indicating the E or H states), and high SNAIL1 and high ZEB1 (indicating the M state). In addition, we included another level characterizing very high concentrations of SNAIL1 and ZEB1. This was proposed because the relative protein abundances of SNAIL1 and ZEB1 in response to 4 ng/ml of TGF- $\beta$  present higher levels than 2 ng/ml. To the regulators of SNAIL1 and ZEB1, MiRNAs 34 and 200, respectively, were assigned 2 levels. ZEB1/miR-200 and SNAIL1/miR-34 circuits were observed to present 2 distinct stable states [8]: high SNAIL1 and low miR-34, and low SNAIL1 and high miR-34; high ZEB1 and low miR-200, and low ZEB1 and high miR-200. This suggested to consider 2 states for both miRNAs. The remaining components of the proposed network were defined as Boolean variables due to the lack of quantitative information. Specification of the logical rule controlling each component and the corresponding published experimental validation are provided in Table S1 (PubMed links included). A summary of the modeling assumptions are presented in the supporting information. The model file is provided in supplementary file S1.

## Wild-type case of the model

To analyze the dynamics of the different levels of the input exogenous TGF- $\beta$  (referred as Ex\_TGFB in the model), we determined the wild-type case (WT) of the model by considering all possible initial conditions in the simulation (Fig. 2). The phenotypic characterization of each state was performed according to phenotype markers [8, 9, 10, 11, 12, 13, 21, 32, 36]: activation of miR-200, 34, 1199, 340, CDH1, GRHL2, CD24, and ESRP1 and inactivation of the remaining components was assigned to the epithelial phenotype. Activation of SNAIL1, TGFB, ZEB1, VIM, CD44s, HAS2 and inactivation of the remaining components was assigned to the mesenchymal phenotype. The hybrid state was assigned to the case when VIM and CDH1 remain in their intermediate states [8]. The WT shows strong multistable (stochastic) behavior, i.e., the



**Figure 2.** Wild-type results of the model. Each line represents a steady state of the model. Yellow, orange, and red cells represent inactive, intermediate, and active states of the components, respectively. White, light green, green, dark green, and very dark green denote levels of the input Ex\_TGFβ corresponding to absence, low, intermediate, high, and very high concentration, respectively. The model phenotypes are interpreted according to the output state (EMT): inactive, intermediate, and high values of the EMT node represent Epithelial (dark blue), Hybrid (half colored cells) and Mesenchymal (dark yellow) phenotypes, respectively.

coexistence of different phenotypes (stable states or attractors) for a given input value. The detailed information of the states of the model can be found in Table 1.

For absence and low Ex\_TGFβ levels, three phenotypes (attractors) were observed: the E, H, and M states. Regarding the intermediate level, Ex\_TGFβ induced 4 attractors (a quadri-stability) corresponding to the E, H and M states. Its high level destabilized the E state, maintaining the stability of H and M states. The highest level of Ex\_TGFβ triggers a deterministic behavior, inducing only the M phenotype.

In the wild-type (WT) model, our approach identified 10 functional circuits, i.e., circuits that have influence on the dynamics (Table 2). Nine of which have already been validated by experimental studies, which supports the model, and a predicted one. We further analyzed the functionality context of circuits according to input levels. In this way, we observed that most circuits are functional for all activity levels (0, 1, 2, 3, and 4), except the circuits TGFβ/SNAIL1/miR-200, SNAIL1/miR-34 and SNAIL1 self-inhibition (Table 2). The functional role of these circuits depends on specific activity levels of the input. The sorting of circuit functionality according to the input level is important to analyze the wild-type dynamics, results are shown in (Table 2). Once the functionality of a circuit is characterized for one input level, its role for other input levels is identical.

In a previous work limited to describe only the coexistence of the phenotypes in EMT, we studied the functional role of eight of these circuits [14]. In summary, the SNAIL1/miR-34

circuit is responsible for the E to H transition, whereas the ZEB1-mediated circuits involving miR-340, miR-1199, miR-200, and GRHL2 are associated to the H to M transition. The circuits ZEB1/HA/HAS2 and ZEB1/ESRP1/CD44s contribute to stabilize the M state via sustained activation of ZEB1, the main regulator of the M state.

## Model dynamics

To characterize the system dynamics, we investigated the qualitative behavior of the trajectories in the state space. However, the number of states in the state transition graph (i.e., the graph

**Table 1.** Definition of the model attractors and related phenotypes. References supporting the definitions are listed.

Model state	considered phenotype	References
<b>Active:</b> miR-190, miR-34, miR-340, miR-1199, miR-129, miR-200, CDH1 (highest level), GRHL2, ESRP1 <b>Inactive:</b> TGFB, SNAIL1, TWIST1, ZEB1, CD44s, HAS2, HA, VIM	Epithelial	[8, 9, 10, 11, 12, 13, 21, 32, 36]
<b>Active:</b> miR-190, miR-340, miR-1199, miR-129, miR-200, CDH1 (intermediate level), GRHL2, ESRP1, VIM (intermediate level), SNAIL1 (high level) <b>Inactive:</b> miR-34, TGFB, TWIST1, ZEB1, CD44s, HAS2, HA	Hybrid	[8, 9, 10, 11, 12, 13, 21, 32, 36]
<b>Active:</b> miR-190, miR-340, miR-1199, TWIST1, miR-200, CDH1 (intermediate level), GRHL2, ESRP1, VIM (intermediate level), SNAIL1 (high level) <b>Inactive:</b> miR-34, TGFB, miR-129, ZEB1, CD44s, HAS2, HA	Hybrid	[8, 9, 10, 11, 12, 13, 21, 32, 36]
<b>Active:</b> TGFB, ZEB1 (high/highest levels), CD44s, HAS2, HA, TWIST1, VIM (highest level), SNAIL1 (high/highest levels) <b>Inactive:</b> miR-34, miR-190, miR-340, miR-1199, miR-200, CDH1, GRHL2, ESRP1, miR-129	Mesenchymal	[8, 9, 10, 11, 12, 13, 21, 32, 36]

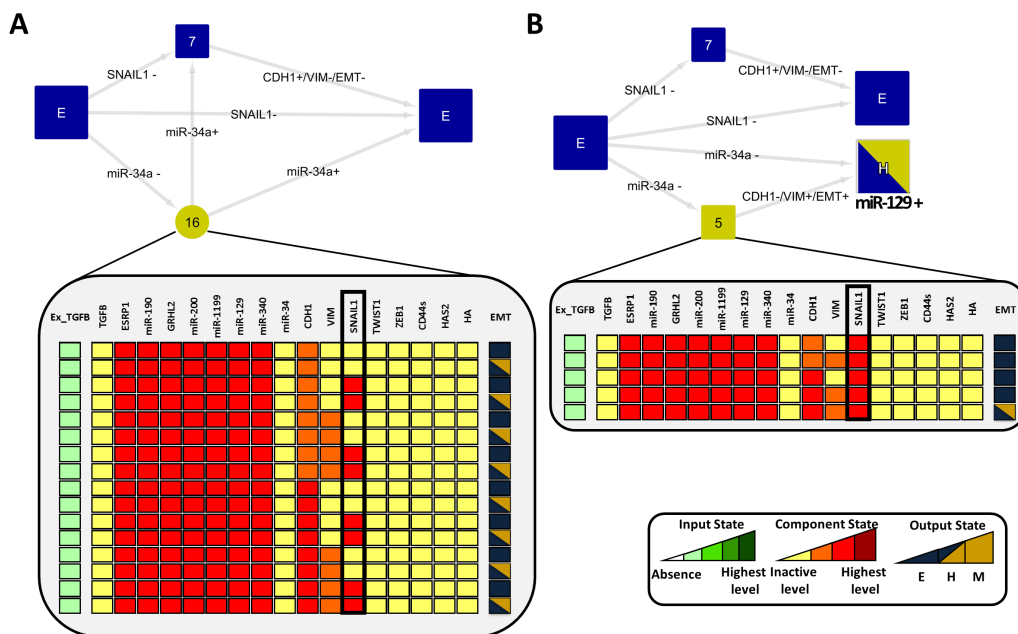
**Table 2.** Functional circuits of the wild-type dynamics of the model and the corresponding input levels required for their functionalities. Experimental reference of each circuit is presented.

Positive Circuits	Input level context	References
GRHL2 / ZEB1	Any	[21]
ZEB1 / miR_340	Any	[11]
ZEB1 / miR_1199	Any	[12]
TWIST1 / miR-129	Any	[13]
TGFB / SNAIL1 / miR_200	0	Prediction
SNAIL1 / miR_34	Range(0-1), 2	[9]
ZEB1 / HAS2 / HA	Any	[36]
ZEB1 / miR_200	Any	[10]
ZEB1 / ESRP1 / CD44s	Any	[32]
Negative Circuit		
SNAIL1	1	[37]



connecting all network states visited in the evolution) increases exponentially with the number of network nodes. Then, to deal with this complexity we used the Hierarchical Transition Graph (HTG) algorithm to visualize the transitions. The method produces a more compact representation of the trajectories. For more details, see the last paragraph of section Methods and references therein.

We found that SNAIL1 self-inhibition is involved in the stability of the E phenotype for the low level of the input  $Ex\_TGFB=1$ . To do that, we started from a state of the network that was neighbor to state E, a neighbor state is defined as an initial network state where all nodes are in the E state but one which is set to the M state. Then, we let the system evolve examining which circuits were actively involved and the trajectory performed in the space state (Fig. 3). That was repeated for each node. We found that unperturbed SNAIL1 circuit triggered trajectories that ended only at the E state (Fig. 3A). The removal of the self-inhibition of SNAIL1 creates a bistable state between E and H (Fig. 3B), implying that this negative circuit is responsible for the stability of the E state for low dose. Upon perturbations of the circuit miR-34/SNAIL1, the bistability of E and H is removed suggesting that this circuit is involved in the stability of state H (see Supplementary file S2).



**Figure 3.** Analysis of lower input active level. In HTG, most-left states represent initial states of the simulation, whereas most-right ones denote final states. Arrows representing transitions between the states are labeled with the updated components. Plus and minus signs indicate that the component is increasing or decreasing its activity level, respectively. Blue and yellow transient states denote that incoming transitions are related to E and M phenotypes, respectively. The number of states that compose each transient is shown. (A) We observed that the unperturbed SNAIL1 circuit generated trajectories where the final state was only the E state. The HTG shows that SNAIL1 transient oscillations can also contribute to stabilize the E state. The depicted 16 states shows that system can oscillate between E and H states until it finally remains at the E state. (B) LoF of SNAIL1 self-inhibition triggers the transition to the H state destabilizing the E state.

For the intermediate level ( $Ex\_TGFB=2$ ), the SNAIL1/miR-34 circuit contributes to stabilize the E state as shown in our previous work [14]. Interestingly, two distinct H states appear and their difference is only the activation of TWIST1. The H state with inactive TWIST1 becomes stable as an effect of the circuit TWIST1/miR-129 and the higher level of the input  $Ex\_TGFB=2$ .

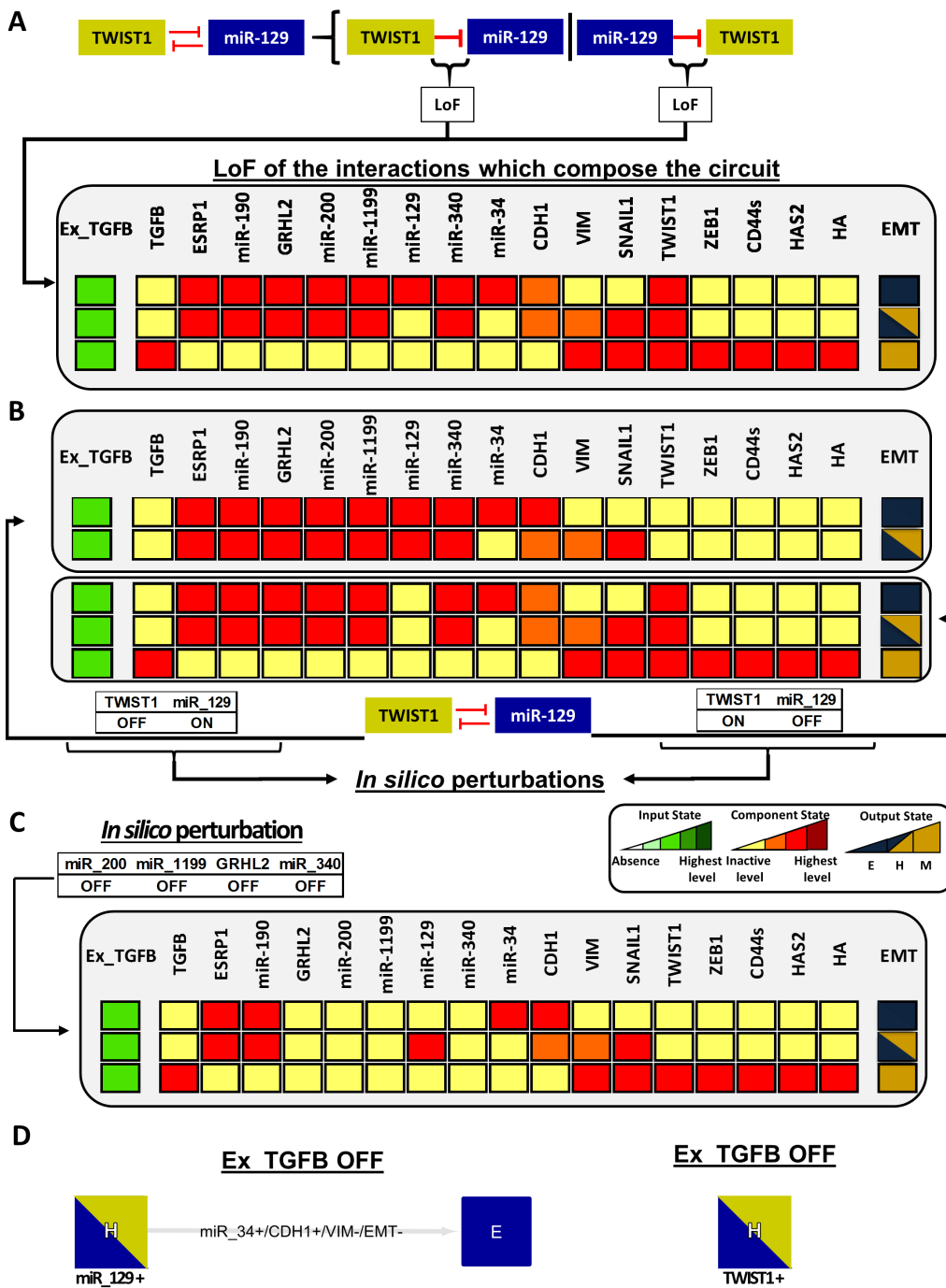
As we did for the low input, we determined that the circuit TWIST1/miR-129 contributes for the bistability of the H states because the LoF of the circuit interactions destroys this behavior (Fig. 4A). By performing individual perturbations on the circuit interactions, we observed that the inhibitory interaction of miR-129 on TWIST1 is responsible for the stabilization of the H state. In this way, the stability of this specific H state is directly related to the ability of miR-129 to inhibit TWIST1 expression during TGF- $\beta$  signaling. Hence, miR-129 should be considered as a phenotypic stability factor (PSF) of the H phenotype.

In addition to this analysis, we perturbed the circuits components. We observed that a bistable switch (TWIST1 ON/miR-129 OFF and TWIST1 OFF/miR-129 ON) controls the stability of the H states (Fig. 4B). Then, we concluded that this circuit should be added to the set of circuits involved in the stability of the H state which include ZEB1/GHRL2, ZEB1/miR-1199, ZEB1/miR-340, and ZEB1/miR-200 (see section wild-type case of the model). Perturbations of the latter circuits (Fig. 4C) abrogate the H state with activated TWIST1. GHRL2, and miRNAs 1199, 340, and 200 negatively regulate the expression of ZEB1 in the model. Their LoF favor the ZEB1 up-regulation which, in turn, destabilizes the H state with up-regulated TWIST1. The remaining H state is stable due to activation of miR-129, as previously discussed. We also evaluated whether the transition between the E and the H states is reversible or irreversible (Fig. 4D). Starting from the H state with miR-129 activated, the system transits to the E state under Ex\_TGFB OFF, whereas for the H state with TWIST1 activated the effect is not observed (Fig. 4D). These results suggest that the bistable switches associated with TWIST1/miR-129 circuit can control the reversibility of the transition.

