



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
Maestría y Doctorado en Ciencias Bioquímicas

Efecto de las estrategias de control sobre la dinámica de
la red de regulación genética de la transición
epitelio-mesénquima

TESIS

QUE PARA OPTAR POR EL GRADO DE:
Maestro en Ciencias

PRESENTA:
Luis Rodrigo Moreno Morton

Alessio Franci
Facultad de Ciencias

Maria Elena Álvarez-Buylla Roces
Instituto de Ecología

Luis Antonio Mendoza Sierra
Instituto de Investigaciones Biomédicas

Ciudad de México. Noviembre, 2020.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo recibido a través de la Beca Nacional para el estudio de la maestría, otorgada a Luis Rodrigo Moreno Morton con CVU 847460.

Este trabajo fue realizado con el apoyo de los proyectos DGAPA-UNAM PAPIIT IN102420 y CONACyT Ciencia Básica A1-S-10610.

Agradecimientos Personales

A mi tutor principal Alessio Franci, que tuvo la paciencia de explicarme cuando no entendía y animarme cuando me sentía incapaz.

A los miembros de mi comité tutor, Luis Mendoza Sierra y Elena Álvarez-Buylla Roces por sus contribuciones al proyecto.

Al Dr. Pablo Padilla Longoria, por sus contribuciones al trabajo como invitado externo del comité tutor.

A mi madre, padre y hermano, por apoyarme y tenerme paciencia (y por jalarme las orejas cuando me descuidaba con la burocracia).

Al Dr. Rudolf Marinus Buijs, que me recibió en la licenciatura e intentó guiarme en el mundo de la ciencia, y a quien le quedé debiendo un examen de grado.

A los amigos nuevos, los viejos y los de siempre, por apoyarme cuando los necesitaba. Cada uno de ustedes sabe lo que hizo.

Resumen

El control de la regulación molecular de procesos biológicos es muy interesante desde un punto de vista clínico y de ciencia básica, ya que permite comprender de manera más clara cómo éstos podrían ser manipulados. La regulación de la transición epitelio-mesénquima, un proceso importante en el desarrollo, reparación de tejidos, y evolución del cáncer, es de particular interés debido a la importancia clínica del cáncer. En este trabajo, nos propusimos encontrar estrategias de control para la manipulación de la transición epitelio-mesénquima, con el fin de proponer experimentos que lleven al desarrollo de tratamientos.

En el trabajo se hizo una conversión de un modelo booleano previamente publicado a un sistema de ecuaciones diferenciales. Estos brindan la ventaja de que existe una teoría de sistemas, estabilidad y control, además de que permiten realizar un control más fino por tener estados y parámetros continuos. Esto se llevó a cabo simplificando el sistema booleano, y utilizando funciones equivalentes a las reglas booleanas propuestas. Finalmente, se encontró un conjunto de parámetros que reproduce de manera cualitativa el fenómeno biológico y se caracterizaron sus equilibrios. Por medio del teorema de la función implícita, se pudo calcular la sensibilidad de los equilibrios a la variación de los parámetros y proponer un cambio de éstos tal que el tamaño de la cuenca de atracción del equilibrio epitelial incrementara.

Los resultados de este trabajo se pueden utilizar para proponer experimentos *in silico*, *in vitro* o *in vivo*, los cuales podrían ser utilizados para el desarrollo de estrategias de control de la transición epitelio-mesénquima.

Índice general

1. Introducción	9
1.1. Redes de Regulación Genética	9
1.2. Transición Epitelio-Mesénquima	11
1.3. Red de la Transición Epitelio-Mesénquima	12
1.3.1. Propiedades Dinámicas	13
1.3.2. Preguntas inconclusas	14
1.4. Resumen	14
2. Modelos de Estudio	17
2.1. Modelos Booleanos	17
2.2. Interpretación de las Redes Booleanas	21
2.3. Limitaciones de los Modelos Booleanos	22
2.4. Ecuaciones Diferenciales	23
2.5. Enfoque Mixto	24
3. Construcción del Sistema de Estudio	29
3.1. Reducción de dimensiones del Modelo Booleano	29
3.1.1. Simplificación de Reglas Lógicas	30
3.1.2. Eliminación de Genes No Determinantes	30

3.1.3. Contexto de Funcionalidad	31
3.2. Conversión a Ecuaciones Diferenciales	33
3.2.1. Ecuación de Hill	33
3.2.2. Equivalencias entre operaciones	34
3.3. Simplificación del Modelo Continuo	36
3.4. Dinámica	37
4. Análisis de Sensibilidad	43
4.1. Análisis de sensibilidad en el sistema bidimensional	43
4.2. Análisis Tetradimensional	49
4.2.1. Parametrización del Sistema	49
4.2.2. Análisis de Sensibilidad	52
5. Discusión	57
5.1. Resultados	57
5.2. Comparación con otras aproximaciones	59
5.3. Interpretación Biológica	59
5.4. Estrategias de Control	60
A. Reglas de Actualización Booleanas	67
B. Definiciones	73
C. Sistema sin la influencia de <i>Snai2</i>	75
D. Análisis de Sensibilidad	79

Capítulo 1

Introducción

1.1. Redes de Regulación Genética

En 1961 se publicó la estructura del operón de *lac* [1]. Este artículo fue el primero de muchos en el que se muestra la regulación de la expresión de genes por otros genes y metabolitos, y permitió la comprensión de la regulación del metabolismo de lactosa. Es un sistema relativamente sencillo cuya dinámica depende del estado de cuatro elementos (fig. 1.1):

1. El conjunto de genes *lacZYA*.
2. La proteína lacI.
3. La región operadora río arriba de *lacZYA*.
4. La lactosa.

El grupo de genes *lacZYA* se encargan en conjunto del metabolismo de la lactosa, codificando proteínas para su transporte (*lacY*) y catabolismo (*lacZ*, *lacA*). La expresión de estos genes se puede regular por la unión de la proteína represora lacI en el sitio operador, del cual se puede despegar en presencia de lactosa. Esto permite que los genes del metabolismo de la lactosa se sinteticen únicamente en *presencia* de lactosa, evitando una inversión energética en proteínas que no van a ser utilizadas.

El trabajo de Monod y Jacob introdujo al mundo biológico el concepto de las redes de regulación genética (RRGs) [2]: un conjunto de genes que interaccionan entre si, y cuyas interacciones resultan en un fenotipo. Este tipo de enfoque permite adoptar una visión sistémica en la cual se atribuye un fenotipo a la interacción entre varios elementos y que enfatiza a las interacciones, en lugar

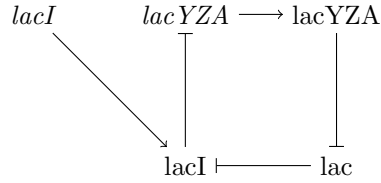


Figura 1.1: Esquema representativo del operón de *lac*. Se muestran los componentes principales del operón y sus interacciones.

de una visión genocéntrica la cual asocia el fenotipo a un conjunto inmutable de genes, que hace énfasis sobre la presencia o ausencia de ellos. Para el caso anterior, por ejemplo, las interacciones son las activaciones e inhibiciones entre genes y metabolitos, y el fenotipo es la capacidad, o falta de ella, de metabolizar lactosa.

Un modelo de RRG tan sencillo como el utilizado para estudiar el operón de *lac* se puede entender de manera casi intuitiva debido a la simplicidad de sus interacciones y su pequeño tamaño. Sin embargo, modelos de este tipo suelen ser únicamente primeras aproximaciones a la comprensión del fenómeno real, e ignoran una gran cantidad de elementos e interacciones que incrementan drásticamente la complejidad del sistema y dificultan su estudio. Un ejemplo de esto es la RRG presentada en la siguiente sección. Además, sistemas aparentemente sencillos pueden tener comportamientos sumamente complejos y contraintuitivos [3, 4]. El incremento en complejidad de los modelos de RRGs trae consigo la necesidad de utilizar herramientas de análisis que permitan entender sus comportamientos. Una de estas herramientas son los *sistemas dinámicos*.

Un sistema dinámico es un formalismo matemático que permite estudiar el cambio del estado de diferentes variables a lo largo del tiempo. Existen diferentes estrategias para la construcción y análisis de los sistemas dinámicos, pero todas brindan la ventaja del uso de herramientas analíticas para el estudio de las RRGs. Esto permite no solo obtener información útil sobre los sistemas, sino que además permite realizar experimentos *in silico* y generar hipótesis que puedan ser corroboradas con experimentos *in vitro* o *in vivo*. Estos a su vez se pueden utilizar como retroalimentación para formar mejores modelos.