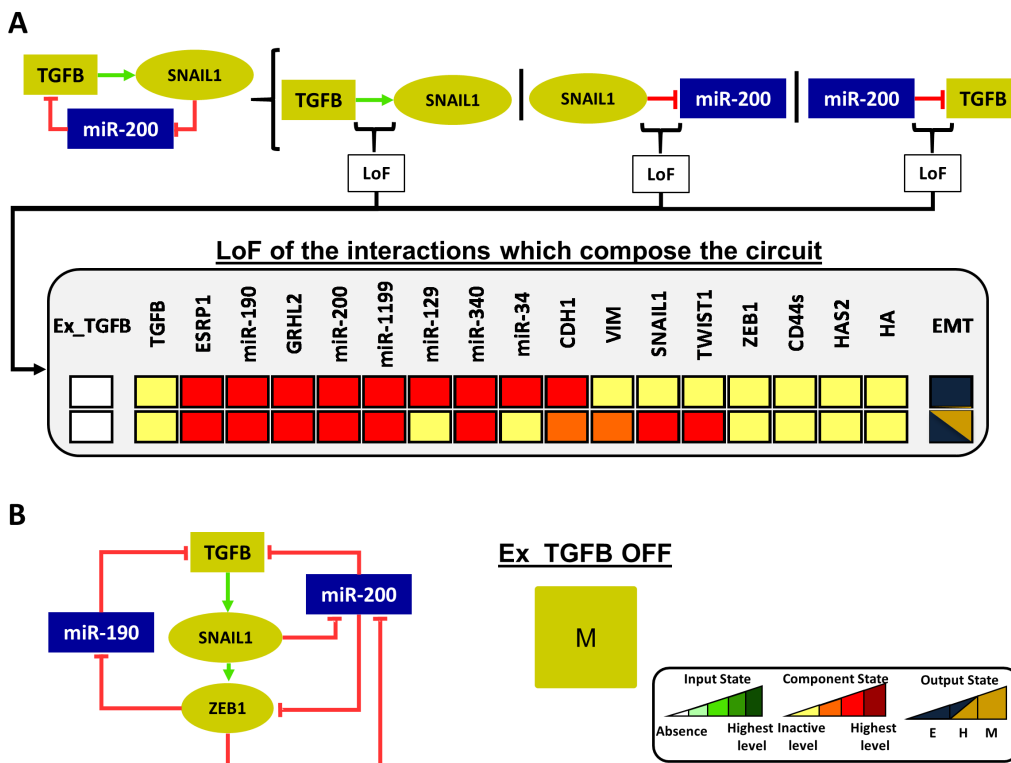
Regarding the circuit TGFB/SNAIL1/miR-200, not studied in the experimental literature for MCF10A cells, it is only functional for Ex\_TGFB=0 (Table 2). The perturbations show that this circuit is involved in the stabilization of the M state (Fig. 5A). Node perturbations of the circuit TGFB/SNAIL1/miR-200 were not performed because they affect other functional circuits that share components such as ZEB1/miR-200 and SNAIL1/miR-34 and so, we cannot evaluate its specific role in the model dynamics. Specifically, the inhibitory interaction of SNAIL1 on miR-200 is responsible for the stabilization of the M state. LoF of this interaction favors the miR-200 up-regulation which, in turn, inhibits ZEB1 expression destabilizing the M state. As performed for the TWIST1/miR-129 circuit, we evaluated the reversibility of the transition from the M state (Fig. 5B). In this way, we set the M state as an initial condition of the simulation for Ex\_TGFB OFF. Thus, we observed that the system does not transit back. These results suggest that the TGFB/SNAIL1/miR-200 circuit is also involved in the irreversibility of EMT. As ZEB1 is the main regulator of the M state in the model (see the last paragraph of section Wild-type case of the model), the functionality of the TGFB/SNAIL1/miR-200 circuit is associated with the ZEB1 regulation. Once ZEB1 is activated, it inhibits miR-200 and miR-190, allowing the activation of TGFB (Fig. 5B; most-left subnetwork). This mechanism induces again the ZEB1 activation due to TGFB-induced SNAIL1. Thus, the circuit regulates the sustained ZEB1 activation and, thus, maintaining the stability of the M state.

Both circuits, TGFB/SNAIL1/miR-200 and TWIST1/miR-129, are able to produce an irreversibility effect on the transition. The activation of TGFB and TWIST1 in the respective circuits leads to the stabilization of the M and H states, respectively, explaining the stability of these phenotypes for the input levels OFF and low (due to preservation of the E state caused by SNAIL1 self-inhibition; Fig. 3) in Fig. 2. Since all possible initial states were considered in the simulation in Fig. 2, those which presented TWIST1 and/or TGFB in active states led to stabilization of the H and/or M states, respectively.

As mentioned in section 3.3, the obtained stable states agree with known markers of the E, H, and M states. In terms of the transition processes, the simulations demonstrated a sequential stepwise dynamics of EMT (Fig. 6). First, the E state transits to the H state, which is controlled by the SNAIL1/miR-34 circuit (Fig. 6A). This result is consistent with our previous study [14]. In addition to this finding, our previous work showed that once the system reaches the H state, the transition to the M state is possible. However, the presence of the circuit TWIST1/miR-129 in this proposed network triggered the stabilization of other H state (Fig. 6B). From the activation



**Figure 4.** Perturbation analysis of TWIST1/miR-129 circuit for intermediate level with active input. Dark red, red, orange, and yellow colors denote the model component levels highest, high, intermediary and inactivation, respectively. (A) LoF of the interactions which compose the TWIST1/miR-129 circuit. The results show the loss of the H state with the absence of TWIST1. (B) The perturbations of the TWIST1/miR-129 circuit show that the bistable switches TWIST1 OFF/miR-200 ON and TWIST1 ON/miR-129 OFF control the bistability between the H states found in Fig. 2 for the intermediate level of Ex\_TGFB. (C) LoF of miR-200, miR-1199, GRHL2, and miR-340 destabilize the H state under presence of activated TWIST1. (D) Under Ex\_TGFB OFF and using H as the initial state with activated miR-129, the system transits to the E state, whereas using the H state with activated TWIST1 the system does not transit to the E state.



**Figure 5.** Perturbation analysis of TGFβ/SNAIL1/miR-200 circuit with off input. The colors yellow, orange, and red denote the model component levels: off, intermediary, and high, respectively. (A) LoF of the circuit interactions leads to the loss of stability of the M state. (B) According to the network, TGFβ can only be activated through ZEB1 activation that, in turn, inhibits miR-200 and miR-190, inducing the circuit functionality. Once TGFβ is activated, it sustains the ZEB1 activation via SNAIL1 inducing the M state even with Ex\_TGFβ OFF.

of TWIST1, the system can transit to the M state via the LoF of GRHL2, miR-200, miR-1199, and miR-340 (Fig. 6C). Collectively, these results show that the model reproduces a stepwise dynamics of EMT consistent with the literature [8]. A schematic figure showing a summary of these results is shown in Fig. 6D.

The aim of the present study was to expand the model including different levels of activity for the input component Ex\_TGFβ. This approach was motivated by the study of Zhang *et al.* [8] as mentioned in section 3.1. Our work showed that SNAIL1 self-inhibition might present an important role at low level of the input maintaining the stability of the E state (see Fig. 3). Also, we demonstrated that in the intermediate level of the input, the TWIST1/miR-129 circuit can stabilize two H states producing a quadri-stability. At its 2 highest levels, the input triggered two different behaviors. The high level is tristable with the E state absent, while the very high level induces only the M state. Indeed, by using the E state with activated SNAIL1 as the initial condition, we observed a transition to the H state with activated miR-129 (Fig. 7A), while the very high level induces a direct transition to the M state (Fig. 7B). Taken together, our results agree with the experimental observations by Zhang *et al.* [8], which show the stepwise process of EMT induced by different concentrations of TGF-β (Fig. 7C).

We further performed a perturbation analysis in order to investigate the correspondence between the resulting stable states of the model and observed experimental phenotypes. In Table 3, we present the agreement between the node perturbations and experimental observations. The

stable states resulting from the model perturbation can be found in the Supplementary file S2. In addition to the perturbation analysis, we summarize the remaining predictions and experimental support of modeling results Table 5. Collectively, these results suggest an excellent agreement between the model and experimental results.

**Table 3.** List of node perturbations and the corresponding abrogated states. The corresponding experimental validation with the MCF-10A cell line for each perturbation is listed.

Stimulus/Perturbation	Abrogated phenotypes	References
miR-34 GoF	H, M	[9]
miR-34 LoF	E	[9]
miR-1199 GoF	M	[12]
miR-200 GoF	M	[38]
miR-340 GoF	M	[11]
GRHL2 GoF	M	[39]
miR-129 GoF	H, M	[13]
TWIST1 LoF	H, M	[40]
SNAIL1 GoF (level 2)	E, H	[41]
SNAIL1 GoF (level 1)	E	[41]
SNAIL1 LoF	H, M	[41]
ZEB1 GoF (level 2)	E, H	[42]
ZEB1 GoF (level 1)	E	[42]
ZEB1 LoF	M	[42]

## Analysis of robustness of the results

In order to perform a systematic evaluation of the model robustness, we checked the network response when the model is confronted with topology failures. A failure is simulated by the removal of a single interaction from the network in Fig. 1 and quantified by the number of states and circuit functionality changes produced when compared to the WT. However, the removal of a node can be seen as a stronger interference in the network than deleting an interaction, as the former simulates the simultaneous removal of all interactions pointing at that nodes. Accordingly, deleting nodes implies some stronger variation from the original result. Table 4 shows the LoF of each interaction incoming to a given node and the changes produced in WT model states and circuit functionality. For example, the first row of the table corresponds to the LoF of each interaction incoming to the SNAIL1 node. This perturbation produces 9 changes in the initial model states of the WT. We obtained an average value of 1.55 errors for introduced failure. Table 4 shows that miR-129 has a strong influence on the network dynamics. For more details about the robustness of the model, see the S2-Table.

## Proposed experimental design based on the model

The theoretical analysis of the present study can provide guidance for experimental studies based on the model provided that the experimental data is obtained at the equilibrium concentration of the molecules, see Table 5. From this perspective, we suggest a systematic measurement of the changes in abundance of proteins, mRNAs and microRNAs of the model components under TGF- $\beta$  stimulus in MCF10A cells. Flow cytometry can be performed to test the bistable switch of the TWIST1/miR-129 circuit as performed by Zhang and colleagues [8] to analyze the SNAIL1/miR-34 and ZEB1/miR-200 circuits. The possible outcome of this experiment would be subpopulations of cells with high and low amounts of TWIST1 and miR-129. Such different amounts might suggest the functionality of TWIST1/miR-129 during the transition. Subpopulations with high

**Table 4.** Robustness analysis of the model. The deletion of each incoming interaction of a node was simulated and the number of node states and circuit functionality affected were computed.

Deletion of each incoming interaction to the node	Number of errors
SNAIL1	9
TWIST1	2
miR-129	16
miR-34	4
TGFB	0
miR-190	0
miR-1199	4
ZEB1	5
miR-200	8
ESRP1	0
CDH1	0
CD44s	0
HAS2	0
GRHL2	4
HA	0
VIM	0
miR-340	4

**Table 5.** Summary of findings of the present study and experimental and theoretical support found in the literature.

Findings	References
The EMT occurs through a stepwise process controlled by multiple circuits	[8]
The H state is characterized by the simultaneous presence of E and M phenotypic makers.	[8]
TGF- $\beta$ initiates the EMT via activation of SNAIL1 inducing the H states, while subsequent induction of ZEB1 maintains the M state	[8, 43]
The SNAIL1 negative circuit can preserve the E state.	[8, 37, 44]
The SNAIL1/miR-34 circuit controls the E-to-H transition via two bistable switches.	[8, 14, 43]
The TWIST1/miR-129 circuit is involved in the induction of an H state through an irreversible bistable switch.	Prediction
MiR-129 is a phenotypic stability factor of the H state	Prediction
The ZEB1-mediated circuits involving miR-340, miR-1199, GRHL2, and miR-200 account for the stabilization of one of the H states.	[8, 14, 43, 45]
The TWIST1 ON and miR-129 OFF bistable switch results from the functionality of ZEB1/miR-340, ZEB1/miR-1199, ZEB1/miR-200, and ZEB1/GRHL2 circuits.	Prediction
The TGFB/SNAIL1/miR-200 circuit is involved in the maintenance of the M state	Prediction
The ZEB1/HAS2/HA and ZEB1/ESRP1/CD44s circuits induce sustained ZEB1 activation maintaining the stability of the M state	[14, 32, 36]

and low amounts of TWIST1 and miR-129, respectively, can unravel the role of miR-1199, miR-200, miR-340, and, GRHL2 as predicted by the present study. Whether these molecules are active in these subpopulations, their role in the stabilization might be verified. Regarding the TGFB/SNAIL1/miR-200 circuit, the measurement of ZEB1 expression resulting from the

TGFB/SNAIL1/miR-200 axis might be a good strategy to verify the maintenance of the M state due to activity of the molecules in MCF10A cells. The possible outcome of this experiment would be sustained ZEB1 expression caused by the miR-200 downregulation-induced autocrine TGFB activation.

## Discussion

The presence of EMT-inducing molecules in the tumor microenvironment favors cancer metastatization. Cells stabilized in H and M states are able to migrate into circulatory systems originating metastasis in the host's body. Consequently, a complete understanding of the EMT process is indispensable to treat such a complex disease. A large number of studies reported that TGF- $\beta$ , one of the main EMT inducers, is frequently dysregulated in cancer [46, 47, 48]. Such dysregulation can trigger the initiation of EMT in epithelial cancer cells inducing metastasis. The TGF- $\beta$  pathway is characterized by its regulatory core composed by ZEB1 and SNAIL1 [8], and the control of the activity of these molecules is fundamental for the regulation of the mechanistic process of EMT. A recent paper published by our group suggested that new molecules could work as phenotypic stability factors [14]. However, the current model presents a more complex dynamics via consideration of different levels of exogenous TGF- $\beta$ . The modeling was mainly based on SNAIL1, the only target of the input in the proposed network. The proposal of its controlling logical equation took into consideration the effect of its regulators for each level of the response to exogenous TGF- $\beta$ . For instance, the influence of miR-34 and SNAIL1 self-inhibition on the logical equation were considered for the low dose (Ex\_TGFB=1) due to their ability to preserve the epithelial state, which is predominately observed in epithelial cells treated with concentrations of 0.5 ng/ml of TGF- $\beta$  [8]. The same idea was performed for the remaining levels of the input related to the logical equation controlling SNAIL1. The proposed logical equations were validated through comparison between the steady states produced by the model with experimental data. The only disagreement is related to the low dose of the input. The presence of both E and H states were observed experimentally [8], whereas the model only presented the E state. This result is associated with the direct influence of SNAIL1 self-inhibition as observed in Fig. 3. Indeed, Lu and colleagues [44], via a quantitative modeling, showed that the SNAIL1 negative circuit reduces the variation of the SNAIL1 level, triggering a monostable dynamics in the SNAIL1/miR-34 circuit corresponding to the E state. However, Zhang and colleagues [8] revealed that SNAIL1/miR-34 is bistable (characterizing the E and H states) even when the effect of SNAIL1 self-inhibition is considered. The authors revised a previously published model of EMT [43] due to its agreement with experimental data in order to include the SNAIL1 self-inhibition and other relevant biochemical information. In this way, our model is limited to show only the E state for the low dose of the input; however, this does not exclude the importance of our findings on the SNAIL1 negative circuit. Taken together, the results suggest that our model is a useful tool to study the different phenotypes induced by TGF- $\beta$  during EMT. The wild-type simulations showed stability of the E state only for the low dose and controlled by the SNAIL1 self-inhibitory circuit. LoF of this self-inhibitory interaction led to the transition to the H state. This result supports the idea of a threshold related to sustained activation of SNAIL1, as suggested by Peiro and colleagues [37].

In the present study, we used a discrete modeling approach to simulate the EMT process. Our results describe a dynamics consistent with the study of Zhang *et al.* [8]. Moreover, our model is coherent with other computational studies of EMT. Tian *et al.* [43] proposed a continuous model which predicts the stability of three stable states associated with E, H and M states. In addition, this work showed that mechanistic process of EMT occurs via two coupled bistable switches. The first reversible switch is associated with the SNAIL1/miR-34 circuit regulating the initiation of EMT, whereas the second irreversible one is based on the ZEB1/miR-200 circuit controlling the transition to the M state. The switches have the same behavior as those described in the present work for the SNAIL1/miR-34 and ZEB1/miR-200 circuits. The SNAIL1/miR-34 switch controls the E-to-H transition, whereas ZEB1/miR-200 regulates the transition to the M state. It was further

demonstrated that EMT proceeds through a stepwise process in which the E cell first transits to the H state and then to the M state [43]. This is observed in our model. The authors also reported that autocrine TGF- $\beta$  and miR-200 can contribute to the irreversibility observed in the second switch controlled by the miR-200/ZEB1 circuit. According to Tian *et al.* [43], the irreversibility is relative to the capacity of miR-200 to inhibit TGF- $\beta$ . That is, the second switch is irreversible if TGF- $\beta$  activation or or miR-200 inhibition is large enough [43]. Our results show that the inhibitory effect of SNAIL1 on miR-200 is responsible for the stabilization of the M state, which is coherent with the study of Tian and colleagues [43].

It has been widely studied theoretically and experimentally that cells can undergo either partial (hybrid) or complete EMT. Although our study shows the presence of two hybrid states during TGF- $\beta$ -induced EMT, this does not exclude the possibility of the existence of additional hybrid states. Font-Clos and colleagues reported the existence of multiple hybrid states during the EMT process [49]. The authors used a Boolean modeling approach, however the initial conditions differ from those used in the present study. Their network contemplated the crosstalk of several inducer pathways of EMT. In this way, the existence of multiple hybrid phenotypes related to this network may be associated with this crosstalk. Another computational study by Huang and colleagues [50] showed the presence of two hybrid states and a proposed network with 22 components including CDH1, ZEB1, VIM, and SNAIL1. These H states have correspondence with those observed in our study. Both H states were characterized with intermediate expression of CDH1 and VIM, and active SNAIL1. However, one of these states presents active ZEB1, which is not observed in our study.

The circuit analysis identified a new circuit, TGFB/SNAIL1/miR-200, which is not characterized in the literature in terms of its functionality. This circuit is associated to autocrine mechanisms in which the cell itself induces the expression of TGF- $\beta$  favoring the EMT induction. These mechanisms present an important role during EMT [51, 52, 53]. In the present study, the results showed that the new circuit contributes to the stability of the M state. The study of Gregory and colleagues reported an autocrine mechanism composed by endogenous TGF- $\beta$  and the ZEB1/miR-200 circuit [54]. The authors showed that TGF- $\beta$  is able to drive sustained ZEB1 expression via SNAIL1 activation favoring the stabilization of the M state. Our results describe this exact function, the activation of TGFB and SNAIL1 leads to the induction of ZEB1 and the inhibition of miR-200 activity, stabilizing the M state. Thus, our findings are consistent with published biochemical information. Unraveling the functionality of the new circuit might help to extend the knowledge of the mechanistic process of EMT. Hence, the control of such a functionality through the inhibition of the circuit components or the induction of the circuit controllers such as miR-190 might prevent cells of acquiring a complete mesenchymal phenotype.

The TWIST1/miR-129 double-negative circuit seems to present an important role during EMT. According to Fig. 4, the circuit can control the transition between the two H states via a bistable switch. The negative correlation between TWIST1 and miR-129 observed in this switch is consistent with experimental data using MCF10A cells [13] showing that miR-129 is negatively correlated with TWIST1 and SNAIL1 expression [13]. Also, this work showed that miR-129 expression was not reduced in TWIST1 or SNAIL1-depleted MCF10A cells treated with TGF- $\beta$ , indicating that both are required to inhibit the expression of miR-129. Indeed, we observed in Fig. 4 that when TWIST1 and SNAIL1 are active, miR-129 is not present. Another experimental evidence that supports our finding on the bistable switch of TWIST1/miR-129 circuit is related to the cooperation of SNAIL1 and TWIST1 during EMT. According to Tran and colleagues [55], SNAIL1 is uniquely required to initiate the TGF- $\beta$ -induced EMT, whereas TWIST1 is able to maintain the later phase of EMT characterized by the down and up-regulation of CDH1 and VIM, respectively. The bistable switch of the TWIST1/miR-129 circuit might be associated to the mechanism studied by Tran and colleagues [55]. One state of the bistable switch is characterized by the presence of SNAIL1, whereas in the other state both TWIST1 and SNAIL1 are active. Moreover, Yu and colleagues showed that miR-129 downregulation-induced TWIST1 activation in the maintenance of the EMT in breast epithelial cells [13]. In this way, the irreversibility of

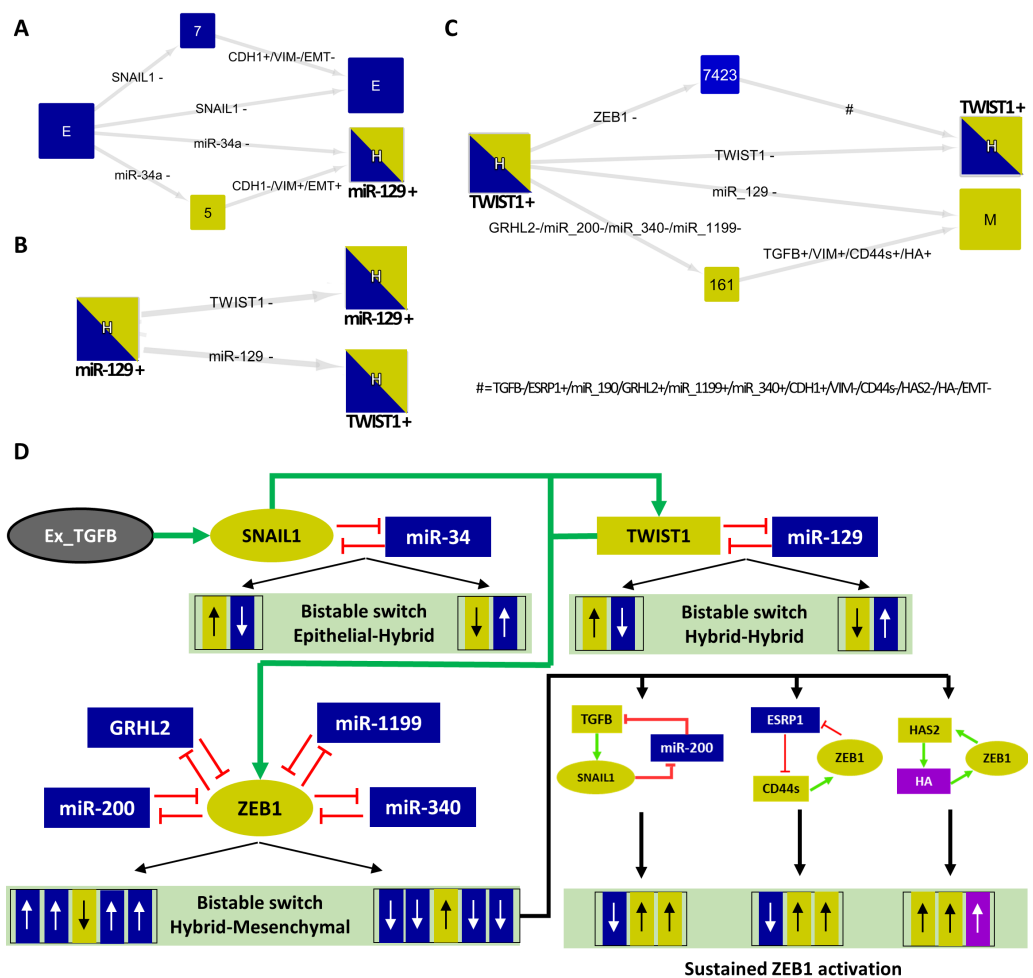


the transition associated with the presence of TWIST1 observed in our study also seems to be experimentally reliable. Regarding miR-129, its ability to stabilize the H state is an important finding of our study. This feature characterizes miR-129 as a PSF of the H phenotype. The identification of PSFs of the H state can expand the knowledge of the EMT-related hybrid phenotype. In this context, Jolly and colleagues identified two PSFs during EMT process via a mathematical modeling: GRHL2 and miR-145. The authors further observed that the knockdown of GRHL2 in H1975 cells leads to a full EMT, which emphasizes the importance of GRHL2 in the stabilization of the H state [45]. In this context, Watanabe and colleagues showed that cells can be maintained in a hybrid phenotype by the transcription factor OVOL, which characterizes this transcription factor as another PSF of the H state [56]. With the possible experimental validation of this prediction, this could help new strategies to treat the metastatic process. Drugs which could inhibit the expression of miR-129 might decrease the number of cells in the H phenotype in the carcinoma, since cells of this phenotype present up to 50 times higher metastatic potential than cells in the M state [4]. The possible lower number of cells in the H state triggered by miR-129 depletion might help to reduce the migration of breast cancer cells to other tissues via metastasis.

EMT is known to be reversible via the mesenchymal-epithelial transition (MET) process. However, many studies have suggested that, under specific conditions, EMT can become irreversible [57, 58]. Such conditions are based on E cells with continuous exposure to TGF- $\beta$  for a sufficiently long period of time. Jia and colleagues have reported that prolonged TGF- $\beta$  exposure (12-15 days) can give rise to irreversible mesenchymal cells [57]. The authors also reported that MCF10A cells underwent a stronger EMT process, characterizing the loss of E-cadherin (CDH1) and the up-regulation of ZEB1 expression. Moreover, Watanabe *et al.* [58] have shown that ZEB1 is essential to control the reversibility of EMT in MCF10A cells. In addition, the authors demonstrated that once the cells have reached the M state, EMT is irreversible, which is consistent with the mathematical model published by Tian and colleagues [43] and experimentally observed by Zhang *et al.* [8]. The E and M states of our model characterize a complete phenotypic transition. As seen in Fig. 5B, the M state does not transit back to the E state, which is consistent with previously discussed results from the literature. Besides, our findings on the TGFB/SNAIL1/miR-200 circuit-related reversibility predicts that this circuit might sustain ZEB1 activation which controls the irreversibility. In this way, our model also describes ZEB1 as an important player of the reversibility. Regarding TWIST1, also predicted as responsible for irreversibility in the model, there is a lack of experimental results on MCF10A cells. Nevertheless, Dragoi and colleagues reported that TWIST1 induces irreversible changes in the mesenchymal state of mammary epithelial cells (HMLE), resulting in the loss of dependency on TGF- $\beta$  [59]. This supports our findings on TWIST1 as an important player in the irreversibility.

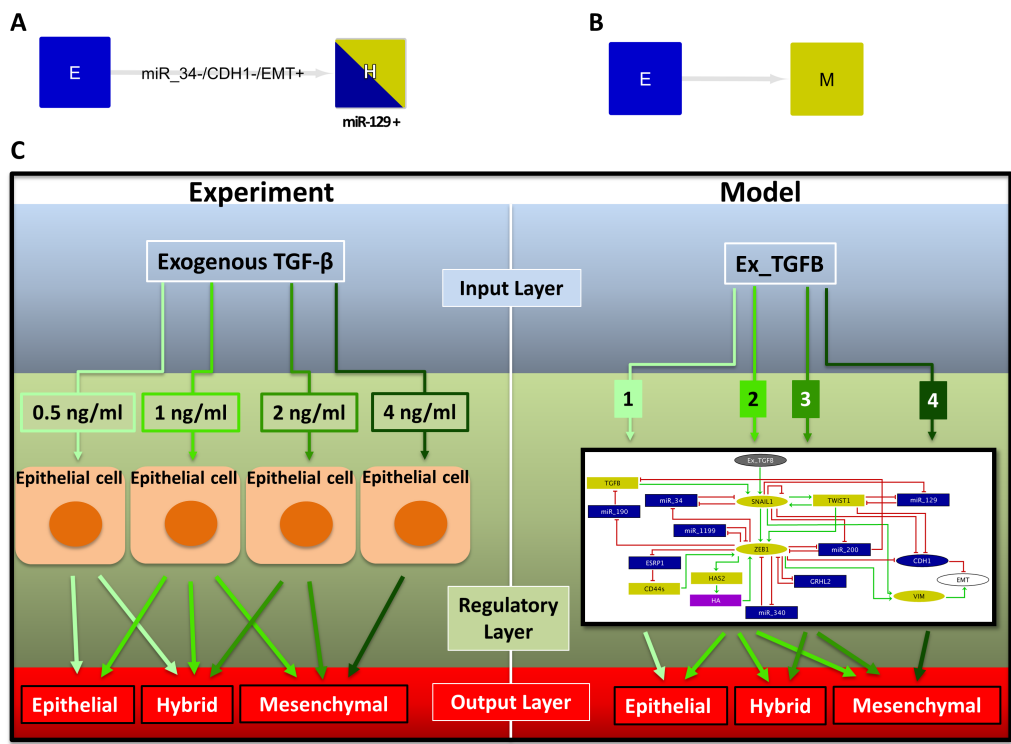
Synthetic biological networks are able to feature various dynamical behaviors such as multistability and oscillations through functional circuits. The model shows 9 functional positive circuits responsible for multistability and a negative one. Some circuits are responsible for the stabilization of specific transitions such as the SNAIL1/miR-34 (E-to-H transition), and TWIST1/miR-129 (H-to-H transition) circuits. The ZEB1-mediated circuits involving the protein GRHL2, and the miRNAs 1199, 200, and 340 control the H-to-M transition. The remaining positive circuits of the model sustain the stabilization of the M state. Our results show that loss of functionality of the positive circuits can affect the multistability of the model. For instance, LoFs of miR-34 and ZEB1 abrogate the E and M states, respectively. These perturbations lead to loss of functionality of their respective circuits affecting the multistability of the model. This suggests that the multistability is correlated with the number of positive circuits in the network, which is supported by the study of Hari *et al.* [60]. By considering all input levels, the multistability of the model triggered 6 H states, while the number of E and M states is lower. This is a signature of a frustrated dynamics for the H case according to the literature [61]. Indeed, Font-Clos and colleagues reported that the E and M states are less frustrated than the H state by using a Boolean modeling approach to study the EMT dynamics [49].

In summary, our findings provide new insights on the complex dynamics of TGF- $\beta$ -induced EMT via consideration of different concentrations of this molecule. The circuit



**Figure 6.** Stepwise process of the model dynamics via activation of multiple circuits. HTGs are displayed as in Fig. 3. The simulations were performed for Ex\_TGFB=2. (A) The initial state E (considering active SNAIL1) induces a bistability between the E and H (with the presence of miR-129) states. (B) Subsequently, the initial state H (considering active TWIST1) triggers another bistability between the H states. (C) Considering the H state with the presence of TWIST1 and additional activation of ZEB1 as initial condition, the simulation showed another multistability between H and M states. (D) Up and down arrows represent up and down-regulation, respectively. Yellow and blue colors denote E and M markers. The induction of SNAIL1/miR-34 circuit by Ex\_TGFB triggers bistable switches that control the transition E to H. The activation of SNAIL1 induces the functionality of TWIST1/miR-129 circuit. The switch TWIST1 OFF/miR-129 ON stabilizes one of the H states, whereas its counterpart (TWIST1 ON/miR-129 OFF) triggers the transition to another H state which is stable due to miR-200-miR-1199-GRHL2-miR-340-ZEB1 circuits. Active TWIST1 can induce ZEB1 and subsequently inactivate miR-200, GRHL2, miR-1199, and miR-340 and inducing the transition to the M state which is sustained by TGFB/miR-200/SNAIL1, ESRP1/CD44s/ZEB1 and HAS2/HA/ZEB1 circuits.

analysis demonstrated new important features about EMT dynamics such as functionalities of SNAIL1, TGFB/miR-200/SNAIL1, and TWIST1/miR-129 circuits during the transition. Further investigations of these features will extend our comprehension of TGF-β-induced EMT in MCF10A breast cells, which might help improve strategies for breast cancer treatment.



**Figure 7.** Schematic comparison between model and experimental observations. (A) The high levels of Ex\_TGFB induces a transition from the E to the H state with activated miR-129. (B) The very high level of Ex\_TGFB induces a direct transition from the E to the M state. (C) The model reproduces the experimental approach by Zhang *et al.* [8] where epithelial cells are treated with concentrations of 0.5, 1, and 2 ng/mL of exogenous TGF-β. Following the treatment, the stable phenotypes agree with the stable states found in the model dynamics. The disagreement related to the lower level of Ex\_TGFB is explained in section Discussion of the paper.

Data Accessibility

This article has no additional data.

Author Contributions

DAS, SG, and JCMM conceived the experiment(s); DAS and SG conducted the experiment(s); DAS, SG, and JCMM analyzed the results. All authors reviewed the manuscript.

Competing interests

The authors declare no conflict of interest.

Funding

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

## Acknowledgment

D.A. Silveira and S. Gupta acknowledge partial support from CNPq and CAPES, respectively.

## Supporting Information

- Supplementary Table S1: Logical rules of the model components and the corresponding published experimental validation.
- Supplementary file S1: The GINsim 3.0.0b code of the model.
- Supplementary file S2: Perturbation analysis of the model dynamics
- Supplementary Table S2: Analysis of model robustness

## References

- 1 Brabletz T, Kalluri R, Nieto MA, Weinberg RA. EMT in cancer. *Nature Reviews Cancer*. 2018;18(2):128.
- 2 Jolly MK, Somarelli JA, Sheth M, Biddle A, Tripathi SC, Armstrong AJ, et al. Hybrid epithelial/mesenchymal phenotypes promote metastasis and therapy resistance across carcinomas. *Pharmacology & therapeutics*. 2018.
- 3 Kröger C, Afeyan A, Mraz J, Eaton EN, Reinhardt F, Khodor YL, et al. Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells. *Proceedings of the National Academy of Sciences*. 2019;116(15):7353–7362.
- 4 Aceto N, Bardia A, Miyamoto DT, Donaldson MC, Wittner BS, Spencer JA, et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell*. 2014;158(5):1110–1122.
- 5 Tsai JH, Yang J. Epithelial–mesenchymal plasticity in carcinoma metastasis. *Genes & development*. 2013;27(20):2192–2206.
- 6 Yeung KT, Yang J. Epithelial–mesenchymal transition in tumor metastasis. *Molecular oncology*. 2017;11(1):28–39.
- 7 Jing Y, Han Z, Zhang S, Liu Y, Wei L. Epithelial-Mesenchymal Transition in tumor microenvironment. *Cell & bioscience*. 2011;1(1):29.
- 8 Zhang J, Tian XJ, Zhang H, Teng Y, Li R, Bai F, et al. TGF- $\beta$ -induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci Signal*. 2014;7(345):ra91–ra91.
- 9 Siemens H, Jackstadt R, Hüntgen S, Kaller M, Menssen A, Götz U, et al. miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. *Cell cycle*. 2011;10(24):4256–4271.
- 10 Brabletz S, Brabletz T. The ZEB/miR-200 feedback loop—a motor of cellular plasticity in development and cancer? *EMBO reports*. 2010;11(9):670–677.
- 11 Hou LK, Yu Y, Xie YG, Wang J, Mao JF, Zhang B, et al. miR-340 and ZEB1 negative feedback loop regulates TGF- $\beta$ -mediated breast cancer progression. *Oncotarget*. 2016;7(18):26016.
- 12 Diepenbruck M, Tiede S, Saxena M, Ivanek R, Kalathur RKR, Lüönd F, et al. miR-1199-5p and Zeb1 function in a double-negative feedback loop potentially coordinating EMT and tumour metastasis. *Nature communications*. 2017;8(1):1168.
- 13 Yu Y, Zhao Y, Sun XH, Ge J, Zhang B, Wang X, et al. Down-regulation of miR-129-5p via the Twist1-Snail feedback loop stimulates the epithelial-mesenchymal transition and is associated with poor prognosis in breast cancer. *Oncotarget*. 2015;6(33):34423.
- 14 Silveira DA, Mombach JCM. Dynamics of the feedback loops required for the phenotypic stabilization in the epithelial-mesenchymal transition. *The FEBS Journal*. 2020;287(3):578–588.
- 15 Le Novère N. Quantitative and logic modelling of molecular and gene networks. *Nature Reviews Genetics*. 2015;16(3):146–158.
- 16 Gupta S, Silveira DA, Barbé-Tuana FM, Mombach JCM. Integrative data modeling from lung and lymphatic cancer predicts functional roles for miR-34a and miR-16 in cell fate regulation. *Scientific Reports*. 2020;10(1):1–8.