En general, el análisis de un sistema dinámico de miles de elementos distintos, como podría ser un organismo completo, es sumamente complicado. En estos casos, se busca simplificar el sistema hasta que sea más sencillo de estudiar, manteniendo ciertas características de interés, dependiendo del enfoque de estudio. Las RRGs usualmente pueden ser descompuestas en partes de menor tamaño con comportamiento relativamente independiente del resto de la red [5, 6]. Cada una de estas partes es llamada *módulo*: un subgrafo cuyos elementos están más conectados entre sí que con elementos fuera de él. Esta desconexión con el resto de la red confiere a los módulos de cierta independencia dinámica [6]. Si se

logra entender la dinámica interna de un módulo, así como su relación entrada-salida con otros elementos de la red, se puede substituir a todos los nodos del módulo por un nodo único que captura sus propiedades dinámicas. Extendiendo esto a la red completa, se puede descomponer la red genética global en una red de módulos de menor escala, en la cual se enfatizan las propiedades dinámicas y funcionales de cada módulo, y sus relaciones entre ellas.

La modularidad en las redes biológicas emerge de manera natural por el proceso evolutivo [5, 6]. El origen de nuevos genes se debe a la duplicación y mutación de un gen «padre». Entonces, el gen «hijo» hereda las interacciones del padre, y las altera al mutar. De esta manera, genes que estaban conectados con el padre siguen estando conectados en medida similar con el hijo. Por otro lado, desde un punto de vista de adquisición o modificación de funciones, la modularidad emerge como consecuencia de una optimización del tiempo de adaptación: ya que un módulo es relativamente independiente del otro, es más fácil adaptar una RRG modular para que adquiera nuevas propiedades ya que los módulos funcionales están aislados de los demás [5, 7]. Si este no fuera el caso, modificar una red para adquirir o alterar una función alteraría el funcionamiento de todas las otras funciones de la red.

Aunque la modularidad funcional y estructural de una RRG no son exactamente la misma [7], la descomposición en módulos nos permite estudiar redes pequeñas asociadas a un fenotipo de interés. Con esto en mente, se han construido redes asociadas a diferentes fenómenos biológicos, como la que se explica a continuación.

1.2. Transición Eitelio-Mesénquima

La transición epitelio-mesénquima (TEM) es la transición de un fenotipo adherido a un tejido celular (epitelio), a un fenotipo libre con capacidades migratorias. Esta transición, y su transición opuesta de mesenquimal a epitelial, son pasos sumamente importantes durante el desarrollo [8], el cierre de heridas y la reparación de tejidos [9]. Adicionalmente, se ha reportado el rol de estas transiciones en la metástasis de células tumorales [10, 11, 12]. Entender los procesos que regulan esta transición, así como las posibles estrategias para manipularla, es interesante tanto desde un punto de vista de desarrollo y biología celular, como desde un punto de vista farmacológico para proponer tratamientos para cáncer.

Las células epiteliales están caracterizadas por una estrecha adhesión intercelular mediada por cadherinas como E-cadherina o claudina, y una polaridad apico-basal. Esto permite que funcionen como una barrera exterior o interior de órganos, ya que impiden el flujo libre de moléculas y dan una dirección al flujo de moléculas reguladas [8, 11]. Por otro lado, las células mesenquimales son

células móviles fusiformes con propiedades semi-troncales [11], con interacciones entre ellas en puntos focales. Las células mesenquimales se pueden marcar por medio de N-cadherina y vimentina [8, 10, 12].

La TEM, siendo un fenómeno que cambia drásticamente el fenotipo celular, involucra una gran cantidad de moléculas que medían la expresión de diferentes genes para el rearrreglo del citoesqueleto, la desintegración de la lámina basal y las uniones intercelulares [8, 11]. Sin embargo, la toma de decisión del fenotipo se puede descomponer en la interacción entre dos grupos de genes: un grupo que refuerza el fenotipo epitelial, compuesto por genes como *ESE2*, *HER2* y *EGF*, y otro que lleva al fenotipo mesenquimal, compuesto por genes como *Snai1* y *Snai2*, *Zeb* y *Twist* [13, 10]. Estas vías de señalización forman un asa de co-inhibición, de tal manera que al prender una, se apaga la otra y hay un cambio en el fenotipo.

De manera similar a lo que sucede con el operón de *lac* mencionado al inicio de este capítulo, estas vías de señalización tienen interacciones con otros fenómenos biológicos que afectan su dinámica [13]. Entre estos, se encuentran módulos funcionales como los de la regulación del ciclo celular, la inflamación, o la senescencia celular. Esto implica el estudio de un sistema mucho más grande y complejo. Sin embargo, éste también se puede descomponer en módulos más pequeños para estudiar su dinámica de manera más completa.

1.3. Red de Regulación Genética de la Transición Epitelio-Mesenquima

El trabajo que se presenta está basado en un modelo de la RRG de la TEM publicado por Méndez-López *et al.* [13]. En éste se estudia un modelo booleano de 41 nodos, separados en seis módulos funcionales:

1. Regulación del fenotipo epitelial
2. Regulación del fenotipo mesenquimal
3. Regulación del fenotipo senescente
4. Regulación de la inflamación
5. Regulación del ciclo celular
6. Silenciamiento epigenético

cada uno con su dinámica local, y con interacción con el resto de los módulos. La relevancia de los módulos epitelial y mesenquimal se discute más arriba. El resto

Cuadro 1.1: *Equilibrios del modelo booleano de la RRG de TEM [13].*

Gen	Epitelial	Mesenquimal	Senescente
<i>Snai2</i>	0	1	0
<i>ESE2</i>	1	0	1
<i>NFκB</i>	1	1	1
<i>p16</i>	0	0	1
<i>p53</i>	0	0	0
<i>Ciclina</i>	1	1	0
<i>TELasas</i>	0	1	1
<i>Rb</i>	0	0	1
<i>E2F</i>	1	0	0

de los módulos se incorpora porque son fenómenos relevantes en la progresión y regulación de la TEM: *e.g.* el silenciamiento epigenético es sumamente importante durante la TEM para silenciar los genes epiteliales, y es desencadenado por los genes de señalización mesenquimal [14]. Utilizando datos de interacción reportados en la literatura, se generaron reglas de actualización booleanas (ver sección 2.1 y apéndice A) en la cual se enlistan las condiciones celulares que permiten la síntesis de cada gen en la red. Ya que una red con 41 nodos es computacionalmente difícil de estudiar aún con un enfoque booleano, la red se simplificó a una de menor tamaño con nueve nodos representativos de cinco de los módulos funcionales de la red extendida (se quita el módulo de silenciamiento epigenético).

El estudio de esta red tiene la particularidad de que permite estudiar el fenómeno de la transición como una interacción entre diferentes módulos funcionales. Otros modelos como los propuestos en [15, 16] estudian la TEM como un fenómeno aislado del resto de los procesos biológicos. Aunque estudios de esta naturaleza son muy útiles para entender el proceso con mucho detalle, un estudio de las interacciones de la transición con otros módulos funcionales nos da una idea más clara sobre su regulación en una escala más general.

1.3.1. Propiedades Dinámicas

De las simulaciones de la red booleana se obtienen tres atractores distintos, cuyos estados corresponden a los perfiles de expresión genética de los fenotipos epitelial, mesenquimal y senescente (cuadro 1.1). Adicionalmente, en el artículo se reportan los atractores obtenidos al hacer «mutaciones» sobre los genes *Snai2*, *ESE2* y *p16*, en las cuales se emula su *knock-in* (KI) o *knock-out* (KO). Los resultados obtenidos con estos experimentos *in silico* corresponden a lo reportado en la bibliografía (datos no mostrados), lo cual da validez al modelo.

Como *proxy* de métrica de estabilidad, se estudiaron las probabilidades de

transición entre equilibrios al introducir ruido aleatorio. En estos experimentos, se observó que si se inicia en el equilibrio epitelial, la mayoría de las simulaciones estocásticas acaba en el fenotipo mesenquimal por una serie de transiciones $epi \rightarrow sen \rightarrow mes$ (fig. 1.2). Esto no solo indica una mayor estabilidad del fenotipo mesenquimal, sino que además concuerda con observaciones reportadas en la literatura, en las cuales se menciona un estado senescente entre los estados epitelial y mesenquimal durante la TEM.

1.3.2. Preguntas inconclusas

El modelo propuesto por Méndez-López *et al.* [13] refleja de manera satisfactoria las observaciones reportadas en la bibliografía. Para su validación, se muestran experimentos *in silico* que emulan en efecto y forma experimentos de *KI* y *KO* previamente reportados. Adicionalmente, se hacen experimentos novedosos para generar nuevas hipótesis sobre la regulación de la TEM. Sin embargo, se dejan de lado preguntas sobre el control de la dinámica del sistema.

La proposición de un modelo dinámico que reproduce las propiedades de un sistema biológico abre preguntas sobre cómo éste puede ser controlado. En el trabajo original se explora esta pregunta por la realización de experimentos de *KI* y *KO*, los cuales se pueden utilizar para manipular al sistema y llevarlo a algún estado deseado. Sin embargo, la aproximación de sistemas booleanos limita mucho las estrategias de control que se pueden utilizar debido a la restricción de estados y la falta de parámetros. Por otro lado, abordar esta pregunta utilizando ecuaciones diferenciales, permitiría no solo realizar manipulaciones suaves sobre el sistema, como serían variaciones pequeñas en la concentración de moléculas o valores de parámetros, sino que además permitiría utilizar métodos de la teoría de sistemas dinámicos clásica, como análisis de estabilidad y sensibilidad, y teoría de bifurcaciones.

1.4. Resumen de Métodos y Conclusiones del trabajo

Para hacer la conversión a un sistema de ecuaciones diferenciales, se hizo una simplificación de las reglas y una reducción de dimensiones utilizando equivalencias del álgebra booleana y métodos propuestos en la bibliografía [17, 18, 19]. En particular, se simplificó las reglas en la medida posible, de tal manera que se eliminasen elementos cuyo estado no afecta resultado de la regla lógica. La eliminación de genes se verificó evaluando su contexto de funcionalidad en la regla booleana en cuestión. Por otro lado, para la eliminación de los genes, se eliminó aquellos que no afectasen a otros, que fueran únicamente intermediarios

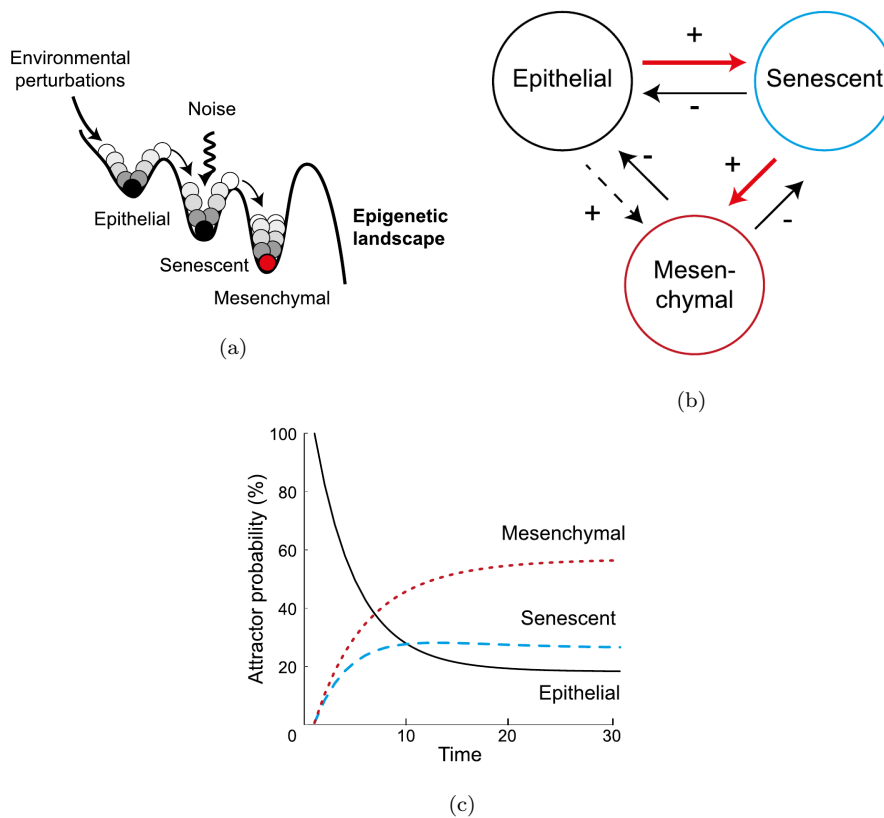


Figura 1.2: Representación de las transiciones del modelo booleano estocástico propuesto en [13]. 1.2a muestra una representación de la trayectoria más probable en el paisaje epigenético. La profundidad de los valles y el tamaño de las colinas entre ellos representan la estabilidad relativa y la energía de transición entre estados, respectivamente. 1.2b muestra los atractores y las posibles transiciones entre ellos (flechas). El signo asociado a cada flecha indica si la tasa neta de transición calculada es positiva o negativa. 1.2c muestra la probabilidad de obtener cada uno de los atractores como función del tiempo, asumiendo que se inició en el atractor epitelial. Tomado de Méndez-López et al., 2017.

entre genes, o que no tuvieran autorregulación, debido a que estos no son clave para la determinación de los puntos de equilibrio.

Una vez obtenida una red booleana de menor dimensión, se hizo una traducción a sistemas de ecuaciones diferenciales por medio de funciones de Hill. Dado que la ecuación de Hill son un modelo genérico de interacción entre biomoléculas, se puede utilizar para describir la dinámica del sistema de manera continua sin necesidad de conocer la interacción exacta entre las moléculas. Al sistema obtenido de esta manera, y a una simplificación adicional a dos dimensiones, se les hizo una caracterización de los equilibrios. A estos, se les realizó un análisis de sensibilidad a la variación de parámetros para encontrar una modificación paramétrica que permitiera incrementar la cuenca de atracción del equilibrio epitelial.

Capítulo 2

Modelos booleanos y continuos para el estudio de redes de regulación genéticas

Existe una gran diversidad de formalismos matemáticos que se pueden utilizar para estudiar sistemas dinámicos. A continuación, se explican los formalismos de mayor relevancia para este trabajo.

2.1. Modelos booleanos como herramienta natural para el estudio de sistemas biológicos

Los métodos más sencillos para el estudio de presencia de RNA y proteínas, como PCR y Western Blotts, arrojan resultados cualitativos. Aunque existen alternativas para obtener resultados cuantitativos utilizando estas metodologías, éstas tienden a ser utilizadas con menor frecuencia debido al incremento en la complejidad y costo de los experimentos. Por estas limitaciones, la bibliografía biomédica cuenta con muchos artículos en los que únicamente se reportan presencia o ausencia de una biomolécula en un contexto celular.

Una pregunta frecuente en el estudio de las RRGs es el efecto de un elemento regulador sobre los demás. Esto se puede estudiar realizando experimentos donde se altera la expresión del regulador de interés para medir el cambio de

otras moléculas río abajo. Debido a que es muy difícil hacer un control exacto sobre la concentración de una molécula en un contexto particular, este tipo de experimentos se lleva a cabo saturando o depletando al sistema del regulador, *i.e.* realizando *KI* o *KO* de los genes. En otras palabras, si quiero conocer el efecto de *genA* sobre *genB*, puedo sobreexpresar o abatir la expresión de *genA* para estudiar su impacto sobre *genB* y ver si funciona como un promotor o inhibidor de su expresión. Conociendo esto con los elementos reguladores de un fenotipo, se puede comenzar a tener una intuición sobre la dirección de cambio dado cierto estado del sistema. Aunque esto se evalúa de manera cualitativa, sigue dando información valiosa sobre su dinámica. Esta información se puede reforzar, dependiendo del contexto, estudiando reacciones o sitios de unión a DNA de proteínas, entre otras cosas.

Medir experimentalmente la presencia o ausencia de una biomolécula en cada instante de tiempo, no solo es de utilidad dudosa, sino experimentalmente inviable. En lugar de esto, los cambios en la expresión se estudian después de un lapso de tiempo para permitir que el sistema llegue a un equilibrio dinámico, el cual es de mayor interés que la dinámica transitoria. Esto implica que, para fines de investigación básica de la regulación, únicamente nos interesan los cambios en tiempos discretos, y se pueden ignorar los cambios en tiempo continuo.

Integrando la información que se tiene de la bibliografía y la que se puede obtener de manera experimental, *i.e.* genes en estados binarios que cambian en tiempos discretos, nos podemos preguntar cómo cambia un conjunto de genes dentro de una RRG. Una buena herramienta para el estudio de sistemas con estas características son los sistemas booleanos como

$$x_i(t+1) = \mathcal{B}_i(x_1(t), \dots, x_i(t), \dots, x_n(t)) \quad (2.1)$$

donde x_i es un gen de la red con n elementos, y $\mathcal{B}_i : [0, 1]^n \rightarrow [0, 1]$ una función booleana.

Una función booleana como la que se muestra en (2.1) toma un vector de valores booleanos (0 ó 1, apagado o encendido, respectivamente), que representan el estado de la red, y regresa un valor booleano que representa el estado siguiente del gen de interés. En el contexto de las RRGs, están compuestas por las operaciones NOT, AND y OR, cuyas tablas de verdad se muestran en 2.1.

Un ejemplo sencillo de una red booleana se puede ver en la red tridimensional

$$\begin{aligned} A(t+1) &= \neg A(t) \vee B(t) \\ B(t+1) &= A(t) \vee (B(t) \wedge C(t)) \\ C(t+1) &= A(t) \wedge C(t) \end{aligned} \quad (2.2)$$

Por simplicidad de notación, se utilizará X' en lugar de $X(t+1)$, y se obviará la dependencia del tiempo en el lado derecho de la ecuación. Aunque es posible que

Cuadro 2.1: Tablas de verdad de las operaciones lógicas NOT, OR y AND, respectivamente.