- 17 Gupta S, Silveira DA, Mombach JCM. ATM/miR-34a-5p axis regulates a p21-dependent senescence-apoptosis switch in non-small cell lung cancer: a Boolean model of G1/S checkpoint regulation. *FEBS Letters*. 2020;594(2):227–239.
- 18 Rossato VV, Silveira DA, Gupta S, Mombach JCM. Towards the contribution of the p38MAPK pathway to the dual role of TGF $\beta$  in cancer: A boolean model approach. *Computers in biology and medicine*. 2019;104:235–240.
- 19 Gupta S, Silveira DA, Mombach JCM. Modeling the role of microRNA-449a in the regulation of the G2/M cell cycle checkpoint in prostate LNCaP cells under ionizing radiation. *PloS one*. 2018;13(7):e0200768.
- 20 Mombach JC, Vendrusculo B, Bugs CA. A model for p38MAPK-induced astrocyte senescence. *PloS one*. 2015;10(5).
- 21 Mooney SM, Talebian V, Jolly MK, Jia D, Gromala M, Levine H, et al. The GRHL2/ZEB feedback loop—A key axis in the regulation of EMT in breast cancer. *Journal of cellular biochemistry*. 2017;118(9):2559–2570.
- 22 Mombach JC, Bugs CA, Chaouiya C. Modelling the onset of senescence at the G1/S cell cycle checkpoint. *BMC genomics*. 2014;15(7):S7.
- 23 Gupta S, Silveira DA, Barbé-Tuana FM, Mombach JCM. Integrative data modeling from lung and lymphatic cancer predicts functional roles for miR-34a and miR-16 in cell fate regulation. *Scientific Reports*. 2020;10(2511):2511.
- 24 Gupta S, Silveira DA, Mombach JCM. Towards DNA-damage induced autophagy: A Boolean model of p53-induced cell fate mechanisms. *DNA Repair*. 2020. Available from: <https://doi.org/10.1016/j.dnarep.2020.102971>.
- 25 Silveira DA, Gupta S, Mombach JCM. p53/E2F1/miR-25 axis regulates apoptosis induction in glioblastoma cells: a qualitative model. *Journal of Physics: Complexity*. 2020;1(3):035001.
- 26 Abou-Jaoudé W, Traynard P, Monteiro PT, Saez-Rodriguez J, Helikar T, Thieffry D, et al. Logical modeling and dynamical analysis of cellular networks. *Frontiers in genetics*. 2016;7:94.
- 27 Thieffry D. Dynamical roles of biological regulatory circuits. *Briefings in bioinformatics*. 2007;8(4):220–225.
- 28 Li D, Gao J. Towards perturbation prediction of biological networks using deep learning. *Scientific Reports*. 2019;9(1):1–9.
- 29 Naldi A, Hernandez C, Abou-Jaoudé W, Monteiro PT, Chaouiya C, Thieffry D. Logical modeling and analysis of cellular regulatory networks with ginsim 3.0. *Frontiers in physiology*. 2018;9.
- 30 Bérenguier D, Chaouiya C, Monteiro PT, Naldi A, Remy E, Thieffry D, et al. Dynamical modeling and analysis of large cellular regulatory networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*. 2013;23(2):025114.
- 31 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 2003;13(11):2498–2504.
- 32 Preca BT, Bajdak K, Mock K, Sundararajan V, Pfannstiel J, Maurer J, et al. A self-enforcing CD 44s/ZEB 1 feedback loop maintains EMT and stemness properties in cancer cells. *International journal of cancer*. 2015;137(11):2566–2577.
- 33 Jolly MK, Preca BT, Tripathi SC, Jia D, George JT, Hanash SM, et al. Interconnected feedback loops among ESRP1, HAS2, and CD44 regulate epithelial-mesenchymal plasticity in cancer. *APL bioengineering*. 2018;2(3):031908.
- 34 Yu Y, Luo W, Yang ZJ, Chi JR, Li YR, Ding Y, et al. miR-190 suppresses breast cancer metastasis by regulation of TGF- $\beta$ -induced epithelial-mesenchymal transition. *Molecular cancer*. 2018;17(1):70.
- 35 Savci-Heijink C, Halfwerk H, Hooijer G, Koster J, Horlings H, Meijer S, et al. Epithelial-to-mesenchymal transition status of primary breast carcinomas and its correlation with metastatic behavior. *Breast cancer research and treatment*. 2019;174(3):649–659.
- 36 Preca BT, Bajdak K, Mock K, Lehmann W, Sundararajan V, Bronsert P, et al. A novel ZEB1/HAS2 positive feedback loop promotes EMT in breast cancer. *Oncotarget*. 2017;8(7):11530.

- 37 Peiro S, Escriva M, Puig I, Barbera MJ, Dave N, Herranz N, et al. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic acids research*. 2006;34(7):2077–2084.
- 38 Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, et al. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nature cell biology*. 2008;10(5):593–601.
- 39 Werner S, Frey S, Riethdorf S, Schulze C, Alawi M, Kling L, et al. Dual roles of the transcription factor grainyhead-like 2 (GRHL2) in breast cancer. *Journal of biological chemistry*. 2013;288(32):22993–23008.
- 40 Nairismägi ML, Vislovukh A, Meng Q, Kratassiouk G, Beldiman C, Petretich M, et al. Translational control of TWIST1 expression in MCF-10A cell lines recapitulating breast cancer progression. *Oncogene*. 2012;31(47):4960–4966.
- 41 Gras B, Jacquaroud L, Wierinckx A, Lamblot C, Fauvet F, Lachuer J, et al. Snail family members unequally trigger EMT and thereby differ in their ability to promote the neoplastic transformation of mammary epithelial cells. *PloS one*. 2014;9(3).
- 42 Chua H, Bhat-Nakshatri P, Clare S, Morimiya A, Badve S, Nakshatri H. NF- $\kappa$ B represses E-cadherin expression and enhances epithelial to mesenchymal transition of mammary epithelial cells: potential involvement of ZEB-1 and ZEB-2. *Oncogene*. 2007;26(5):711.
- 43 Tian XJ, Zhang H, Xing J. Coupled reversible and irreversible bistable switches underlying TGF $\beta$ -induced epithelial to mesenchymal transition. *Biophysical journal*. 2013;105(4):1079–1089.
- 44 Lu M, Jolly MK, Levine H, Onuchic JN, Ben-Jacob E. MicroRNA-based regulation of epithelial–hybrid–mesenchymal fate determination. *Proceedings of the National Academy of Sciences*. 2013;110(45):18144–18149.
- 45 Jolly MK, Tripathi SC, Jia D, Mooney SM, Celiktas M, Hanash SM, et al. Stability of the hybrid epithelial/mesenchymal phenotype. *Oncotarget*. 2016;7(19):27067.
- 46 Huang JJ, Blobel GC. Dichotomous roles of TGF- $\beta$  in human cancer. *Biochemical Society Transactions*. 2016;44(5):1441–1454.
- 47 Hargadon K. Dysregulation of TGF $\beta$ 1 activity in cancer and its influence on the quality of anti-tumor immunity. *Journal of clinical medicine*. 2016;5(9):76.
- 48 Wang X, Inoue S, Gu J, Miyoshi E, Noda K, Li W, et al. Dysregulation of TGF- $\beta$ 1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. *Proceedings of the National Academy of Sciences*. 2005;102(44):15791–15796.
- 49 Font-Clos F, Zapperi S, La Porta CA. Topography of epithelial–mesenchymal plasticity. *Proceedings of the National Academy of Sciences*. 2018;115(23):5902–5907.
- 50 Huang B, Lu M, Jia D, Ben-Jacob E, Levine H, Onuchic JN. Interrogating the topological robustness of gene regulatory circuits by randomization. *PLoS computational biology*. 2017;13(3):e1005456.
- 51 Scheel C, Eaton EN, Li SHJ, Chaffer CL, Reinhardt F, Kah KJ, et al. Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell*. 2011;145(6):926–940.
- 52 Larocca C, Cohen JR, Fernando RI, Huang B, Hamilton DH, Palena C. An autocrine loop between TGF- $\beta$ 1 and the transcription factor brachyury controls the transition of human carcinoma cells into a mesenchymal phenotype. *Molecular cancer therapeutics*. 2013;12(9):1805–1815.
- 53 Dumont N, Bakin AV, Arteaga CL. Autocrine transforming growth factor- $\beta$  signaling mediates Smad-independent motility in human cancer cells. *Journal of Biological Chemistry*. 2003;278(5):3275–3285.
- 54 Gregory PA, Bracken CP, Smith E, Bert AG, Wright JA, Roslan S, et al. An autocrine TGF- $\beta$ /ZEB/miR-200 signaling network regulates establishment and maintenance of epithelial–mesenchymal transition. *Molecular biology of the cell*. 2011;22(10):1686–1698.
- 55 Tran DD, Corsa CAS, Biswas H, Aft RL, Longmore GD. Temporal and spatial cooperation of Snail1 and Twist1 during epithelial–mesenchymal transition predicts for human breast cancer recurrence. *Molecular cancer research*. 2011;9(12):1644–1657.

- 56 Watanabe K, Villarreal-Ponce A, Sun P, Salmans ML, Fallahi M, Andersen B, et al. Mammary morphogenesis and regeneration require the inhibition of EMT at terminal end buds by Ovol2 transcriptional repressor. *Developmental cell*. 2014;29(1):59–74.
- 57 Jia W, Deshmukh A, Mani SA, Jolly MK, Levine H. A possible role for epigenetic feedback regulation in the dynamics of the epithelial–mesenchymal transition (EMT). *Physical biology*. 2019;16(6):066004.
- 58 Watanabe K, Panchy N, Noguchi S, Suzuki H, Hong T. Combinatorial perturbation analysis reveals divergent regulations of mesenchymal genes during epithelial-to-mesenchymal transition. *NPJ systems biology and applications*. 2019;5(1):1–15.
- 59 Dragoi D, Krattenmacher A, Mishra VK, Schmidt JM, Kloos UJ, Meixner LK, et al. Twist1 induces distinct cell states depending on TGFBR1-activation. *Oncotarget*. 2016;7(21):30396.
- 60 Hari K, Sabuwala B, Subramani BV, La Porta CA, Zapperi S, Font-Clos F, et al. Identifying inhibitors of epithelial–mesenchymal plasticity using a network topology-based approach. *NPJ systems biology and applications*. 2020;6(1):1–12.
- 61 Tripathi S, Kessler DA, Levine H. Biological Networks Regulating Cell Fate Choice are Minimally Frustrated. *Physical Review Letters*. 2020;125(8):088101.

## 4 DISCUSSÃO GERAL

O processo da transição epitélio-mesenquimal (TEM) é abordado como um mecanismo necessário para a metástase do câncer (HEERBOTH et al., 2015). Os princípios operacionais da TEM são baseados em células cancerosas epiteliais que sofrem uma alteração fenotípica mediada por um estímulo externo. Este mecanismo converte células epiteliais do câncer em células mesenquimais, permitindo a migração e invasão de outros tecidos do hospedeiro. Além disso, a estabilização de fenótipos híbridos podem aumentar o potencial metastático do câncer. Conseqüentemente, uma compreensão completa do processo da TEM é indispensável para tratar tal doença com alta complexidade.

Com a imensa complexidade dos sistemas biológicos em termos de sua natureza interativa, o desenvolvimento de modelos baseados em redes desses sistemas foi reconhecido como uma abordagem valiosa para estudar muitos processos biológicos (BURK et al., 2008; HEERBOTH et al., 2015; JOLLY et al., 2016; MENDES et al., 2018). Predições testáveis experimentalmente podem ser obtidas usando este método, as quais estão associadas a aspectos menos compreendidos da integração de análise de rede e modelagem dinâmica (ABOU-JAOUDE et al., 2016). Uma das principais características dos modelos qualitativos é estudar a influência de circuitos regulatórios (também conhecidos como loops de feedback), que são caminhos fechados entre dois ou mais nós da rede. Determinar a função reguladora de um circuito na rede pode ajudar a elucidar a decisão do destino da célula.

Um grande número de estudos relatou que o TGF- $\beta$ , um dos principais indutores da TEM, é frequentemente desregulado no câncer (HARGADON, 2016; HUANG; BLOBE, 2016; WANG et al., 2005). A via TGF- $\beta$  é caracterizada por seu núcleo regulatório composto por ZEB1 e SNAIL1 (ZHANG et al., 2014), e o controle da atividade dessas moléculas é fundamental para a regulação dessa transição. Através da ativação dos receptores de superfície correspondentes da membrana celular, a molécula TGF- $\beta$  desencadeia a indução SNAIL1 e ZEB1, os quais inibem uma molécula característica do fenótipo epitelial, a E-caderina codificada pelo gene CDH1. A indução dessas moléculas é controlada por circuitos característicos que regulam a dinâmica da TEM. Os circuitos mais estudados nesse processo são compostos por SNAIL1/2 e os microRNAs (miRNAs) 34a/b/c (coletivamente referidos como miR-34) (SIEMENS et al., 2011), e ZEB1/2 e os miRNAs-200a/b/c (coletivamente referidos como miR-200) (BURK et al., 2008). Tais circuitos formam "*loops*" chamados duplo-negativos (inibição mútua entre as moléculas). Recentemente, novas moléculas foram identificadas como reguladoras de ZEB1 durante a TEM induzida por TGF- $\beta$  em câncer de mama. Nesse contexto, foram encontrados dois novos circuitos positivos envolvendo novos meca-



nismos autócrinos relacionados à ZEB1: o primeiro envolvendo a molécula HAS2 e o ácido hialurônico sintetizado por HAS2 (HA) e o segundo relacionado à ESRP1 e CD44s (PRECA et al., 2015; PRECA et al., 2017). Ambos os circuitos foram identificados como mecanismos nos quais ZEB1 acelera a indução da TEM. Outra molécula que foi identificada como interagente com ZEB1 é a proteína GRHL2. Esta proteína está ligada a ZEB1 por meio de um circuito duplo-negativo, que foi descrito para controlar a progressão do tumor conduzida por TEM. GRHL2 também previne os padrões de expressão do estado M em células de câncer de mama associados a altos níveis de concentração de CD44 e baixos níveis de molécula-cula de CD24 (CD24) (CIEPLY et al., 2013; MOONEY et al., 2017). Além disso, novos miRNAs também foram descritos como interagentes com ZEB1, como miR-1199-5p, miR-340 e miR-190. O primeiro e o segundo participam de um circuito duplo-negativo com ZEB1 que controla a progressão do câncer de mama mediado por TEM (DIEPENBRUCK et al., 2017; HOU et al., 2016). O terceiro foi identificado como supressor do sinal de TGF- $\beta$  durante a TEM (YU et al., 2018). Além desse achado, o miR-190 também foi descrito como sendo transcricionalmente suprimido por ZEB1 (YU et al., 2018). Assim, essas novas descobertas revelam um cenário regulatório mais complexo para TEM responsivo ao TGF- $\beta$  e modelagem atualizada de demanda.

Baseando-se nas considerações acima descritas, o trabalho apresentado na seção 2 intitulado "*Dynamics of the feedback loops required for the phenotypic stabilization in the epithelial-mesenchymal transition*" tem como proposta um novo modelo da TEM baseado em dados recentes de moléculas e circuitos relacionados aos fatores de transcrição (FTs) SNAIL1 e ZEB1 (SILVEIRA; MOMBACH, 2020). Nesse estudo foram implementados ao modelo proposto por Zhang *et al.* 2014 (ZHANG et al., 2014) dois circuitos positivos envolvendo ZEB1 e as moléculas HAS2, HA, CD44s e ESRP1. Além disso, os miRNAs 340 e 1199 e a proteína GRHL2 foram introduzidas no modelo através de seus circuitos com ZEB1. O miRNA-190 e a molécula CD24 também foram consideradas na rede devido às suas importantes funções na inibição da sinalização de TGF- $\beta$  e ZEB1, respectivamente. Todos os dados implementados no modelo foram baseados na linhagem de célula de mama MCF10A.

Através da construção da rede regulatória, o modelo foi implementado no software GINsim para simulações da dinâmica usando o formalismo lógico. As regras lógicas foram atribuídas as componentes de acordo com observações teóricas e experimentais associadas ao processo da TEM. Por exemplo, de acordo com o estudo proposto por Zhang *et al.* 2014 (ZHANG et al., 2014), o aumento da atividade de SNAIL1 não induzia a diminuição de expressão do miR-200 através de sua interação de inibição. Após a ativação de ambos SNAIL1 e ZEB1, o miR-200 apresentava diminuição nos seus níveis. Esse resultado indica que esses FTs são essenciais na inibição do miR-200. Dessa forma, na construção da regra lógica do miR-200 foi im-

plementado o operador *OR* para inibição de SNAIL1 e ZEB1 de modo a reproduzir o efeito biológico observado.

A dinâmica tri-estável da TEM induzida por TGF- $\beta$  é observada experimentalmente em células epiteliais (MCF10A) tratadas com diferentes concentrações de TGF- $\beta$  de forma exógena (ZHANG et al., 2014). Os resultados observados no modelo indicam um comportamento semelhante relacionado à ativação do *input* (Ex\_TGFB), apresentando melhor concordância quantitativa com resultados associados com a concentração de 0,5 ng/ml de TGF- $\beta$  aplicada em células epiteliais MCF10A. Ambos, experimento e o modelo, encontram uma população de células do tipo H ligeiramente maior do que as outras duas (E e M).

A análise de perturbação das componentes dos circuitos funcionais previu um comportamento de comutação em todos os circuitos positivos, que parece ser responsável pelas transições entre os fenótipos. Os resultados sugerem que o circuito SNAIL1/miR-34 pode controlar a transição de E para H. Este processo foi observado na linhagem celular MCF10A (ZHANG et al., 2014). Os outros circuitos mediados por ZEB1 foram observados controlando a transição de H para M. Em particular, o circuito ZEB1/miR-200 foi experimentalmente observado induzindo essa transição (ZHANG et al., 2014). Além disso, o circuito ZEB1/HAS2/HA foi capaz de estabilizar o estado M de acordo com o modelo. Esse resultado foi observado experimentalmente (PRECA et al., 2017). Além disso, os resultados indicam que os circuitos ZEB1/miR-200/miR-340/miR-1199/GRHL2 têm um papel funcional na estabilização do estado H. Os circuitos envolvendo ZEB1, miR-200, GRHL2 e CD24 já foram observados experimentalmente com esta característica (MOONEY et al., 2017; ZHANG et al., 2014; JOLLY et al., 2016).

Os miRNAs 340 e 1199 foram observados como controladores da expressão de ZEB1 durante a sinalização de TGF- $\beta$  em células MCF10A. Além disso, ambos os miRNAs foram descritos como inibidores de migração e invasão devido à inativação de ZEB1. De acordo com o modelo, essas características estão presentes, incluindo novas relacionadas à estabilização do estado H, o qual pode apresentar potencial metastático 50 vezes maior do que as células no estado M (ACETO et al., 2014). De acordo com Zhang et al. (2014), as MCF10A, tratadas com 0,5 ng / ml de hTGF- $\beta$ 1, sofrem uma alta alteração fenotípica em direção ao fenótipo H. Hou e colegas demonstraram que esta mesma concentração de TGF- $\beta$ 1 não reduz significativamente a expressão do miR-340 em células MCF10A. Além disso, o estudo sugere que o circuito negativo ZEB1 / miR-340 torna-se funcional devido ao tratamento com TGF- $\beta$ 1 de forma dependente da dose e do tempo. Esses resultados indicam fortemente que miR-340 é importante para a estabilização do fenótipo H. Em relação ao miR-1199, Diepenbruck et al. (2017) demonstraram que na presença de TGF $\beta$  esta molécula apresenta funções semelhantes aos membros da família do miR-200 miR-200b-3p e

miR-429-3. A alta da expressão desses miRNAs desestabiliza a transição para o estado mesenquimal em células MCF10A. MiRNAs 1199 e 340 também apresentam um papel importante em linhas de células de câncer de mama. A regulação positiva da expressão de miR-340 inibe a progressão do câncer de mama na linha celular de câncer de mama MDA-MB-231 (HOU et al., 2016), enquanto o miR-1199 foi verificado para inibir ZEB1 na linhagem celular de câncer de mama de murino Py2T (DIEPENBRUCK et al., 2017). Além disso, na presença de TGF- $\beta$ , miR-1199 atrasa TEM em células Py2T (DIEPENBRUCK et al., 2017). Assim, nossas previsões sobre miR-340 e miR-1199 são consistentes com dados experimentais.

De maneira a expandir a análise da TEM durante a sinalização de TGF- $\beta$ , outro trabalho foi proposto intitulado "*Systems biology approach suggests new miRNAs as phenotypic stability factors in the epithelial-mesenchymal transition*" (seção 3) e tem como proposta um modelo baseado no trabalho da seção 2, contemplando múltiplos níveis de atividade para as componentes. O modelo apresenta uma dinâmica mais complexa por meio da consideração de diferentes níveis de TGF- $\beta$  exógeno. A modelagem foi baseada principalmente em SNAIL1, o único destino da entrada na rede proposta. A proposta de sua equação lógica de controle levou em consideração o efeito de seus reguladores para cada nível de resposta ao TGF- $\beta$  exógeno. Por exemplo, a influência da auto-inibição de miR-34 e SNAIL1 na equação lógica foi considerada para a dose baixa (Ex\_TGFB = 1) devido à sua capacidade de preservar o estado epitelial, que é predominantemente observado em células epiteliais tratadas com concentrações de 0,5 ng/ml de TGF- $\beta$  (ZHANG et al., 2014). A mesma ideia foi realizada para os demais níveis da entrada relacionados à equação lógica que controla SNAIL1. As equações lógicas propostas foram validadas através da comparação entre os estados estacionários produzidos pelo modelo com dados experimentais. A única discordância está relacionada à baixa dose do insumo. A presença dos estados E e H foi observada experimentalmente (ZHANG et al., 2014), enquanto o modelo apresentou apenas o estado E. Este resultado está associado à influência direta da auto-inibição de SNAIL. Na verdade, Lu e colegas (LU et al., 2013), por meio de uma modelagem quantitativa, mostraram que o circuito negativo SNAIL1 reduz a variação do nível de SNAIL1, desencadeando uma dinâmica mono-estável no circuito SNAIL1/miR-34 correspondente ao estado E. No entanto, Zhang e colegas (ZHANG et al., 2014) revelaram que SNAIL1/miR-34 é biestável (caracterizando os estados E e H) mesmo quando o efeito da auto-inibição de SNAIL1 é considerado. Os autores revisaram um modelo publicado anteriormente da TEM (TIAN; ZHANG; XING, 2013) devido à sua concordância com dados experimentais, a fim de incluir a auto-inibição de SNAIL1 e outras informações bioquímicas relevantes. Desta forma, o modelo se limita a mostrar apenas o estado E para a baixa dose da entrada; no entanto, isso não exclui a importância dos achados do presente estudo associados ao circuito negativo da SNAIL1.

Coletivamente, os resultados sugerem que o modelo é uma ferramenta útil para estudar os diferentes fenótipos induzidos por TGF- $\beta$  durante a TEM. As simulações do tipo selvagem mostraram estabilidade do estado E apenas para a dose baixa e controlado pelo circuito auto-inibitório SNAIL1. PF dessa interação auto-inibitória levou à transição para o estado H. Este resultado suporta a ideia de um limiar relacionado à ativação sustentada de SNAIL1, conforme sugerido por Peiro e colegas (PEIRO et al., 2006).

Os resultados descrevem uma dinâmica consistente com o estudo de Zhang *et al.* (ZHANG et al., 2014). Além disso, o modelo concorda com outros estudos computacionais relacionados à TEM. Tian *et al.* (TIAN; ZHANG; XING, 2013) propôs um modelo contínuo que prevê a estabilidade de três estados estáveis associados aos estados E, H e M. Além disso, o presente estudo mostrou que o processo mecânico de TEM ocorre por meio de duas chaves biestáveis acopladas. A primeira chave reversível está associada ao circuito SNAIL1/miR-34 que regula a iniciação de TEM, enquanto a segunda chave irreversível é baseada no circuito ZEB1/miR-200 que controla a transição para o estado M. As chaves possuem o mesmo comportamento descrito no presente trabalho para os circuitos SNAIL1/miR-34 e ZEB1/miR-200. A chave SNAIL1/miR-34 controla a transição E para H, enquanto ZEB1/miR-200 regula a transição para o estado M. Foi ainda demonstrado que a TEM prossegue através de um processo gradual no qual a célula E primeiro transita para o Estado H e depois para o estado M (TIAN; ZHANG; XING, 2013). Isso é observado no modelo. Os autores também relataram que o TGF- $\beta$  autócrina e miR-200 podem contribuir para a irreversibilidade observada na segunda troca controlada pelo circuito miR-200/ZEB1. De acordo com Tian, Zhang e Xing (2013), a irreversibilidade é relativa à capacidade do miR-200 de inibir o TGF- $\beta$ . Ou seja, a segunda opção é irreversível se a ativação da TGF- $\beta$  ou a inibição do miR-200 é grande o suficiente (TIAN; ZHANG; XING, 2013). Os resultados do presente estudo mostram que o efeito inibitório de SNAIL1 no miR-200 é responsável pela estabilização do estado M. Tal resultado concorda com o estudo de Tian e colaboradores (TIAN; ZHANG; XING, 2013).

A transição parcial (híbrida) durante a TEM está sendo amplamente estudado na literatura devido ao seu alto potencial metastático. Embora o modelo proposto apresente dois estados híbridos, isso não exclui a possibilidade da existência de estados híbridos adicionais. Font-Clos e colegas relataram a existência de múltiplos estados híbridos durante a TEM (FONT-CLOS; ZAPPERI; PORTA, 2018). Os autores utilizaram uma abordagem de modelagem booleana, porém as condições iniciais diferem das utilizadas no presente trabalho. A rede desse estudo contemplou várias vias indutoras da TEM. Desta forma, a existência de múltiplos fenótipos híbridos relacionados a esta rede pode estar associada a esta metodologia. Outro estudo computacional de Huang e colegas (HUANG et al., 2017) demonstrou a presença de dois estados

híbridos através de uma rede proposta com 22 componentes, incluindo CDH1, ZEB1, VIM e SNAIL1. Esses estados H têm correspondência com os observados no modelo proposto. Ambos os estados H foram caracterizados com expressão intermediária de CDH1 e VIM, e SNAIL1 ativo. Porém, um desses estados apresenta a ativação de ZEB1, que não foi observada no presente estudo.

O modelo identificou um novo circuito, TGFB/SNAIL1/miR-200, que não é caracterizado na literatura quanto à sua funcionalidade. Este circuito está associado à expressão autócrina, na qual a própria célula induz a expressão de TGF- $\beta$  favorecendo a indução da TEM. Esses mecanismos apresentam um papel importante durante a TEM (SCHEEL et al., 2011; LAROCCA et al., 2013; DUMONT; BAKIN; ARTEAGA, 2003). No presente estudo, os resultados mostraram que o novo circuito contribui para a estabilidade do estado M. O estudo de Gregory e colegas relatou uma expressão autócrina composta por TGF- $\beta$  endógena e o circuito ZEB1/miR-200 (GREGORY et al., 2011). Os autores demonstraram que TGF- $\beta$  é capaz de induzir a expressão sustentada de ZEB1 por meio da ativação de SNAIL1, favorecendo a estabilização do estado M. Os resultados do modelo descrevem exatamente esta função; a ativação de TGFB e SNAIL1 leva à indução de ZEB1 e à inibição da atividade do miR-200, estabilizando o estado M. Portanto, os resultados são consistentes com as informações bioquímicas publicadas. Desvendar a funcionalidade do novo circuito pode ajudar a ampliar o conhecimento do processo mecânico da TEM. Portanto, o controle de tal funcionalidade por meio da inibição dos componentes do circuito ou da indução dos controladores do circuito, como miR-190, pode impedir que as células adquiram um fenótipo mesenquimal completo.

O circuito positivo TWIST1/miR-129 pode apresentar um papel importante durante a TEM. O circuito pode controlar a transição entre os dois estados H de acordo com o presente estudo. A correlação negativa entre TWIST1 e miR-129 observada neste controle é consistente com dados experimentais usando células MCF10A (YU et al., 2015). Isso demonstra que o miR-129 está negativamente correlacionado com TWIST1 e a expressão de SNAIL1 (YU et al., 2015). Além disso, este trabalho mostrou que a expressão de miR-129 não foi reduzida em células MCF10A deficientes de TWIST1 ou SNAIL e tratadas com TGF- $\beta$ , indicando que ambos são necessários para inibir a expressão de miR-129. De fato, os resultados demonstram que quando TWIST1 e SNAIL1 estão ativos, o miR-129 não está presente. Outra evidência experimental que é consistente com o resultado do modelo sobre a relação negativa entre TWIST1 e miR-129 está relacionada à cooperação de SNAIL1 e TWIST1 durante a TEM. De acordo com Tran e colegas (TRAN et al., 2011), SNAIL1 é exclusivamente necessária para iniciar a TEM induzida por TGF  $\beta$ , enquanto TWIST1 é capaz de sustentar fases mais avançadas da TEM. Isso é caracterizado pela regulação negativa e positiva de CDH1 e VIM, respectivamente. O comportamento apresentado pelo

circuito TWIST1/miR-129 no modelo pode estar associado ao mecanismo estudado por Tran e colegas (TRAN et al., 2011). Um estado do *switch* biestável (comportamento do circuito na estabilização dos estados H) é caracterizado pela presença de SNAIL1, enquanto no outro estado TWIST1 e SNAIL1 estão ativos. Além disso, Yu e colegas mostraram que a ativação de TWIST1 induzida pela regulação negativa do miR-129 é importante para a manutenção da TEM em células epiteliais de mama (YU et al., 2015). Dessa forma, a irreversibilidade da transição associada à presença de TWIST1 observada no presente estudo também parece ser experimentalmente confiável. Com relação ao miR-129, sua capacidade de estabilizar o estado H é um achado importante. Este comportamento caracteriza o miR-129 como um fator de estabilidade fenotípica (FEF) do fenótipo H. A identificação de FEFs do estado H pode expandir o conhecimento do fenótipo híbrido relacionado à TEM. Nesse contexto, Jolly e colegas identificaram dois FEFs durante o processo de TEM por meio de uma modelagem matemática: GRHL2 e miR-145. Os autores observaram ainda que a inativação de GRHL2 em células H1975 leva a uma TEM completa. Isso enfatiza a importância de GRHL2 na estabilização do estado H (JOLLY et al., 2016). Nesse contexto, Watanabe e colaboradores mostraram que as células podem ser mantidas em um fenótipo híbrido pelo fator de transcrição OVOL, o que caracteriza esse fator de transcrição como outro FEF do estado H (WATANABE et al., 2014). Com a possível validação experimental dessa previsão, isso poderia ajudar em novas estratégias para tratar o processo metastático. Drogas que poderiam inibir a expressão do miR-129 podem diminuir o número de células do fenótipo H no carcinoma, uma vez que as células desse fenótipo apresentam potencial metastático até 50 vezes maior do que as células no estado M (ACETO et al., 2014). O possível menor número de células no estado H desencadeado pela depleção de miR-129 pode ajudar a reduzir a migração de células de câncer de mama para outros tecidos por meio de metástases.

A TEM é conhecida por ser reversível por meio do processo de transição de uma célula mesenquimal para uma epitelial. No entanto, muitos estudos têm sugerido que, em condições específicas, a TEM pode se tornar irreversível (JIA et al., 2019; WATANABE et al., 2019). Tais condições são baseadas em células E com exposição contínua a TGF- $\beta$  por um período de tempo suficientemente longo. Jia e colegas relataram que a exposição prolongada a TGF- $\beta$  (12-15 dias) pode dar origem a células mesenquimais irreversíveis (JIA et al., 2019). Os autores também relataram que as células MCF10A passaram por um processo TEM mais contundente, caracterizando a perda de E-caderina (CDH1) e a regulação positiva da expressão de ZEB1. Além disso, Watanabe *et al.* (WATANABE et al., 2019) mostraram que ZEB1 é essencial para controlar a reversibilidade de TEM em células MCF10A. Além disso, os autores demonstraram que uma vez que as células atingem o estado M, a TEM é irreversível, que é consistente com o modelo matemático publicado por Tian e colegas (TIAN;

ZHANG; XING, 2013) e experimentalmente observado por Zhang *et al.* (ZHANG *et al.*, 2014). Os estados E e M do modelo do presente estudo caracterizam uma transição fenotípica completa. O estado M não transita de volta para o estado E, que é consistente com os resultados discutidos anteriormente na literatura. Além disso, os resultados sobre a reversibilidade relacionada ao circuito TGFB/SNAIL1/miR-200 prediz que este circuito pode sustentar a ativação de ZEB1 que controla a irreversibilidade. Desta forma, o modelo também descreve ZEB1 como um importante jogador da reversibilidade. Em relação à TWIST1, também previsto como responsável pela irreversibilidade no modelo, faltam resultados experimentais em células MCF10A. No entanto, Dragoi e colegas relataram que TWIST1 induz alterações irreversíveis no estado mesenquimal das células epiteliais mamárias, resultando na perda de dependência de TGF- $\beta$  (DRAGOI *et al.*, 2016). Isso é consistente com os achados do presente estudo sobre a TWIST1 ser importante na irreversibilidade.

Redes biológicas são capazes de apresentar vários comportamentos dinâmicos, como multiestabilidade e oscilações através dos circuitos funcionais. O modelo mostra 9 circuitos positivos funcionais responsáveis pela multiestabilidade e um negativo. Alguns circuitos são responsáveis pela estabilização de transições específicas, como os circuitos SNAIL1/miR-34 (transição E para H) e TWIST1/miR-129 (transição entre os estados H). Os circuitos mediados por ZEB1 envolvendo a proteína GRHL2 e os miRNAs 1199, 200 e 340 controlam a transição entre H e M. Os demais circuitos positivos do modelo sustentam a estabilização do estado M. Os resultados mostram que a perda de funcionalidade dos circuitos positivos pode afetar a multiestabilidade do modelo. Por exemplo, as PFs do miR-34 e ZEB1 anulam os estados E e M, respectivamente. Essas perturbações levam à perda de funcionalidade de seus respectivos circuitos afetando a multiestabilidade do modelo. Isso sugere que a multiestabilidade está correlacionada com o número de circuitos positivos na rede, que é corroborado pelo estudo de Hari *et al.* (HARI *et al.*, 2020). Ao considerar todos os níveis do *input*, a multiestabilidade do modelo mostrou 6 estados H, enquanto o número de estados E e M é menor. Esta é a assinatura de uma dinâmica frustrada para o caso H de acordo com a literatura (TRIPATHI; KESSLER; LEVINE, 2020). De fato, Font-Clos e colegas relataram que os estados E e M são menos frustrados do que o estado H, usando uma abordagem de modelagem booleana para estudar a dinâmica da TEM (FONT-CLOS; ZAPPERI; PORTA, 2018).

O formalismo lógico, baseado na modelagem discreta, fornece resultados qualitativos da dinâmica de uma rede. Sua aplicação é pertinente em sistemas com informações limitadas sobre as taxas de reações bioquímicas. Entretanto, a discretização do sistema limita o modelo através do negligenciamento do comportamento transitório entre o estado inicial e final de um determinado processo. Por exemplo, um modelo booleano pode descrever apenas a dinâmica das componentes entre dois estados

(ativo e inativo). Níveis intermediários entre a transição de tais estados são negligenciados. Assim, modelos contínuos são mais adequados para uma descrição transitória completa entre os estados,

Coletivamente, os achados do presente trabalho fornecem ideias teóricas de novos mecanismos associados à estabilidade dos estados durante a complexa dinâmica da TEM. Para isto, foram propostos modelos computacionais através de uma modelagem discreta. Tais modelos foram validados usando comparações entre as perturbações das componentes e suas correspondentes análises experimentais. Dessa forma, é possível investigar o comportante resultante e verificar a concordância entre os resultados observados. Ambos os modelos demonstraram uma excelente concordância com os dados experimentais e, assim, podem ser considerados ferramentas confiáveis para estudar a TEM. As importantes previsões teóricas observadas em tais modelos englobam a descoberta de novos FEFs para o estado H (miRNAs 340, 1199, 340 e 129), novos mecanismo autócrinos envolvendo TGF- $\beta$  e a descrição da evolução da dinâmica através da funcionalidade de cada circuito positivo relacionando com a estabilização dos estados observados.



## 5 CONCLUSÃO

A caracterização da transição epitélio-mesenquimal (TEM) através das redes biológicas, sua modelagem e análise de sua dinâmica tem sido amplamente utilizado na literatura. Nesse contexto, a presente tese apresentou dois estudos envolvendo a modelagem da via TGF- $\beta$  relacionada à ativação da TEM, utilizando dados bioquímicos recentes de maneira a expandir a análise dinâmica da rede. Além disso, foram explorados o comportamento dinâmico de microRNAs (miRNAs), a funcionalidade dos circuitos funcionais e as probabilidades de acessibilidade dos atratores. A partir dos resultados obtidos nesses estudos, as seguintes conclusões podem ser formuladas:

- I. Os dois novos modelos computacionais propostos demonstram uma excelente concordância quantitativa com os dados experimentais associados à células MCF10A e, assim, são modelos confiáveis para estudar a TEM em diferentes contextos biológicos;
- II. A descrição da dinâmica dos estados através da funcionalidade coordenada dos circuitos facilitam a compreensão das transições entre os estados, contribuindo para o desenvolvimento de novas estratégias terapêuticas para o câncer;
- III. Os miRNAs 129, 340 e 1199 podem ser importantes na estabilização do fenótipo híbrido;
- IV. A funcionalidade do circuito positivo entre TWIST1 e miR-129 pode controlar transições entre estados H durante a sinalização de TGF- $\beta$ ;
- V. O achado relacionado ao mecanismo autócrino envolvendo TGF- $\beta$ , SNAIL1 e miR-200 sugere uma nova ideia teórica que pode contribuir para o entendimento da irreversibilidade da TEM;
- VI. Todas as previsões simuladas no modelo foram detalhadamente estudadas e apresentaram uma boa fundamentação teórica e, dessa forma, fornecem excelentes perspectivas para futuros experimentos desenvolvidos para testar tais resultados.