A	$\neg A$
0	1
1	0

(a)

A	B	$A \vee B$
0	0	0
0	1	1
1	0	1
1	1	1

(b)

A	B	$A \wedge B$
0	0	0
0	1	0
1	0	0
1	1	1

(c)

Cuadro 2.2: Cuadro de actualizaciones del sistema 2.2. Las seis columnas centrales representan el vector de entrada y salida de las funciones lógicas. Las columnas de los extremos son la representación decimal del número binario que se forma con el vector de estados. Estos van desde 0 hasta $2^n - 1$; todos los estados posibles del sistema.

$ID(t)$	$A(t)$	$B(t)$	$C(t)$	$A(t+1)$	$B(t+1)$	$C(t+1)$	$ID(t+1)$
0	0	0	0	1	0	0	4
1	0	0	1	1	0	0	4
2	0	1	0	1	0	0	4
3	0	1	1	1	1	0	6
4	1	0	0	0	1	0	2
5	1	0	1	0	1	1	3
6	1	1	0	1	1	0	6
7	1	1	1	1	1	1	7

las funciones tomen estados de tiempos más atrás en el pasado (e.g. $A(t+1) = A(t) \wedge A(t-1)$), estos casos no se van a ver en este trabajo.

El estado de esta red en algún tiempo arbitrario se puede expresar utilizando un vector que enlista el estado de cada uno de sus elementos en algún orden determinado. Por ejemplo, si formamos un vector de estados $[A, B, C]$ y queremos referirnos al estado en que solo A esté encendido, el vector tomaría la forma $[1, 0, 0]$. Ya que los estados individuales de cada nodo son discretos, la cantidad total de estados de la red es una combinación finita de los estados posibles de cada uno de sus elementos. Conociendo estos y las reglas de actualización, se pueden construir tablas y grafos donde se muestra el estado siguiente para cada estado inicial (cuadro 2.2, fig. 2.1). Cabe recalcar que la cantidad de estados posibles de la red crece como 2^n , donde n es la cantidad de nodos. Nótese además que a cada estado inicial le corresponde un y solo un estado sucesor. Esto se debe a que las funciones de actualización son deterministas.

Del cuadro 2.2 y la figura 2.1 se pueden identificar estados particulares que son de especial interés. El primero de ellos es un estado en el que se cumple la igualdad

$$x(t+1) = \mathcal{B}(x) = x(t).$$

A un estado que cumple con esta igualdad se le conoce como atractor de punto

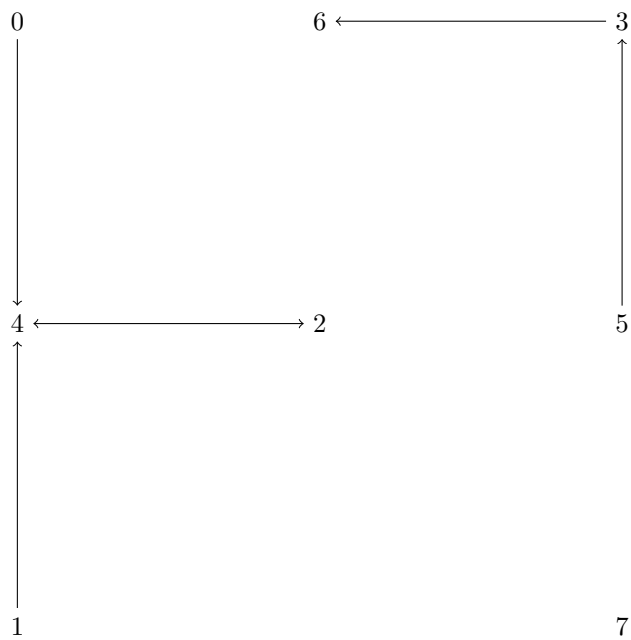


Figura 2.1: *Transiciones entre estados del sistema dinámico booleano 2.2. Cada nodo representa un vector de estados, representado de la misma manera que en el cuadro 2.2. Las flechas representan una transición del estado en el inicio de la flecha al estado en el final. Los estados sin flechas de salida son los equilibrios, y transitan hacia ellos mismos.*

fijo, ya que una vez que el sistema llega a él, permanece en él para todo tiempo futuro. El segundo tipo de estado cumple con la igualdad

$$x(t + \tau) = \mathcal{B}(x).$$

Un estado que cumple con esta igualdad forma parte de un atractor cíclico de periodo τ , ya que el sistema vuelve a pasar por ahí cada τ iteraciones. En el caso particular del sistema 2.2, existen dos atractores de punto fijo (estados 6 y 7) y un atractor cíclico de periodo 2 (estados 2 y 4). El resto de los estados forman parte de las cuencas de atracción del equilibrio al cual conducen. Entonces, el estado 7 tiene una cuenca de tamaño 0, mientras que el estado 6 tiene una cuenca de tamaño 2, compuesta por los estados 5 y 3.

2.2. Interpretación biológica de las redes booleanas

El atractor de una RRG es importante ya que es un estado estable del sistema. Como asumimos que un fenotipo celular es estable dentro de cierto contexto biológico, entonces el atractor de una RRG representa un fenotipo celular [20]. Esta hipótesis ofrece un marco para evaluar si un modelo particular es suficientemente bueno o no: si se obtienen los atractores que corresponden a los fenotipos reportados, el modelo captura las características necesarias del sistema. Por otro lado, si el modelo falla en reproducir los resultados experimentales, ya sea por falta o exceso de atractores, es posible que el modelo se haya construido con alguna hipótesis errónea.

Una vez generado un modelo apropiado, éste se puede utilizar para poner a prueba diferentes hipótesis que pueden resultar difíciles de demostrar experimentalmente, tales como el *KO* de genes letales, o la manipulación simultánea de varios genes. Los resultados de estos experimentos *in silico* deben ser corroborados de manera experimental, ya que provienen de una simplificación de la realidad, pero son valiosos por su cualidad generadora de hipótesis:

- si un experimento demuestra un cambio a un fenotipo deseado, se puede poner a prueba experimentalmente para verificar que las manipulaciones realizadas *in silico* tengan el mismo efecto *in vivo*.
- si un experimento no replica los resultados *in vivo*, entonces el modelo no captura la dinámica lo suficientemente bien. Se pueden proponer cambios al modelo que permitan reproducir los resultados, profundizando el conocimiento del fenómeno.

El modelo booleano propuesto por Mendez-López *et al.* [13] es un modelo que se puede utilizar para estudiar la TEM, ya que permite obtener los atractores

observados experimentales no solo en condiciones silvestres, sino que además reproduce los resultados obtenidos por *KI* o *KO* de genes.

2.3. Limitaciones de los modelos booleanos

Por más que los sistemas booleanos sean una buena primera aproximación de los sistemas biológicos, se debe reconocer que tienen ciertas deficiencias. En primera instancia, las moléculas no están en un «estado encendido» o un «estado apagado», sino que están presentes en concentraciones continuas. Ya que los sistemas booleanos únicamente consideran dos estados, no se puede estudiar el efecto de variaciones pequeñas en la concentración sobre la dinámica global del sistema. Existen, sin embargo, extensiones de los modelos booleanos a sistemas multiestado con actualizaciones lógicas que se pueden utilizar para mitigar esta deficiencia.

Por otro lado, las actualizaciones sincronizadas en tiempo discreto de los sistemas booleanos tienen como hipótesis una escala temporal similar para todos los procesos dentro del sistema modelado. Sin embargo, sabemos que las escalas temporales dentro de un mismo sistema pueden ser muy distintas aún en sistemas sencillos que modelen únicamente un fenómeno; la síntesis y degradación del RNA es mucho más rápida que la síntesis y degradación de una proteína, y a su vez muy distinta a las escalas temporales de una vía de señalización. Existen formas alternativas de actualización de los estados de una red booleana, en las que se eligen nodos al azar dependiendo de una distribución de probabilidad asociada a cada elemento. Estos esquemas pueden ser desde muy sencillos, en los que se eligen los genes completamente al azar, hasta arbitrariamente complicados en los que se forman grupos de los cuales se selecciona un gen para actualizar, o hasta grupos de selección azarosa y otros de actualización sincronizada. Aunque estos esquemas de actualización ayudan a sobrellevar este problema, un modelo previamente simple se puede volver sumamente complejo al considerar escalas temporales distintas.

Finalmente, en los últimos años se han introducido métodos experimentales que permiten hacer estudios cuantitativos de las concentraciones de biomoléculas. Esto permite no solo tener una aproximación de las concentraciones reales de las biomoléculas, sino que además permite estimar parámetros de interacción entre elementos reguladores del sistema. Estos parámetros de regulación, los cuales no son considerados en una aproximación booleana, son muy interesantes ya que permiten la manipulación de la dinámica del sistema sin necesidad de modificar la expresión genética o vida media de las sustancias. Esto permite tener una herramienta adicional de control sobre el fenotipo, ya sea dentro del esquema biológico, o de posibles esquemas terapéuticos o tecnológicos. A diferencia de las críticas anteriores, no existe un modelo booleano con el que se pueda estudiar el efecto causado por la variación de parámetros.