Além dos estudos apresentados na presente tese de doutorado, outros seis artigos foram publicados em periódicos científicos durante a realização do projeto. Tais artigos estão listados no apêndice D.

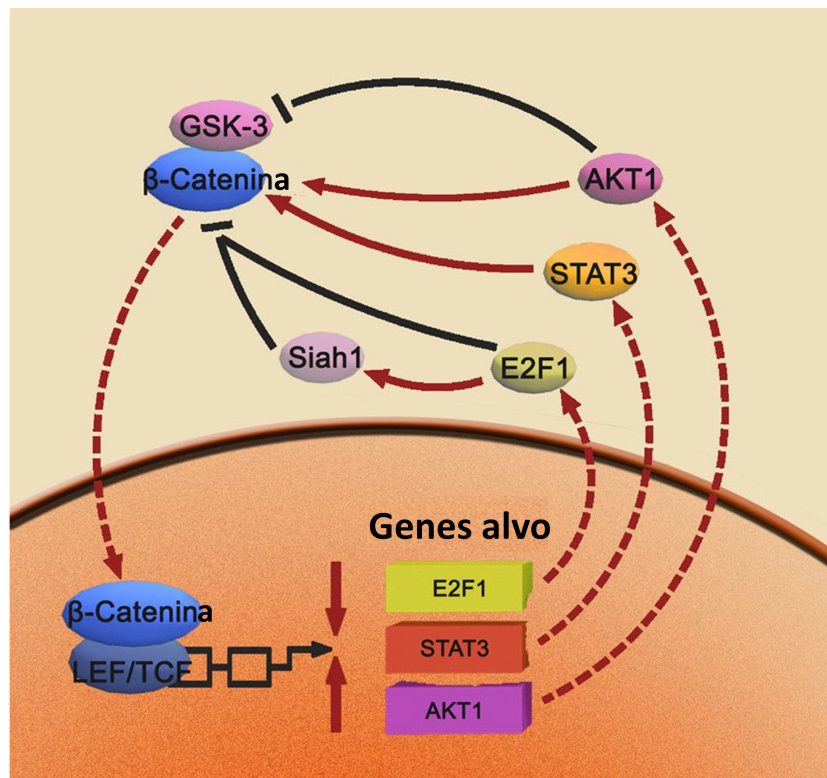
## 5.1 PERSPECTIVAS FUTURAS

Os artigos apresentados nesta tese de doutorado são as primeiras etapas para o desenvolvimento de novas pesquisas associadas aos mecanismos de indução da TEM. Um importante aspecto nos modelos dos artigos é a possibilidade da implementação de RNAs longos não-codificantes (LncRNAs), os quais são RNAs não codificantes que podem regular miRNAs. De acordo com Lin, Lin e Yang (2016), LncRNAs podem apresentar um importante papel durante a TEM.

A modelagem de novas vias de ativação da TEM também amplifica o conhecimento científico sobre tal processo. A via da família de moléculas Wnt, por exemplo, é constantemente estudada na TEM devido a sua frequente ativação em câncer (ZHAN; RINDTORFF; BOUTROS, 2017). A rede que governa a sinalização em cascata da Wnt apresenta uma complexa variedade de circuitos através de uma de suas moléculas alvo chamada  $\beta$ -catenina (figura 5.1). A partir da ativação da via da WNT, a molécula GSK-3 é inativada e permite a migração da  $\beta$ -catenina até o núcleo celular, que, por sua vez, liga-se a LEF/TCF de maneira a induzir os genes alvo STAT3 e AKT1. A partir disso, as proteínas STAT3 e AKT1 induzem a migração de  $\beta$ -catenina para o núcleo formando, assim, um circuito positivo entre essas moléculas. Por outro lado, a E2F1, inibidora de  $\beta$ -catenina, forma um circuito negativo nessa rede através da interação direta ou indireta via Siah1 com  $\beta$ -catenina, que por sua vez, pode inibir E2F1 através de seu complexo com LEF/TCF.

Devido a complexidade da rede em termos das interação e circuitos formados, a via Wnt torna-se um processo biológico relevante para uma abordagem de modelagem computacional. Dessa forma, a análise de sua dinâmica pode evidenciar novos mecanismos relacionados à TEM em câncer e, assim, auxiliar no aprimoramento de novas estratégias para o seu tratamento.

Figura 5.1 – Circuitos entre a componente da via Wnt  $\beta$ -catenina e suas moléculas alvo. O complexo  $\beta$ -catenina/LEF/TCF promove a expressão AKT1 e STAT3. A ativação de AKT1 induz a fosforilação de  $\beta$ -catenina de forma direta ou indireta via fosforilação de GSK-3. Por sua vez, ativação de STAT3 aumenta a acumulação  $\beta$ -catenina no núcleo celular, aumentando desse modo sua atividade transcricional. Além disso, a atividade da molécula de E2F1 é reprimida pela expressão do complexo  $\beta$ -catenina/LEF/TCF, enquanto que E2F1 pode inibir a transcrição de  $\beta$ -catenina diretamente ou através da regulação positiva de Siah1.



Fonte: Adaptado de Huang et al. (2010).

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABBA, M. et al. Microrna regulation of epithelial to mesenchymal transition. **Journal of clinical medicine**, Multidisciplinary Digital Publishing Institute, v. 5, n. 1, p. 8, 2016.
- ABOU-JAOUDÉ, W. et al. Logical Modeling and Dynamical Analysis of Cellular Networks. **Frontiers in Genetics**, v. 7, p. 94, 2016.
- ACETO, N. et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. **Cell**, Elsevier, v. 158, n. 5, p. 1110–1122, 2014.
- ARNOSTI, D. N.; AY, A. Boolean modeling of gene regulatory networks: Driesch redux. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 109, n. 45, p. 18239–18240, 2012.
- ARNOUX, V. et al. Cutaneous wound reepithelialization. rise and fall of epithelial phenotype. **Springer US**, p. 111–34, 2005.
- BANZHAF, W.; KUO, P. D. Network motifs in natural and artificial transcriptional regulatory networks. **Journal of Biological Physics and Chemistry**, Citeseer, v. 4, p. 85–92, 2004.
- BRABLETZ, S.; BRABLETZ, T. The zeb/mir-200 feedback loop a motor of cellular plasticity in development and cancer? **EMBO reports**, John Wiley & Sons, Ltd, v. 11, n. 9, p. 670–677, 2010.
- BRABLETZ, T. et al. Emt in cancer. **Nature Reviews Cancer**, Nature Publishing Group, v. 18, n. 2, p. 128, 2018.
- BURGER, G. A.; DANEN, E. H.; BELTMAN, J. B. Deciphering epithelial–mesenchymal transition regulatory networks in cancer through computational approaches. **Frontiers in oncology**, Frontiers, v. 7, p. 162, 2017.
- BURK, U. et al. A reciprocal repression between zeb1 and members of the mir-200 family promotes emt and invasion in cancer cells. **EMBO reports**, John Wiley & Sons, Ltd Chichester, UK, v. 9, n. 6, p. 582–589, 2008.
- CAI, Y. et al. A brief review on the mechanisms of miRNA regulation. **Genomics, proteomics & bioinformatics**, Elsevier, v. 7, n. 4, p. 147–154, 2009.
- CALZONE, L.; BARILLOT, E.; ZINOVYEV, A. Logical versus kinetic modeling of biological networks: applications in cancer research. **Current Opinion in Chemical Engineering**, Elsevier, v. 21, p. 22–31, 2018.
- CHAOUIYA, C.; REMY, E. **Logical modelling of regulatory networks, methods and applications**. [S.l.]: Springer, 2013.
- CHEN, Y.-R.; HUANG, H.-C.; LIN, C.-C. Regulatory feedback loops bridge the human gene regulatory network and regulate carcinogenesis. **Briefings in bioinformatics**, 2017.

CHOI, S. et al. A nuclear phosphoinositide kinase complex regulates p53. **Nature cell biology**, Nature Publishing Group, p. 1, 2019.

CIEPLY, B. et al. Epithelial–mesenchymal transition and tumor suppression are controlled by a reciprocal feedback loop between zeb1 and grainyhead-like-2. **Cancer research**, AACR, v. 73, n. 20, p. 6299–6309, 2013.

CRAENE, B. D.; BERX, G. Regulatory networks defining EMT during cancer initiation and progression. **Nature Reviews Cancer**, Nature Publishing Group, v. 13, n. 2, p. 97, 2013.

DALCHAU, N. et al. Computing with biological switches and clocks. **Natural computing**, Springer, v. 17, n. 4, p. 761–779, 2018.

DAVIDSON, E.; LEVIN, M. Gene regulatory networks. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 102, n. 14, p. 4935–4935, 2005.

DENU, J. M. et al. Form and function in protein dephosphorylation. **Cell**, Elsevier, v. 87, n. 3, p. 361–364, 1996.

DIEPENBRUCK, M. et al. miR-1199-5p and Zeb1 function in a double-negative feedback loop potentially coordinating EMT and tumour metastasis. **Nature communications**, Nature Publishing Group, v. 8, n. 1, p. 1168, 2017.

DRAGOI, D. et al. Twist1 induces distinct cell states depending on tgfr1-activation. **Oncotarget**, Impact Journals, LLC, v. 7, n. 21, p. 30396, 2016.

DUMONT, N.; BAKIN, A. V.; ARTEAGA, C. L. Autocrine transforming growth factor- $\beta$  signaling mediates smad-independent motility in human cancer cells. **Journal of Biological Chemistry**, ASBMB, v. 278, n. 5, p. 3275–3285, 2003.

FABREGAT, A. et al. The reactome pathway knowledgebase. **Nucleic acids research**, Oxford University Press, v. 46, n. D1, p. D649–D655, 2018.

FONT-CLOS, F.; ZAPPERI, S.; PORTA, C. A. L. Topography of epithelial–mesenchymal plasticity. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 115, n. 23, p. 5902–5907, 2018.

GAGNIUC, P. A. **Markov chains: from theory to implementation and experimentation**. [S.I.]: John Wiley & Sons, 2017.

GARG, M. Epithelial, mesenchymal and hybrid epithelial/mesenchymal phenotypes and their clinical relevance in cancer metastasis. **Expert reviews in molecular medicine**, Cambridge University Press, v. 19, 2017.

GITTER, A. et al. Linking the signaling cascades and dynamic regulatory networks controlling stress responses. **Genome research**, Cold Spring Harbor Lab, v. 23, n. 2, p. 365–376, 2013.

GONZALEZ, A. G. et al. GINsim: a software suite for the qualitative modelling, simulation and analysis of regulatory networks. **Biosystems**, Elsevier, v. 84, n. 2, p. 91–100, 2006.

GREGORY, P. A. et al. An autocrine  $\text{tgf-}\beta/\text{zeb}/\text{mir-200}$  signaling network regulates establishment and maintenance of epithelial-mesenchymal transition. **Molecular biology of the cell**, Am Soc Cell Biol, v. 22, n. 10, p. 1686–1698, 2011.

GRINSTEAD, C. M.; SNELL, J. L. **Introduction to probability**. [S.l.]: American Mathematical Soc., 2012.

GUPTA, S. et al. Integrative data modeling from lung and lymphatic cancer predicts functional roles for mir-34a and mir-16 in cell fate regulation. **Scientific Reports**, Nature Publishing Group, v. 10, n. 1, p. 1–8, 2020.

GUPTA, S.; SILVEIRA, D. A.; MOMBACH, J. C. M. Modeling the role of microRNA-449a in the regulation of the G2/M cell cycle checkpoint in prostate LNCaP cells under ionizing radiation. **PloS one**, Public Library of Science, v. 13, n. 7, p. e0200768, 2018.

GUPTA, S.; SILVEIRA, D. A.; MOMBACH, J. C. M. ATM/miR-34a-5p axis regulates a p21-dependent senescence-apoptosis switch in non-small cell lung cancer: a boolean model of g1/s checkpoint regulation. **FEBS letters**, Wiley Online Library, v. 594, n. 2, p. 227–239, 2020.

GUPTA, S.; SILVEIRA, D. A.; MOMBACH, J. C. M. Towards dna-damage induced autophagy: A boolean model of p53-induced cell fate mechanisms. **DNA repair**, Elsevier, v. 96, p. 102971, 2020.

HAAPANIEMI, E. et al. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. **Nature medicine**, Nature Publishing Group, v. 24, n. 7, p. 927, 2018.

HALLINAN, J.; JACKWAY, P. T. Network motifs, feedback loops and the dynamics of genetic regulatory networks. In: IEEE. **2005 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology**. [S.l.], 2005. p. 1–7.

HARGADON, K. Dysregulation of  $\text{tgf}\beta 1$  activity in cancer and its influence on the quality of anti-tumor immunity. **Journal of clinical medicine**, Multidisciplinary Digital Publishing Institute, v. 5, n. 9, p. 76, 2016.

HARI, K. et al. Identifying inhibitors of epithelial–mesenchymal plasticity using a network topology-based approach. **NPJ systems biology and applications**, Nature Publishing Group, v. 6, n. 1, p. 1–12, 2020.

HAY, E. D. An overview of epithelio-mesenchymal transformation. **Cells Tissues Organs**, Karger Publishers, v. 154, n. 1, p. 8–20, 1995.

HEERBOTH, S. et al. Emt and tumor metastasis. **Clinical and translational medicine**, SpringerOpen, v. 4, n. 1, p. 1–13, 2015.

HELIKAR, T. et al. The cell Collective: toward an open and collaborative approach to systems biology. **BMC systems biology**, BioMed Central, v. 6, n. 1, p. 96, 2012.

HOLLSTEIN, M. et al. p53 mutations in human cancers. **Science**, American Association for the Advancement of Science, v. 253, n. 5015, p. 49–53, 1991.

HOU, L.-K. et al. miR-340 and ZEB1 negative feedback loop regulates TGF- $\beta$ -mediated breast cancer progression. **Oncotarget**, Impact Journals, LLC, v. 7, n. 18, p. 26016, 2016.

HUANG, B. et al. Interrogating the topological robustness of gene regulatory circuits by randomization. **PLoS computational biology**, Public Library of Science, v. 13, n. 3, p. e1005456, 2017.

HUANG, J. J.; BLOBE, G. C. Dichotomous roles of  $\text{tgf-}\beta$  in human cancer. **Biochemical Society Transactions**, Portland Press Limited, v. 44, n. 5, p. 1441–1454, 2016.

HUANG, K. et al. MicroRNA roles in beta-catenin pathway. **Molecular cancer**, BioMed Central, v. 9, n. 1, p. 252, 2010.

ISER, I. C. et al. The epithelial-to-mesenchymal transition-like process in glioblastoma: An updated systematic review and in silico investigation. **Medicinal research reviews**, Wiley Online Library, v. 37, n. 2, p. 271–313, 2017.

ISSLER, M. V. C.; MOMBACH, J. C. M. MicroRNA-16 feedback loop with p53 and Wip1 can regulate cell fate determination between apoptosis and senescence in DNA damage response. **PloS one**, Public Library of Science, v. 12, n. 10, p. e0185794, 2017.

JIA, D. et al. Distinguishing mechanisms underlying EMT tristability. **Cancer convergence**, Nature Publishing Group, v. 1, n. 1, p. 2, 2017.

JIA, W. et al. A possible role for epigenetic feedback regulation in the dynamics of the epithelial–mesenchymal transition (emt). **Physical biology**, IOP Publishing, v. 16, n. 6, p. 066004, 2019.

JOLLY, M. K. et al. Stability of the hybrid epithelial/mesenchymal phenotype. **Oncotarget**, Impact Journals, LLC, v. 7, n. 19, p. 27067, 2016.

JONOSKA, N.; SAITO, M. **Discrete and topological models in molecular biology**. [S.l.]: Springer, 2013.

KALLURI, R.; WEINBERG, R. A. The basics of epithelial-mesenchymal transition. **The Journal of clinical investigation**, Am Soc Clin Investig, v. 119, n. 6, p. 1420–1428, 2009.

KANEHISA, M. et al. Kegg: integrating viruses and cellular organisms. **Nucleic Acids Research**, Oxford University Press, v. 49, n. D1, p. D545–D551, 2021.

KAUFFMAN, S. A. Metabolic stability and epigenesis in randomly constructed genetic nets. **Journal of theoretical biology**, Elsevier, v. 22, n. 3, p. 437–467, 1969.

KAUFMAN, M.; SOULE, C.; THOMAS, R. A new necessary condition on interaction graphs for multistationarity. **Journal of Theoretical Biology**, Elsevier, v. 248, n. 4, p. 675–685, 2007.

KORPAL, M.; KANG, Y. The emerging role of miR-200 family of microRNAs in epithelial-mesenchymal transition and cancer metastasis. **RNA biology**, Taylor & Francis, v. 5, n. 3, p. 115–119, 2008.

KROESE, D. P. et al. Why the Monte Carlo method is so important today. **Wiley Interdisciplinary Reviews: Computational Statistics**, Wiley Online Library, v. 6, n. 6, p. 386–392, 2014.

LAROCCA, C. et al. An autocrine loop between  $\text{tgf-}\beta 1$  and the transcription factor brachyury controls the transition of human carcinoma cells into a mesenchymal phenotype. **Molecular cancer therapeutics**, AACR, v. 12, n. 9, p. 1805–1815, 2013.

LATCHMAN, D. S. Transcription factors: an overview. **The international journal of biochemistry & cell biology**, Elsevier, v. 29, n. 12, p. 1305–1312, 1997.

LEE, J. M. et al. The epithelial–mesenchymal transition: new insights in signaling, development, and disease. **J Cell Biol**, Rockefeller University Press, v. 172, n. 7, p. 973–981, 2006.

LEE, R. C.; FEINBAUM, R. L.; AMBROS, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. **cell**, Elsevier, v. 75, n. 5, p. 843–854, 1993.

LI, R. et al. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. **Cell stem cell**, Elsevier, v. 7, n. 1, p. 51–63, 2010.

LI, Y.; VARALA, K.; CORUZZI, G. M. From milliseconds to lifetimes: tracking the dynamic behavior of transcription factors in gene networks. **Trends in Genetics**, Elsevier, v. 31, n. 9, p. 509–515, 2015.

LIN, C.-W.; LIN, P.-Y.; YANG, P.-C. Noncoding RNAs in tumor epithelial-to-mesenchymal transition. **Stem cells international**, Hindawi, v. 2016, 2016.

LING, H.; FABBRI, M.; CALIN, G. A. MicroRNAs and other non-coding RNAs as targets for anticancer drug development. **Nature reviews Drug discovery**, Nature Publishing Group, v. 12, n. 11, p. 847, 2013.

LIU, D.; ALBERGANTE, L.; NEWMAN, T. J. Universal attenuators and their interactions with feedback loops in gene regulatory networks. **Nucleic acids research**, Oxford University Press, v. 45, n. 12, p. 7078–7093, 2017.

LODISH, H. et al. **Molecular Cell Biology**. [S.l.]: W. H. Freeman and Company, 2013.

LU, M. et al. MicroRNA-based regulation of epithelial–hybrid–mesenchymal fate determination. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 110, n. 45, p. 18144–18149, 2013.

MARCUCCI, F.; STASSI, G.; MARIA, R. D. Epithelial–mesenchymal transition: a new target in anticancer drug discovery. **Nature reviews Drug discovery**, Nature Publishing Group, v. 15, n. 5, p. 311, 2016.

MENDES, N. D. et al. Estimating attractor reachability in asynchronous logical models. **Frontiers in physiology**, Frontiers Media SA, v. 9, 2018.

MENDEZ, M. G.; KOJIMA, S.-I.; GOLDMAN, R. D. Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition. **The FA-**



**SEB Journal**, Federation of American Societies for Experimental Biology, v. 24, n. 6, p. 1838–1851, 2010.

MICALIZZI, D. S.; FARABAUGH, S. M.; FORD, H. L. Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression. **Journal of mammary gland biology and neoplasia**, Springer, v. 15, n. 2, p. 117–134, 2010.

MILO, R. et al. Network motifs: simple building blocks of complex networks. **Science**, American Association for the Advancement of Science, v. 298, n. 5594, p. 824–827, 2002.

MOLL, U. M.; PETRENKO, O. The MDM2-p53 interaction. **Molecular cancer research**, AACR, v. 1, n. 14, p. 1001–1008, 2003.

MOONEY, S. M. et al. The grhl2/zeb feedback loop a key axis in the regulation of emt in breast cancer. **Journal of cellular biochemistry**, Wiley Online Library, v. 118, n. 9, p. 2559–2570, 2017.

MÜSSEL, C.; HOPFENSITZ, M.; KESTLER, H. A. BoolNetan R package for generation, reconstruction and analysis of Boolean networks. **Bioinformatics**, Oxford University Press, v. 26, n. 10, p. 1378–1380, 2010.

NALDI, A. et al. Logical modeling and analysis of cellular regulatory networks with GINsim 3.0. **Frontiers in physiology**, Frontiers Media SA, v. 9, 2018.

NALDI, A.; THIEFFRY, D.; CHAOUIYA, C. Decision diagrams for the representation and analysis of logical models of genetic networks. In: SPRINGER. **International Conference on Computational Methods in Systems Biology**. [S.l.], 2007. p. 233–247.

NOVERE, N. L. Quantitative and logic modelling of molecular and gene networks. **Nature Reviews Genetics**, Nature Publishing Group, v. 16, n. 3, p. 146, 2015.

O'BRIEN, J. et al. Overview of microRNA biogenesis, mechanisms of actions, and circulation. **Frontiers in endocrinology**, Frontiers, v. 9, p. 402, 2018.

OUGHTRED, R. et al. The biogrid database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. **Protein Science**, Wiley Online Library, v. 30, n. 1, p. 187–200, 2021.

PAIS, R. J. F. T. G. **How cells initiate Epithelial-to-Mesenchymal Transition?** 2018. 187 p. Tese (Doutorado em Biologica de sistemas) — Universidade Nova de Lisboa, Lisboa, 2018.

PEIRO, S. et al. Snail1 transcriptional repressor binds to its own promoter and controls its expression. **Nucleic acids research**, Oxford University Press, v. 34, n. 7, p. 2077–2084, 2006.

PIGOLOTTI, S.; KRISHNA, S.; JENSEN, M. H. Oscillation patterns in negative feedback loops. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 104, n. 16, p. 6533–6537, 2007.

POVEY, S. et al. The hugo gene nomenclature committee (hgnc). **Human genetics**, Springer-Verlag, v. 109, n. 6, p. 678–680, 2001.

PRECA, B.-T. et al. A self-enforcing cd 44s/zeb 1 feedback loop maintains emt and stemness properties in cancer cells. **International journal of cancer**, Wiley Online Library, v. 137, n. 11, p. 2566–2577, 2015.

PRECA, B.-T. et al. A novel zeb1/has2 positive feedback loop promotes emt in breast cancer. **Oncotarget**, Impact Journals, LLC, v. 8, n. 7, p. 11530, 2017.

RAY, P. D.; HUANG, B.-W.; TSUJI, Y. Reactive oxygen species (ros) homeostasis and redox regulation in cellular signaling. **Cellular signalling**, Elsevier, v. 24, n. 5, p. 981–990, 2012.

REMY, E. et al. A modeling approach to explain mutually exclusive and co-occurring genetic alterations in bladder tumorigenesis. **Cancer research**, AACR, v. 75, n. 19, p. 4042–4052, 2015.

REVENU, C.; GILMOUR, D. EMT 2.0: shaping epithelia through collective migration. **Current opinion in genetics & development**, Elsevier, v. 19, n. 4, p. 338–342, 2009.

RODRIGUEZ-BOULAN, E.; MACARA, I. G. Organization and execution of the epithelial polarity programme. **Nature reviews Molecular cell biology**, Nature Publishing Group, v. 15, n. 4, p. 225, 2014.

ROSSATO, V. V. et al. Towards the contribution of the p38MAPK pathway to the dual role of TGF $\beta$  in cancer: A boolean model approach. **Computers in biology and medicine**, Elsevier, v. 104, p. 235–240, 2019.

RUBIN, C. S.; ROSEN, O. M. Protein phosphorylation. **Annual review of biochemistry**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 44, n. 1, p. 831–887, 1975.

SAMAVARCHI-TEHRANI, P. et al. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. **Cell stem cell**, Elsevier, v. 7, n. 1, p. 64–77, 2010.

SÁNCHEZ, L.; CHAOUIYA, C. Primary sex determination of placental mammals: a modelling study uncovers dynamical developmental constraints in the formation of Sertoli and granulosa cells. **BMC systems biology**, BioMed Central, v. 10, n. 1, p. 37, 2016.

SAVAGNER, P. Epithelial–mesenchymal transitions: from cell plasticity to concept elasticity. In: **Current topics in developmental biology**. [S.l.]: Elsevier, 2015. v. 112, p. 273–300.

SCHEEL, C. et al. Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. **Cell**, Elsevier, v. 145, n. 6, p. 926–940, 2011.

SHALGI, R. et al. Global and local architecture of the mammalian microRNA–transcription factor regulatory network. **PLoS computational biology**, Public Library of Science, v. 3, n. 7, p. e131, 2007.

SHIN, Y.-J. et al. Post-translational regulation enables robust p53 regulation. **BMC systems biology**, BioMed Central, v. 7, n. 1, p. 83, 2013.

SHOOK, D.; KELLER, R. Mechanisms, mechanics and function of epithelial–mesenchymal transitions in early development. **Mechanisms of development**, Elsevier, v. 120, n. 11, p. 1351–1383, 2003.

SIEMENS, H. et al. miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. **Cell cycle**, Taylor & Francis, v. 10, n. 24, p. 4256–4271, 2011.

SILVEIRA, D. A.; GUPTA, S.; MOMBACH, J. C. M. p53/e2f1/mir-25 axis regulates apoptosis induction in glioblastoma cells: a qualitative model. **Journal of Physics: Complexity**, IOP Publishing, v. 1, n. 3, p. 035001, 2020.

SILVEIRA, D. A.; GUPTA, S.; MOMBACH, J. C. M. Systems biology approach suggests new mirnas as phenotypic stability factors in the epithelial–mesenchymal transition. **Journal of the Royal Society Interface**, The Royal Society, v. 17, n. 171, p. 20200693, 2020.

SILVEIRA, D. A.; MOMBACH, J. C. M. Dynamics of the feedback loops required for the phenotypic stabilization in the epithelial-mesenchymal transition. **The FEBS Journal**, Wiley Online Library, v. 287, n. 3, p. 578–588, 2020.

SMITH, A. J. et al. Beyond picomolar affinities: quantitative aspects of noncovalent and covalent binding of drugs to proteins. **Journal of medicinal chemistry**, ACS Publications, v. 52, n. 2, p. 225–233, 2009.

SOULÉ, C. Mathematical approaches to differentiation and gene regulation. **Comptes rendus biologiques**, Elsevier, v. 329, n. 1, p. 13–20, 2006.

THOMAS, R. Boolean formalization of genetic control circuits. **Journal of theoretical biology**, Elsevier, v. 42, n. 3, p. 563–585, 1973.

THOMAS, R.; D'ARI, R. **Biological feedback**. [S.I.]: CRC press, 1990.

TIAN, X.-J.; ZHANG, H.; XING, J. Coupled reversible and irreversible bistable switches underlying TGF $\beta$ -induced epithelial to mesenchymal transition. **Biophysical journal**, Elsevier, v. 105, n. 4, p. 1079–1089, 2013.

TONELLO, E.; FARCOT, E.; CHAOUIYA, C. Local negative circuits and cyclic attractors in Boolean networks with at most five components. **SIAM Journal on Applied Dynamical Systems**, SIAM, v. 18, n. 1, p. 68–79, 2019.

TRAN, D. D. et al. Temporal and spatial cooperation of snail1 and twist1 during epithelial–mesenchymal transition predicts for human breast cancer recurrence. **Molecular cancer research**, AACR, v. 9, n. 12, p. 1644–1657, 2011.

TRIPATHI, S.; KESSLER, D. A.; LEVINE, H. Biological networks regulating cell fate choice are minimally frustrated. **Physical Review Letters**, APS, v. 125, n. 8, p. 088101, 2020.

TWEEDIE, S. et al. Genenames. org: the hgnc and vgnc resources in 2021. **Nucleic Acids Research**, Oxford University Press, v. 49, n. D1, p. D939–D946, 2021.

VIJESH, N.; CHAKRABARTI, S. K.; SREEKUMAR, J. Modeling of gene regulatory networks: a review. **Journal of Biomedical Science and Engineering**, Scientific Research Publishing, v. 6, n. 02, p. 223, 2013.

WALHOUT, A. J. Gene-centered regulatory network mapping. In: **Methods in cell biology**. [S.l.]: Elsevier, 2011. v. 106, p. 271–288.

WANG, X. et al. Dysregulation of  $\text{tgf-}\beta\text{1}$  receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 102, n. 44, p. 15791–15796, 2005.

WATANABE, K. et al. Combinatorial perturbation analysis reveals divergent regulations of mesenchymal genes during epithelial-to-mesenchymal transition. **NPJ systems biology and applications**, Nature Publishing Group, v. 5, n. 1, p. 1–15, 2019.

WATANABE, K. et al. Mammary morphogenesis and regeneration require the inhibition of emt at terminal end buds by *ovol2* transcriptional repressor. **Developmental cell**, Elsevier, v. 29, n. 1, p. 59–74, 2014.

XIAO, K. et al. Role of p120-catenin in cadherin trafficking. **Biochimica et Biophysica Acta (BBA)-Molecular Cell Research**, Elsevier, v. 1773, n. 1, p. 8–16, 2007.

YANG, J.; WEINBERG, R. A. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. **Developmental cell**, Elsevier, v. 14, n. 6, p. 818–829, 2008.

YU, Y. et al. mir-190 suppresses breast cancer metastasis by regulation of  $\text{tgf-}\beta\text{-}$ induced epithelial–mesenchymal transition. **Molecular cancer**, BioMed Central, v. 17, n. 1, p. 70, 2018.

YU, Y. et al. Down-regulation of mir-129-5p via the twist1-snail feedback loop stimulates the epithelial-mesenchymal transition and is associated with poor prognosis in breast cancer. **Oncotarget**, Impact Journals, LLC, v. 6, n. 33, p. 34423, 2015.

ZHAN, T.; RINDTORFF, N.; BOUTROS, M. Wnt signaling in cancer. **Oncogene**, Nature Publishing Group, v. 36, n. 11, p. 1461, 2017.

ZHANG, J. et al. TGF- $\beta$ -induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. **Sci. Signal.**, American Association for the Advancement of Science, v. 7, n. 345, p. ra91–ra91, 2014.

## APÊNDICE A – DINÂMICA DE UM ATRATOR CÍCLICO

O exemplo matemático relacionado à dinâmica de um atrator cíclico apresentado nesse apêndice foi baseado no livro de Thomas e d'Ari (1990). As notações foram adequadas de acordo com a referência. Dessa forma, considerou-se termos em negrito como funções e variáveis lógicas e termos sublinhados como genes e seus produtos.

Considerando-se um sistema simples de dois elementos, supondo que o produto  $x$  ativa o gene  $\underline{Y}$  e o produto  $y$  inibe o gene  $\underline{X}$  (circuito negativo simples; veja figura 1.3). Assim,

$$\mathbf{X} = 1 \text{ se } y = 0 \text{ (}\underline{X}\text{ está ligado se } y \text{ é ausente)} \quad (\text{A.1})$$

$$\mathbf{Y} = 1 \text{ se } x = 1 \text{ (}\underline{Y}\text{ está ligado se } x \text{ é presente)} \quad (\text{A.2})$$

A relação lógica do sistema desencadeada pelas funções (A.1) e (A.2) pode ser escrita como:

$$\mathbf{X} = \bar{y} \quad (\text{A.3})$$

$$\mathbf{Y} = x \quad (\text{A.4})$$

em que a barra sobre  $y$  denota o operador lógico NOT.

A relação dos valores das variáveis das eqs. (A.3) e (A.4) são representados na Tabela A.1.

Tabela A.1 – Relação lógica dos valores das variáveis  $x$  e  $y$ , e suas respectivas funções  $\mathbf{X}$  e  $\mathbf{Y}$ .

$x$	$y$	$\mathbf{X}$	$\mathbf{Y}$
0	0	1	0
0	1	0	0
1	1	0	1
1	0	1	1

A parte esquerda da tabela anterior representa todas as combinações possíveis  $2^n$  entre os estados (0 ou 1) de  $x$  e  $y$ .

Denotando-se o estado das variáveis através de um vetor lógico  $xy$ , podendo assumir os valores 00, 01, 11, e 10, e o estado da função lógica como um vetor  $\mathbf{XY}$ , tem-se que no sistema:

$$X = \Phi_1(x, y, \dots) \quad (\text{A.5})$$

$$Y = \Phi_2(x, y, \dots) \quad (\text{A.6})$$

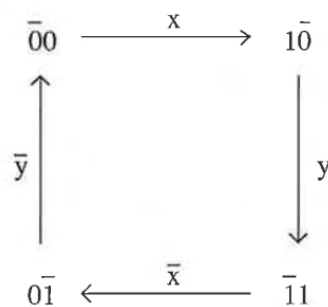
o vetor  $XY$  é obtido através da aplicação da função lógica  $\Phi$  no vetor  $xy$ . Assim,  $XY$  é chamado de imagem do vetor  $xy$ . O estado do sistema é, assim, escrito como  $xy/XY$ . Considerando-se, por exemplo, o estado  $00/01$ , no qual  $\underline{x}$  e  $\underline{y}$  são ausentes, mas o gene  $\underline{X}$  está no estado 1. Dessa forma, é esperado que produto  $\underline{x}$  seja sintetizado, significando que  $x$  transitará de 0 para o valor lógico 1. Assim, é possível simplificar a notação do estado do sistema  $00/01$  para  $\bar{0}0$ , no qual a barra sobre o primeiro número significa que  $x$  irá trocar de estado. Generalizando essa notação, pode-se escrever a Tabela A.2.

Tabela A.2 – Notação em termos da evolução dos estados da Tabela A.1.

$x$	$y$	$X$	$Y$
$\bar{0}$	$0$	1	0
$0$	$\bar{1}$	0	0
$\bar{1}$	$1$	0	1
$1$	$\bar{0}$	1	1

De acordo com a Tabela A.2, a sequência temporal de estados do sistema é representado pela figura A.1.

Figura A.1 – Sequência temporal de estados do atrator cíclico.



Fonte: Thomas e d'Ari (1990).

Portanto, o sistema permanece aprisionado em um atrator cíclico, variando seus estados de acordo com as funções lógicas estabelecidas.

## APÊNDICE B – FUNDAMENTOS TEÓRICOS DO MÉTODO DE CADEIAS DE MARKOV

As cadeias de Markov modelam uma sequência de estados de um determinado para os quais a descrição do estado atual captura totalmente todas as informações que podem influenciar a evolução futura do processo. No caso de cadeias absorventes, existem estados finais (absorventes) que não evoluem para outros. Teoricamente, para uma cadeia de Markov  $X$  de  $l$  estados absorventes e  $t$  estados transitórios, a matriz de transição  $\pi$  pode ser escrita na sua forma canônica como (GRINSTEAD; SNELL, 2012):

$$\pi = \begin{pmatrix} Q & L \\ 0 & I \end{pmatrix} \quad (\text{B.1})$$

em termos:

- (a)  $Q$  é uma matriz  $t \times t$ ;
- (b)  $L$  é uma matriz não nula denotada a partir de  $t \times l$ ;
- (c)  $0$  é uma matriz nula obtida de  $l \times t$ ;
- (d)  $I$  é a matriz identidade definida a partir de  $l \times l$ .