2.4. Ecuaciones Diferenciales para el modelado de Sistemas Biológicos

Un sistema de ecuaciones diferenciales es un conjunto de ecuaciones como el siguiente:

$$\frac{dx}{dt} = F(x, \lambda) \quad (2.3)$$

donde x es el vector de estados en \mathbb{R}^n , λ un vector de parámetros en \mathbb{R}^m y $F(\cdot)$ una función de $\mathbb{R}^n \times \mathbb{R}^m$ a \mathbb{R}^n que representa el cambio instantáneo del estado. Es importante resaltar la diferencia entre los modelos booleanos y continuos, ya que en el primero el resultado de la función representa el estado siguiente, mientras que en los sistemas de ecuaciones diferenciales, el resultado representa el cambio del estado continuo en el tiempo. Este se puede utilizar para obtener el estado siguiente utilizando métodos de integración numérica, pero es importante recordar que las ecuaciones presentadas en ambos casos no son equivalentes.

Contrario al caso de los sistemas booleanos, los elementos del vector de estados pueden tomar cualquier valor real. Esto permite modelar fenómenos más diversos como velocidades, ángulos o concentraciones, en los que un criterio binario es insuficiente. Además, la existencia de parámetros permite modular características cualitativas y cuantitativas de la dinámica del sistema, como se explicará más adelante. Finalmente, el cálculo de la derivada es instantáneo en lugar de en pasos discretos, lo cual permite tener escalas temporales distintas para diferentes variables sin necesidad de utilizar esquemas de actualización complejos.

Un equilibrio de un sistema de ecuaciones diferenciales es cualquier estado en el que la función $F(\cdot)$ de la ecuación 2.3 es igual a cero. A diferencia de los sistemas booleanos, en los que los equilibrios siempre son estables y atractivos, un equilibrio en ecuaciones diferenciales simplemente representa un estado en el que el cambio es cero. Dos formas importantes de clasificar los equilibrios de un sistema de ecuaciones diferenciales son por su estabilidad y su atraktividad. Un equilibrio estable es aquel que, si se empieza en un punto cercano a él, la dinámica va a permanecer cerca de él para todo tiempo futuro. Sin embargo, la estabilidad no es suficiente para hacer un símil con el atractor booleano, ya que las trayectorias no necesariamente se acercan al equilibrio estable. Para que esto suceda, el equilibrio debe además cumplir con el requisito de ser atractivo: para toda trayectoria que empieza cerca del equilibrio, esta no solo se va a mantener cerca, sino que su distancia al punto de equilibrio tiende a cero cuando el tiempo tiende al infinito. Aunque aún dentro de esta clasificación existen subclases, entrar en detalle sobre estas no es relevante para este trabajo. Para ver las definiciones formales de los equilibrios estables, asintóticamente estables y exponencialmente estables, ver el apéndice B.

Como se mencionó anteriormente, una diferencia fundamental entre los sis-

temas booleanos y continuos es la existencia de parámetros de interacción, que pueden modular desde la fuerza interacción entre dos moléculas, hasta la afinidad de una por la otra¹. Estos parámetros, representados por el vector λ en la ecuación 2.3, pueden modificar no solo las propiedades cuantitativas de la dinámica, sino hasta las propiedades cualitativas de los equilibrios como la existencia o estabilidad de los equilibrios. Un ejemplo sencillo de esto se puede ver en el sistema

$$\begin{aligned}\frac{dx}{dt} &= \alpha_1 \frac{y^n}{k_1^n + y^n} - x \\ \frac{dy}{dt} &= \alpha_2 \frac{x^n}{k_2^n + x^n} - y\end{aligned}\tag{2.4}$$

Si se observan las trayectorias del sistema en el plano fase con diferentes valores del parámetro k_1 , se puede ver que la cantidad de equilibrios se reduce al incrementar el valor del parámetro (fig. 2.2). En el primer caso, cuando $k_1 = 0.2$, el sistema tiene dos equilibrios estables en $x = 0$ y $x \sim 1$, separados por un equilibrio inestable cerca del origen. Al incrementar k_1 a 0.4, los equilibrios fuera del origen se acercan, hasta juntarse y cancelarse mutuamente en $k_1 = 0.8$. Esto deja únicamente al equilibrio estable del origen.

Se puede graficar la relación de un parámetro de bifurcación con el valor y estabilidad de un equilibrio (fig. 2.3). En estas gráficas, llamadas *diagramas de bifurcación*, se puede ver cómo cambia la posición de los puntos de equilibrio al variar el valor de algún o algunos parámetros. Un punto interesante de estas gráficas es aquel en el que se gana o pierde uno o más equilibrios, llamado *punto de bifurcación*. En el caso del sistema anterior, el punto de bifurcación se encuentra en $k_1 = 0.5$; para valores menores a 0.5, existen dos equilibrios estables y uno inestable, mientras que para valores mayores existe un solo punto de equilibrio. Estos puntos son de gran interés para el estudio de sistemas dinámicos debido a que son aquellos en los que se da un cambio cualitativo en la dinámica del sistema.

2.5. Enfoque mixto para el estudio de modelos biológicos

Un modelo dinámico que refleja las interacciones precisas entre elementos, así como la dinámica del sistema, es muy difícil de construir. Una aproximación usual es utilizar expresiones de equilibrios químicos por ley de acción de masas [21, 22]. Esta aproximación requiere de un conocimiento extenso no solo del efecto de un gen sobre otro, sino además del orden de reacción de ambas

¹Por la redacción, se puede pensar que me refiero a lo mismo en ambos casos. Debo pensar en otra manera de expresarlo.

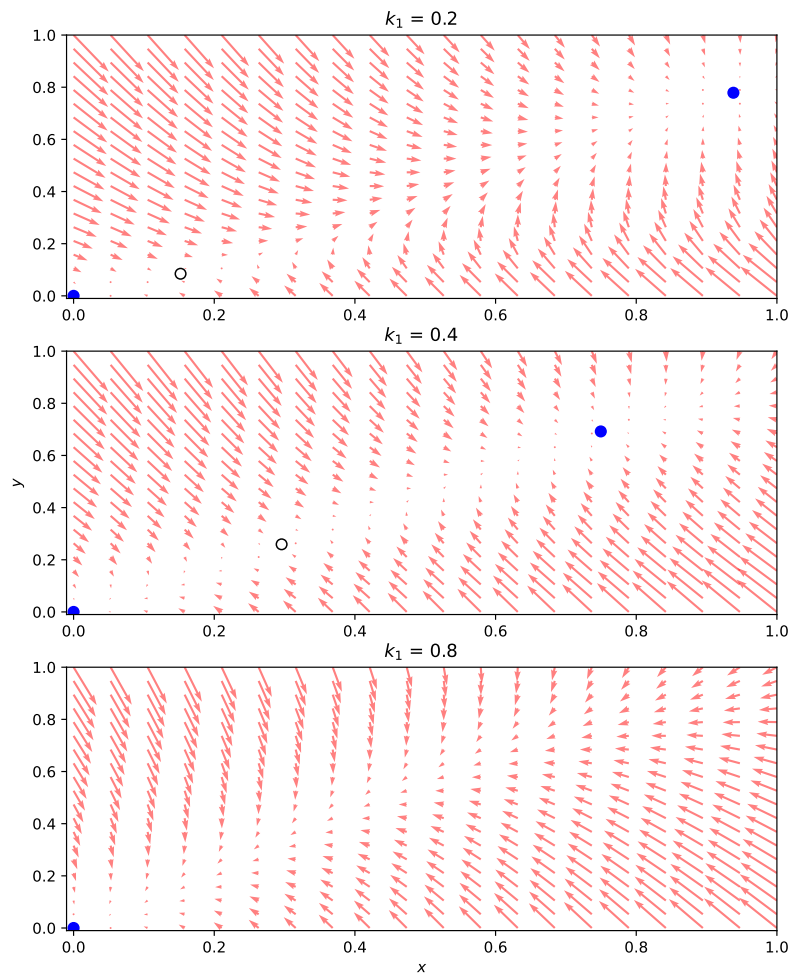


Figura 2.2: Planos fase del sistema 2.4 con tres valores del parámetro k_1 . Los equilibrios estables e inestables de cada plano fase se muestran con puntos azules rellenos y negros vacíos, respectivamente.