Partindo de um estado  $u$  para um estado  $v$  depois de um número de etapas  $k$  na cadeia  $X$ , a matriz de transição pode ser escrita nessa configuração como:

$$\pi^k(u, v) = \begin{pmatrix} Q^k & (\sum_{j=0}^{k-1} Q^j)L \\ 0 & I \end{pmatrix}. \quad (\text{B.2})$$

Além disso, os seguintes teoremas para as cadeias absorventes de Markov são válidos (GRINSTEAD; SNELL, 2012):

- (a)  $Q^k \rightarrow 0$  se  $k \rightarrow \infty$ ;
- (b) a matriz  $I - Q$  possui uma inversa, e

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n Q^k = (I - Q)^{-1}, \quad (\text{B.3})$$

em que a entrada  $(i, j)$  de  $(I - Q)^{-1}$  fornece o número esperado de vezes que o processo está no estado transitório  $v$ , iniciando a partir de  $u$ ;

- (c) A probabilidade da cadeia  $X$  alcançar um estado absorvente  $a$  a partir de um transitório  $u$  é

$$P(X_\infty = a) = (I - Q)^{-1}L. \quad (\text{B.4})$$

As supramencionadas definições matemáticas associadas às cadeias de Markov são a base para os fundamentos teóricos relacionados ao método do Avatar (MENDES et al., 2018). Resumidamente, o algoritmo inicia seguindo trajetórias aleatórias ao longo da dinâmica assíncrona no gráfico de transição de estados (GTE) conforme a eq. (1.3) e o método de Monte Carlo. Uma matriz de transição similar a eq. (B.1) descreve a dinâmica dos estados até a identificação, caso exista, de um ciclo transitente. A partir disso, o algoritmo modifica o GTE (baseando-se nas eqs. (B.2), (B.3) e (B.4)), desfazendo as suas trajetórias de interações e ligando seus estados aos outros vetores alvos do ciclo. As probabilidades de saída dos estados são recalculadas de maneira a assegurar uma avaliação correta das probabilidades de acessibilidade, executando um elevado número de trajetórias aleatórias sobre uma cadeia de Markov  $X$ . Assim, este procedimento equivale a modificar a cadeia e não altera as propriedades assintóticas de saída da simulação (MENDES et al., 2018).



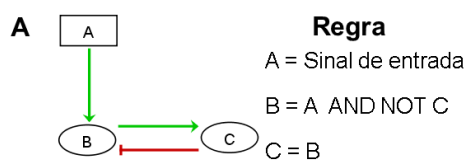
## APÊNDICE C – RELAÇÃO ENTRE REGRAS LÓGICAS E CONTEXTO DE FUNCIONALIDADE DOS CIRCUITOS

As regras lógicas definidas para as componentes do circuito implicam diretamente nas suas restrições de funcionalidade de acordo com os seus reguladores. Cada restrição baseia-se na formulação da regra lógica de modo que a relação entre as funções lógicas das componentes do circuito e dos seus respectivos reguladores sejam determinadas de acordo com suas dependências. Na figura C.1, por exemplo, a rede composta pelos nós A, B e C e suas respectivas regras e funções lógicas possui uma restrição na funcionalidade do circuito negativo B/C através do regulador A. A aplicação do operador AND entre as funções de A e C aplicadas na componente B implica em uma dependência de A na condição de atividade de B. Como a funcionalidade do circuito depende da atividade das componentes que o compõem, a atividade do regulador A é uma restrição para a ativação do nó B. Assim, para uma regulação ativatória entre A e B (figura C.1A), o estado ativo de A é necessário para a funcionalidade do circuito negativo B/C (figura C.1B), enquanto que para uma relação inibitória (figura C.1C), o circuito B/C será funcional se A permanecer inativo (figura C.1D).

A troca de operadores (AND por OR) no sistema da figura C.2 implica na independência de A na condição de atividade de B, desencadeando uma inversão nos valores lógicos dos contextos de funcionalidade relacionados ao nó A. Essa inversão ocorre porque os estados dos regulares seriam suficientes para perturbar os padrões dinâmicos produzidos pelas componentes do circuito (multiestáveis ou oscilatório). Por exemplo, no caso de A ativando B (figura C.2A), o estado inativo de A é necessário para a funcionalidade do circuito negativo B/C (figura C.2B), enquanto que, para uma relação inibitória entre A e B (figura C.2C), o circuito negativo B/C será funcional se A permanecer ativo (figura C.2D).

A figura C.3 apresenta um exemplo mais complexo envolvendo mais de um regulador. Uma rede composta por um circuito A/B/C e seus respectivos reguladores (figura C.3A) possui um comportamento dinâmico de acordo com as regras lógicas estabelecidas (figura C.3B). Com a mesma lógica aplicada as figuras C.1 e C.2, as restrições de cada componente do circuito e seus reguladores, definidas em termos de sub-redes, são obtidas através da relação entre as funções lógicas dos reguladores e componentes do circuito (figuras C.3C, C.3D e C.3E). A interseção entre essas restrições define o contexto de funcionalidade do circuito A/B/C (figura C.3F).

Figura C.1 – Exemplo de relação de regras lógicas e contexto de funcionalidade de circuitos através do uso do operador AND. (A-D) O grafo regulatório é representado por um circuito negativo entre B e C, regulado por A (sinal de entrada) através da componente B. As regras lógicas são definidas para os nós B e C de acordo com o sinal da interação de A e B. Para uma regulação positiva entre A e B (A), o estado ativo de A é necessário para a funcionalidade do circuito A/B, de acordo com a dinâmica do GTE (B). Por outro lado, uma regulação negativa (C) implica no comportamento inverso de A (inativo) associado à funcionalidade do circuito (D).

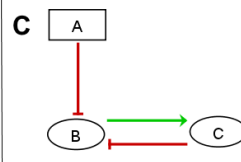


**Regra**  
 A = Sinal de entrada  
 $B = A \text{ AND NOT } C$   
 $C = B$

**Funções lógicas**

$K_B = 1$  se  $(x_A = 1 \text{ AND NOT } x_C = 1)$   
 $K_B = 0$  se não

$K_C = 1$  se  $x_B = 1$   
 $K_C = 0$  se não

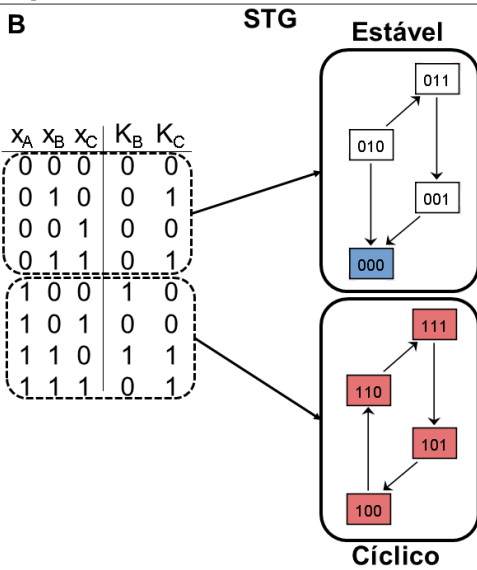


**Regra**  
 A = Sinal de entrada  
 $B = \text{NOT } A \text{ AND NOT } C$   
 $C = B$

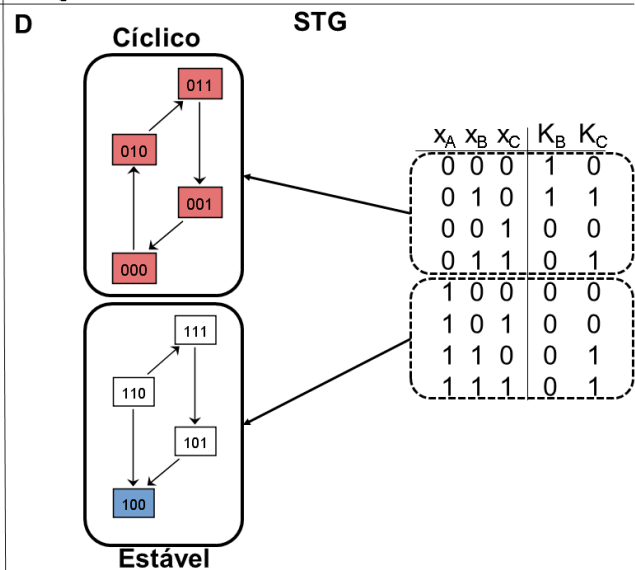
**Funções lógicas**

$K_B = 1$  se  $(\text{NOT } x_A = 1 \text{ AND NOT } x_C = 1)$   
 $K_B = 0$  se não

$K_C = 1$  se  $x_B = 1$   
 $K_C = 0$  se não



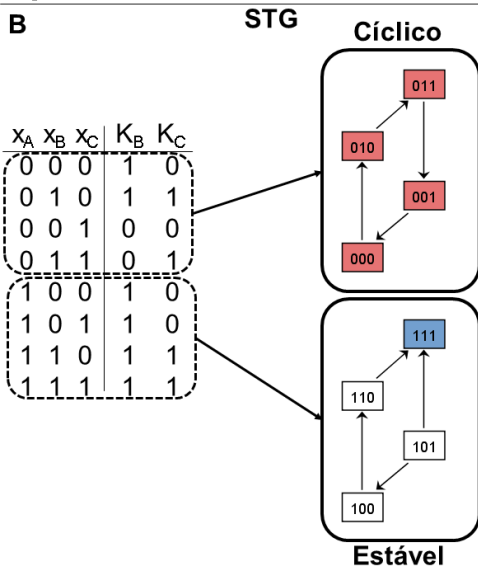
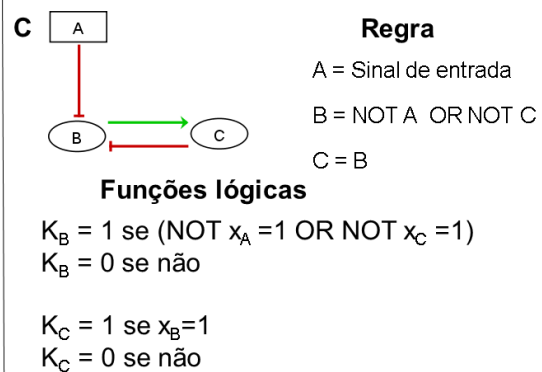
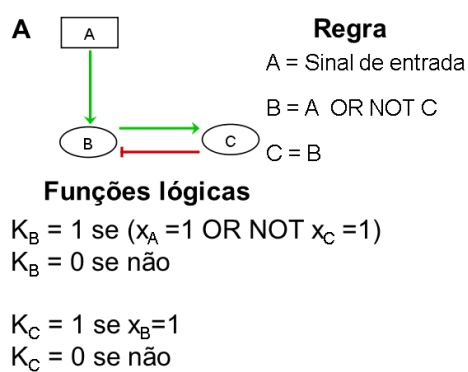
Presença de A é necessária para a funcionalidade do circuito negativo B/C



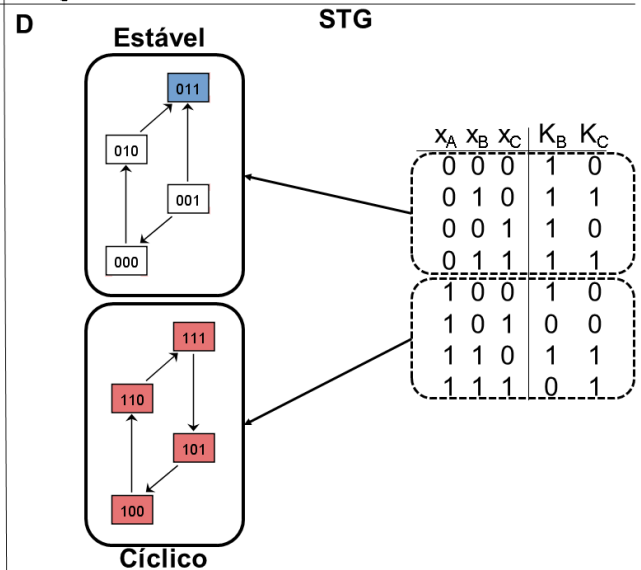
Ausência de A é necessária para a funcionalidade do circuito negativo B/C

Fonte: Próprio autor.

Figura C.2 – Exemplo de relação de regras lógicas e contexto de funcionalidade de circuitos através do uso do operador OR. (A-D) O grafo regulatório é representado por um circuito negativo entre B e C, regulado por A (sinal de entrada) através da componente B. As regras lógicas são definidas para os nós B e C de acordo com o sinal da interação de A e B. Para uma regulação positiva entre A e B (A), o estado inativo de A é necessário para a funcionalidade do circuito A/B, de acordo com a dinâmica do GTE (B). Por outro lado, uma regulação negativa (C) implica no comportamento inverso de A (ativo) associado à funcionalidade do circuito (D).



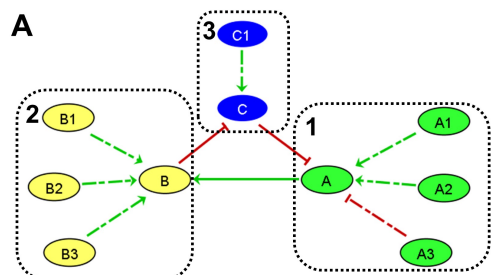
Ausência de A é necessária para a funcionalidade do circuito negativo B/C



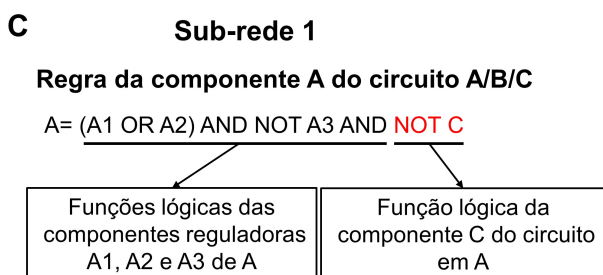
Presença de A é necessária para a funcionalidade do circuito negativo B/C

Fonte: Próprio autor.

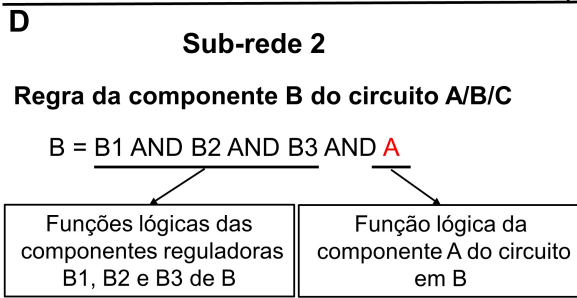
Figura C.3 – Exemplo de relação de regras lógicas e contexto de funcionalidade de circuitos para uma rede complexa. (A) O grafo regulatório é representado por um circuito positivo entre A, B e C, e seus respectivos reguladores A1, A2, A3, B1, B2, B3 e C1 (sinais de entrada). (B) As regras lógicas são definidas para os nós A, B e C. (C-E) Cada componente do circuito e seus respectivos reguladores são subdivididas em sub-redes para facilitar o entendimento. Assim, as restrições de cada componente do circuito são definidas a partir das regras estabelecidas, relacionando a funções lógicas dos reguladores e componentes do circuito (em vermelho) para as sub-redes de A (C), B (D) e C (E). (F) A interseção dessas restrições definem o contexto de funcionalidade do circuito A/B/C.



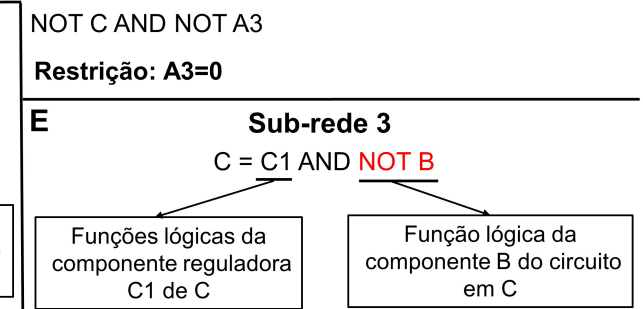
**B Regras**  
 $A = (A1 \text{ OR } A2) \text{ AND NOT } A3 \text{ AND NOT } C$   
 $B = B1 \text{ AND } B2 \text{ AND } B3 \text{ AND } A$   
 $C = C1 \text{ AND NOT } B$



**Relação entre as funções lógicas dos reguladores e componente do circuito**  
 $\text{NOT } C \text{ AND } (A1 \text{ OR } A2)$   
**Restrição: A1=1 e A2=0 ou A1=0 e A2=1**



**Relação entre as funções lógicas dos reguladores e componente do circuito**  
 $A \text{ AND } (B1 \text{ AND } B2 \text{ AND } B3)$   
**Restrição: B1=1 e B2=1 e B3=1**



**Relação entre as funções lógicas dos reguladores e componente do circuito**  
 $\text{NOT } B \text{ AND } C1$   
**Restrição: C1=1**

**F Interseção entre as restrições**  
**Contexto de funcionalidade do circuito A/B/C:**  
 $A2=0, B1=1, B2=1, B3=1, A1=1, C1=1 \text{ e } A3=0$  ou  
 $A2=1, B1=1, B2=1, B3=1, C1=1 \text{ e } A3=0$

Fonte: Próprio autor.

## APÊNDICE D – OUTROS ARTIGOS PUBLICADOS

- Towards DNA-damage induced autophagy: a Boolean model of p53-induced cell fate mechanisms. *DNA Repair*<sup>1</sup> 96, 102971 (2020),  
<http://dx.doi.org/10.1016/j.dnarep.2020.102971>
- p53/E2F1/miR-25 axis regulates apoptosis induction in glioblastoma cells: a qualitative model. *J. Phys. Complex.* 1, 035001 (2020),  
<http://doi.org/10.1088/2632-072X/aba3bb>
- Integrative data modeling from lung and lymphatic cancer predicts functional roles for miR-34a and miR-16 in cell fate regulation. *Scientific Reports*<sup>2</sup> 10, 2511 (2020),  
<http://dx.doi.org/10.1038/s41598-020-59339-y>
- ATM/miR-34a-5p axis regulates a p21-dependent senescence-apoptosis switch in non-small cell lung cancer: a Boolean model of G1/S checkpoint regulation. *FEBS Letters*<sup>3</sup> 594, 227-239 (2019),  
<http://doi.org/10.1002/1873-3468.13615>
- Towards the contribution of the p38MAPK pathway to the dual role of TGF $\beta$  in cancer: A boolean model approach. *Computers in Biology and Medicine*<sup>4</sup> 104, 23524 (2019),  
<http://doi.org/10.1016/j.compbimed.2018.11.025>
- Modeling the role of microRNA-449a in the regulation of the G2/M cell cycle checkpoint in prostate LNCaP cells under ionizing radiation. *PLoS ONE*<sup>5</sup> 13(7): e0200768 (2018),  
<https://doi.org/10.1371/journal.pone.0200768>

---

<sup>1</sup>Fator de impacto (2020): 4.913

<sup>2</sup>Fator de impacto (2020): 4.379

<sup>3</sup>Fator de impacto (2020): 4.124

<sup>4</sup>Fator de impacto (2020): 4.589

<sup>5</sup>Fator de impacto (2020): 3.240