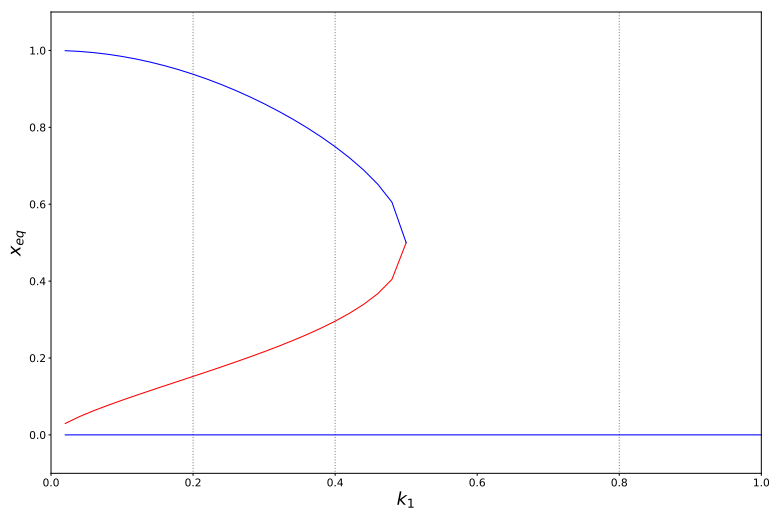


Figura 2.3: Diagrama de bifurcación del punto de equilibrio de la coordenada x del sistema 2.4, con $k_1 \in (0, 1]$. Los equilibrios estables se muestran con una línea azul, mientras que los equilibrios inestables se muestran con una línea roja. Las líneas punteadas marcan los valores de k_1 utilizados para los planos fase de la figura 2.2.

moléculas, usualmente con un conocimiento detallado de afinidades y fuerzas de interacción. Estos parámetros son en muchos casos desconocidos, y en otros virtuales; la interacción modelada no es una interacción real, sino una amalgamación de interacciones que comienzan y acaban en las moléculas de interés. Por otro lado, sistemas puramente cualitativos como los booleanos son relativamente sencillos de construir. Estos modelos capturan propiedades esenciales del sistema, como la naturaleza y dirección de las interacciones, así como la identidad de los atractores.

Una alternativa para la construcción de modelos cuantitativos, es partir de un sistema cualitativo y cuantificarlo [21, 22], utilizando funciones genéricas para la interacción entre elementos, como las funciones de Hill (ver sección 3.2.1) u otras sigmoides. Estas funciones sirven como intermediario entre los sistemas booleanos y mecánicos. En primer lugar, son funciones acotadas entre cero y uno (en el caso de Hill), imitando las propiedades de los sistemas booleanos y la saturabilidad de las interacciones modeladas por ley de acción de masas [22]. En segundo lugar, estas funciones tienen un incremento semi-lineal en concentraciones «intermedias» (fig. 2.4), lo cual emula los umbrales de acción supuestos por los sistemas booleanos y predichos por ley de acción de masas. El uso de estos modelos «híbridos» permite hacer estudios cuantitativos de los procesos modelados, sin necesidad de conocer a detalle las interacciones entre los elementos reguladores. Es importante recordar que estos modelos *no* son modelos mecánicos, y que por ende las interacciones modeladas no reflejan de manera precisa las interacciones reales. Sin embargo, siguen siendo modelos valiosos ya que permiten estudiar el efecto de variaciones pequeñas de parámetros y concentraciones sobre la dinámica del sistema, las cuales no se pueden estudiar utilizando sistemas booleanos.

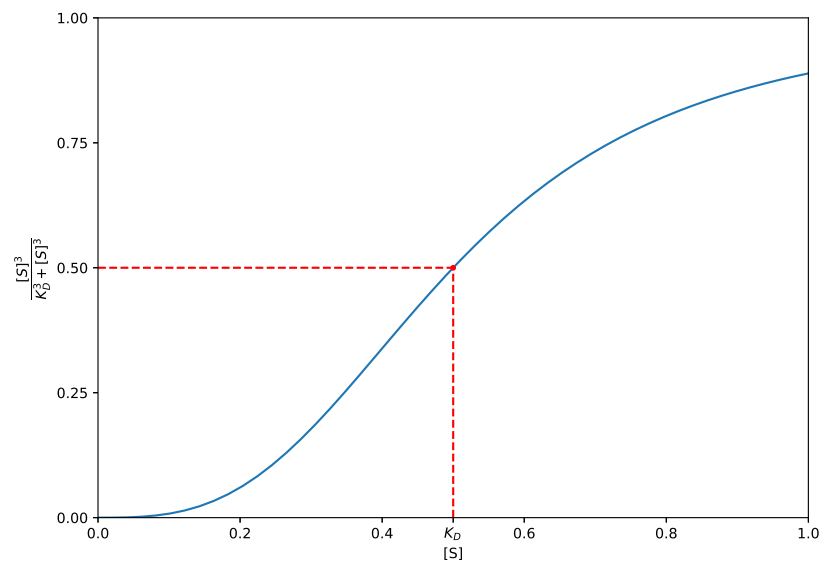


Figura 2.4: Forma de la ecuación de Hill con $S \in [0, 1]$, $K_D = 0.5$ y $n = 3$. Nótese que, por construcción, la función vale la mitad de su valor máximo cuando $S = K_D$.

Capítulo 3

Construcción de un sistema equivalente en ecuaciones diferenciales para la RRG de la transición epitelio-mesénquima

Para estudiar los efectos de la variación de parámetros sobre la dinámica de la transición epitelio-mesénquima, se construyó un sistema de ecuaciones diferenciales que mantuviera la dinámica del sistema booleano. Esto se llevó a cabo en dos partes. En primer lugar, se hizo una reducción de dimensiones del sistema booleano para identificar los genes determinantes de la dinámica del sistema, de tal manera que se tuviera un sistema más fácil de estudiar pero que brindara la misma información que el sistema original. Finalmente, se hizo una traducción por medio de funciones de Hill para trabajar con concentraciones continuas, usando funciones equivalentes a la lógica booleana. A continuación, se describen ambos pasos con mayor detalle.

3.1. Reducción de dimensiones del Modelo Booleano

En la sección 3.1.1 se describe cómo se llegó a un conjunto de reglas lógicas a partir de las reglas de actualización publicadas. Más adelante, en las seccio-

Cuadro 3.1: Comparación de los valores de la función lógica mostrada en (3.1) y (3.2). Las primeras dos columnas de la izquierda representan todas las combinaciones de valores de A y B posibles. Las dos columnas de la derecha muestran el valor de la función (3.1) y (3.2), respectivamente.

A	B	$(A \wedge B) \vee (A \wedge \neg B)$	A
0	0	0	0
0	1	0	0
1	0	1	1
1	1	1	1

nes 3.1.2 y 3.1.3 se describe cómo se hizo la reducción de dimensiones, y cómo se validó la misma.

3.1.1. Simplificación de Reglas Lógicas

Las reglas lógicas publicadas en [13] se construyeron enlistando condiciones celulares que permiten la expresión del gen de interés. Su estructura es una similar a la que se muestra en la ecuación (3.1).

$$(A \wedge B \wedge C) \vee (A \wedge B \wedge \neg C) \vee \dots \vee (\neg A \wedge \neg B \wedge C) \vee (\neg A \wedge \neg B \wedge \neg C) \quad (3.1)$$

la cual se puede simplificar utilizando la equivalencia

$$(A \wedge B) \vee (A \wedge \neg B) \equiv A, \quad (3.2)$$

demostrada en el cuadro 3.1. Se utilizó esta equivalencia para simplificar las reglas de actualización publicadas en [13], seguida de una simplificación final realizada utilizando el comando `simplify [regla]` de los servicios de Wolfram Alpha, para encontrar una función aún más sencilla utilizando otras equivalencias del álgebra booleana (ver apéndice A para ver las reglas originales y simplificadas del sistema). Se corroboró que la simplificación obtenida de esta manera se haya realizado de manera correcta comparando los atractores de la red simplificada con los atractores originales (datos no mostrados).

3.1.2. Eliminación de Genes No Determinantes

Para la eliminación de genes con influencia menor sobre la dinámica, se siguieron tres metodologías:

1. Eliminación de nodos sin salidas: ya que estos nodos no tienen efecto sobre el estado de otros nodos de la red, su estado puede verse únicamente como una salida del sistema e ignorarse en la determinación de estados consecutivos.

2. Eliminación de nodos intermediarios [18, 17, 19]: si un nodo B sirve únicamente como un relevo de información desde el nodo A hasta el nodo C , puede ser eliminado. Su regla de actualización se substituye en la dependencia de B de la regla de C .
3. Eliminación de nodos sin autoregulación [19]: debido a las características estabilizadoras de la autoregulación, un gen cuya regla de actualización no incluye a si mismo puede ser eliminado sin alterar los atractores de la red. De igual forma que en el caso anterior, la regla de actualización del nodo eliminado se substituye en las reglas de actualización de los nodos a los que afecta.

Estas reglas llevaron a la eliminación de los nodos $p53$, $Ciclina$, Rb y $E2F$ por falta de autoregulación, y $TELas$ por ser un nodo intermediario entre $ESE2$, $p16$ y $p53$ (fig. 3.1).

3.1.3. Contexto de Funcionalidad

La simplificación de una regla lógica booleana no tiene un resultado único. Debido a esto, y que la eliminación de nodos depende en parte del resultado de dichas simplificaciones, se buscó corroborar que diferentes vías de simplificación no resultaran en redes simplificadas cualitativamente distintas. Para identificar elementos cuyo estado no afecta el valor de una regla booleana, se puede evaluar el contexto de funcionalidad de cada gen para cada función. Esto indica qué genes se pueden eliminar de las reglas [23, 24]. Buscamos que las simplificaciones eliminen únicamente genes no funcionales.

En una red booleana, el contexto de funcionalidad se estudia formando parejas de vectores de estado idénticos excepto por el gen cuya funcionalidad se quiere estudiar, y comparando el valor de las funciones lógicas dentro de cada pareja. Si dentro de cada pareja la función toma el mismo valor, entonces la variación del elemento no afecta al valor de la función, y puede ser eliminado. Entonces, si se busca estudiar el contexto de funcionalidad del gen C en una función lógica como

$$\mathcal{B}(A, B, C) = ((A \wedge B) \vee (A \wedge C)) \vee ((A \wedge B) \vee (A \wedge \neg C))$$

con vector de estado $[A, B, C]$, se formarán las parejas

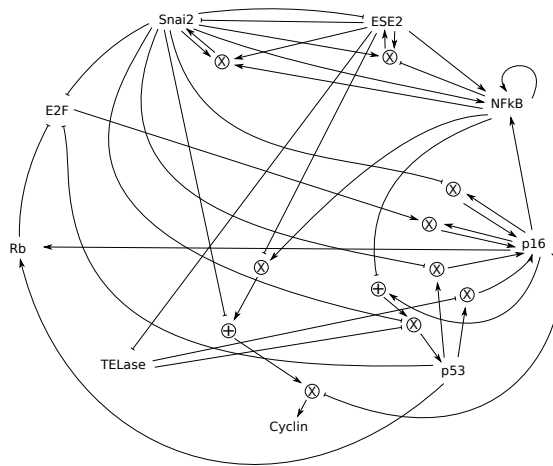
$$\begin{array}{ll} \text{Pareja1} & [0, 0, \mathbf{0}], \quad [0, 0, \mathbf{1}] \\ \text{Pareja2} & [0, 1, \mathbf{0}], \quad [0, 1, \mathbf{1}] \\ \text{Pareja3} & [1, 0, \mathbf{0}], \quad [1, 0, \mathbf{1}] \\ \text{Pareja4} & [1, 1, \mathbf{0}], \quad [1, 1, \mathbf{1}] \end{array}$$

Al comparar los valores de la función lógica en esas parejas, podemos observar que el valor de C no afecta el resultado de la función, ya que las parejas 1 y 2

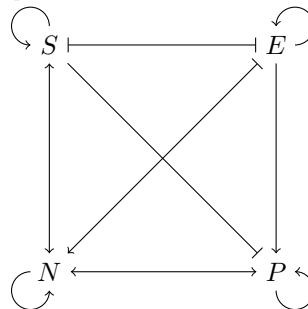
dan resultado 0, y las parejas 3 y 4 dan resultado 1. Al repetir este proceso con el gen A , notamos desde la pareja 1 $([0, 0, 0], [1, 0, 0])$ que el gen sí es funcional.

Utilizando este criterio, se verificó que las simplificaciones de las reglas booleanas eliminaran los mismos elementos, y por ende resultaran en reducciones equivalentes del sistema.

Con los criterios mencionados anteriormente, se llegó al sistema booleano de



(a) Representación gráfica de la red booleana completa. Los símbolos \oplus y \otimes representan las funciones lógicas OR y AND, respectivamente.



(b) Representación gráfica de la red booleana reducida utilizando los métodos mencionados arriba.

Figura 3.1

Cuadro 3.2: *Equilibrios del sistema (3.3) y sus equivalencias con el sistema original. El orden de los genes en el vector de estados del sistema original es [Snai2, ESE2, NFκB, p16, p53, Ciclina, TELasa, Rb, E2F]. El vector de estados en el sistema reducido es [Snai2, ESE2, NFκB, p16].*

Atractor	Sistema Original	Sistema Reducido
Mesenquimal	[1, 0, 1, 0, 0, 1, 1, 0, 0]	[1, 0, 1, 0]
Epitelial	[0, 1, 1, 0, 0, 1, 0, 0, 1]	[0, 1, 1, 0]
Senescente	[0, 1, 1, 1, 1, 0, 0, 1, 0]	[0, 1, 1, 1]

cuatro dimensiones que se muestra en la ecuación (3.3).

$$\begin{aligned}
 S(t+1) &= \neg E \vee (S \wedge E \wedge N) \\
 E(t+1) &= \neg S \vee (S \wedge E \wedge \neg N) \\
 N(t+1) &= S \vee E \vee N \vee P \\
 P(t+1) &= \neg S \wedge ((E \wedge N) \vee P)
 \end{aligned} \tag{3.3}$$

donde S , E , N y P representan los genes *Snai2*, *ESE2*, *NFκB* y *p16*, respectivamente.

Los equilibrios de este sistema y su equivalencia en el sistema original se muestran en el cuadro 3.2.

3.2. Conversión a un sistema equivalente en ecuaciones diferenciales

Habiendo obtenido un sistema booleano equivalente de dimensiones reducidas, se buscó una conversión a un sistema de ecuaciones diferenciales en tiempo continuo. Esto requirió encontrar una forma de expresar los estados booleanos 0 y 1 como un continuo para poder trabajar con un *proxy* de concentración, además de una manera de traducir las reglas lógicas a funciones de \mathbb{R}^n en \mathbb{R} .

3.2.1. Ecuación de Hill

La ecuación de Hill es un modelo general de interacción entre dos biomoléculas. La ecuación tiene la forma

$$\Theta = \frac{L^n}{K_D^n + L^n} \tag{3.4}$$

donde Θ representa la fracción de sitios de unión a ligando ocupados, expresada como función de la concentración de ligando L , una constante de afinidad K_D . La

ecuación de Hill se comporta de manera idéntica a la de Michaelis-Menten, con mínimo en 0, máximo en 1 y punto medio en $L = K_D$. Sin embargo, la ecuación de Hill se puede ver como una generalización de la ecuación de Michaelis-Menten, ya que incluye un factor de cooperatividad n que cambia la pendiente de la sigmoide. Este factor representa el cambio de afinidad proteína-ligando por la asociación o disociación de otra proteína en un complejo. Si $n > 1$, se dice que hay cooperatividad positiva en la interacción, y la afinidad incrementa; cuando se une un ligando en un sitio del complejo, se favorece la unión en otro. Por otro lado, si $n < 1$, la cooperatividad es negativa, y la unión en un sitio previene la unión en otro. En caso de que $n = 1$, la unión de cada ligando es un evento independiente, y la interacción se puede modelar según la ecuación de Michaelis-Menten.

Se decidió utilizar la ecuación de Hill ya que permite hacer una aproximación de las interacciones sin necesidad de conocer a detalle la dinámica de interacción entre los elementos de la red [18]. Dado que las interacciones consideradas en el sistema no necesariamente son entre las proteínas en si, sino entre vías de señalización asociadas a ellas, el uso de las ecuaciones de Hill permite obtener información sobre un sistema con relaciones desconocidas o difíciles de conocer.

3.2.2. Equivalencias entre operaciones

Existen diferentes estrategias para hacer traducciones de sistemas booleanos a ecuaciones diferenciales [25, 26]. En muchos casos, éstas buscan mantener los estados acotados entre 0 y 1, o requieren de ciertas condiciones sobre las reglas lógicas. En particular, el método propuesto en [25] requiere que las reglas lógicas sean de la forma

$$X(t+1) = (A_1 \vee \dots \vee A_i) \wedge \neg(I_1 \vee \dots \vee I_j)$$

donde A e I representan conjuntos de genes activadores e inhibidores, respectivamente. Esto implica que cada uno de los genes activadores es capaz de activar a X sin necesidad de otro activador, siempre y cuando no haya *ningún* gen inhibidor. Ya que nuestras reglas lógicas no cumplen con esta condición, y que no nos interesa mantener la concentración acotada en $[0, 1]$, se tuvo que llegar a una lógica de traducción que satisfaga las propiedades mostradas en el cuadro 2.1. Para el caso de la negación, lo único que se necesita es que la función tome el valor contrario al argumento (cuadro 2.1a): 1 cuando es 0, y 0 cuando es 1. Ya que la ecuación de Hill tiene un máximo en 1, esto se puede lograr haciendo la equivalencia

$$\neg A \equiv 1 - H(A) = \frac{K_d^n}{K_d^n + A^n}$$

donde $H(A)$ representa la ecuación de Hill con argumento A , con sus ctes. de afinidad y cooperatividad correspondientes.

Para la función **OR**, se necesita que la función sea 0 únicamente cuando ambos argumentos son cero (cuadro 2.1b). Esto se logra utilizando una suma, ya que ésta es distinta de cero si al menos uno de los argumentos lo es. Por tal motivo, se hizo la traducción utilizando la equivalencia

$$A \vee B \equiv H(A) + H(B).$$

Para el caso de la función **AND**, cuyo valor es 0 con que alguno de los argumentos sea cero (cuadro 2.1c), se utilizó el producto, que cumple con las mismas propiedades. Esto lleva a la equivalencia

$$A \wedge B \equiv H(A) \times H(B).$$

Utilizando estas reglas para convertir las operaciones booleanas en funciones continuas, se puede transformar un sistema como el ejemplo en la ecuación (2.2) en el sistema de ecuaciones diferenciales

$$\begin{aligned} \frac{dA}{dt} &= \alpha_1 \frac{k_1^n}{k_1^n + A^n} + \alpha_2 \frac{B^n}{k_2^n + B^n} + \beta_A - A \\ \frac{dB}{dt} &= \alpha_3 \frac{A^n}{k_3^n + A^n} + \alpha_4 \left[\frac{B^n}{k_4^n + B^n} \times \frac{C^n}{k_5^n + C^n} \right] + \beta_B - B \\ \frac{dC}{dt} &= \alpha_5 \frac{A^n}{k_6^n + A^n} \times \frac{C^n}{k_7^n + C^n} + \beta_C - C \end{aligned}$$

Aquí, además de ver una traducción directa de las funciones booleanas en las continuas, se agrega un factor de degradación lineal a cada una de las ecuaciones para preservar los equilibrios booleanos en el sistema continuo. Adicionalmente, se considera una síntesis basal β_i para cada una de las biomoléculas. El incorporar una tasa lineal de degradación impide que las concentraciones crezcan de manera infinita. Ya que las ecuaciones de Hill son saturables, independientemente del valor de K_D habrá un valor de concentración más alto para el cual la ecuación de Hill tome el valor de saturación. Si la concentración es mayor a ese valor, entonces la degradación supera a la síntesis y la concentración decrece.

Aplicando estas reglas de conversión a las reglas lógicas presentadas en la ecuación (3.3), se llegó al sistema dinámico de ecuaciones diferenciales que se presenta en la ecuación (3.5).

$$\begin{aligned} \frac{dS}{dt} &= \alpha_1 \frac{k_1^n}{E^n + k_1^n} + \alpha_2 \frac{S^n}{S^n + k_2^n} \times \frac{E^n}{E^n + k_3^n} \times \frac{N^n}{N^n + k_4^n} + \beta_S - S \\ \frac{dE}{dt} &= \alpha_3 \frac{k_5^n}{S^n + k_5^n} + \alpha_4 \frac{E^n}{E^n + k_6^n} \times \frac{S^n}{S^n + k_7^n} \times \frac{k_8^n}{N^n + k_8^n} + \beta_E - E \\ \frac{dN}{dt} &= \alpha_5 \frac{S^n}{S^n + k_9^n} + \alpha_6 \frac{E^n}{E^n + k_{10}^n} + \alpha_7 \frac{N^n}{N^n + k_{11}^n} + \alpha_8 \frac{P^n}{P^n + k_{12}^n} + \beta_N - N \\ \frac{dP}{dt} &= \alpha_9 \frac{k_{13}^n}{S^n + k_{13}^n} \left[\left(\alpha_{10} \frac{E^n}{E^n + k_{14}^n} \times \frac{N^n}{N^n + k_{15}^n} \right) + \alpha_{11} \frac{P^n}{P^n + k_{16}^n} \right] + \beta_P - P. \end{aligned} \tag{3.5}$$

Nótese que cada función de Hill tiene parámetros k_i distintos, y que se adicionan coeficientes de amplificación α_i .

3.3. Reducción de dimensiones del sistema en tiempo continuo

Un sistema como el mostrado en la ecuación (3.5) es sumamente difícil de estudiar. En primer lugar, debido a las no-linealidades, la identificación de los puntos de equilibrio se debe realizar, en la gran mayoría de los casos, utilizando métodos numéricos. Estos métodos están limitados a encontrar únicamente el equilibrio más cercano a una condición inicial dada. Esto implica realizar una búsqueda exhaustiva de condiciones iniciales, o un conocimiento previo de la dinámica del sistema para poder encontrar todos los equilibrios. Por otro lado, debido a que el sistema es de dimensiones altas, no se puede utilizar métodos gráficos para asistir en la búsqueda de los equilibrios.

Ya que en el trabajo se buscó evitar un análisis de fuerza bruta para la identificación de parámetros, se decidió hacer una reducción de dimensiones al sistema en ecuaciones diferenciales para facilitar su estudio. Esto nos permite hacer un análisis metodológicamente más sencillo, cuyos resultados sirven como punto de partida para el análisis del sistema de mayores dimensiones. Como primer paso de la simplificación, se eliminó la ecuación de *NFκB* debido a que su estado es el mismo en todos los equilibrios del sistema booleano (cuadro 3.2). Por ende, el estado de *NFκB* no parece ser crítico en las transiciones entre equilibrios. Ya que se encuentra en un producto en todas las ecuaciones excepto la propia, la variación en su concentración se puede ver como un cambio en los valores de α_i del resto de las ecuaciones.

El siguiente paso es considerar un sistema tridimensional con asas de interacción moduladas por los otros elementos de la red (figura 3.2). Ya que se eliminó *NFκB*, su influencia sobre los otros genes se puede entender como una influencia indirecta de sus genes de entrada sobre sus genes de salida. Por ejemplo, como *p16* activa *NFκB*, y éste activa *Snai2*, esa interacción se puede ver como una activación de *Snai2* por *p16*. De manera análoga, se puede decir que *p16* inhibe a *ESE2* y se autoactiva.

Consideramos que, dada la estructura de las ecuaciones, las inhibiciones mutuas entre los genes *Snai2* y *ESE2* son más fuertes que sus autoactivaciones. Entonces, el ciclo de inhibiciones determina más fuertemente la dinámica del sistema. Si se ignora la dinámica de *Snai2*, se puede considerar la inhibición de *ESE2* por *Snai2* como una activación de *ESE2* sobre si mismo, modulada por *p16*. De manera similar, la modulación negativa de la activación de *p16* por *ESE2*, se puede ver como una activación por *ESE2* modulada positivamente por

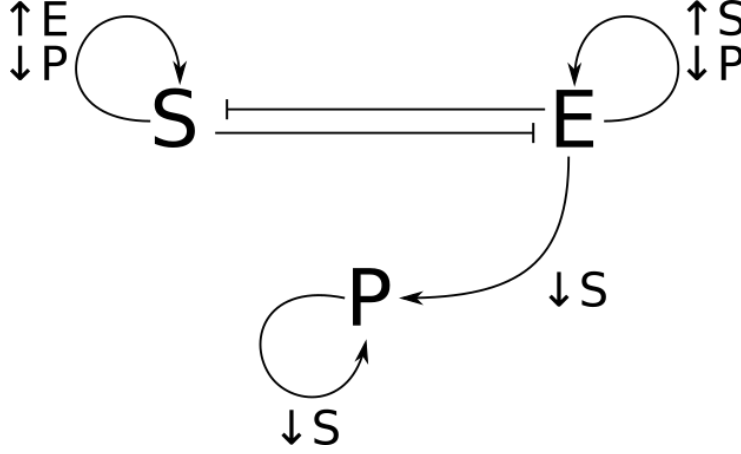


Figura 3.2: Esquema de las interacciones entre *Snai2*, *ESE2* y *p16*. De manera análoga a lo presentado en la ecuación (3.5), *S* representa a *Snai2*, *E* a *ESE2* y *P* a *p16*. Flechas que terminan en \rightarrow representan activación, mientras que las flechas que terminan en \dashv representan inhibición. Las etiquetas junto a las interacciones representan la modulación de la interacción por otros elementos regulatorios. Debido a los efectos cruzados entre elementos regulatorios, podemos considerar, por ejemplo, la inhibición de *p16* por *Snai2* como una activación indirecta de *p16* por *ESE2*.

él mismo. Finalmente, utilizando la misma lógica, se puede ver la modulación negativa de *Snai2* sobre la autoactivación de *p16* como una modulación positiva por parte de *ESE2*. Para seguir considerando de manera más explícita el rol de *Snai2*, se incorporó un factor m que incrementa o reduce la constante de cooperatividad de las ecuaciones dependiendo del rol activador o inhibidor de *Snai2*. Con esto, se llega al modelo bidimensional que se muestra a continuación:

$$\frac{dE}{dt} = \alpha_1 \frac{E^{n/m}}{c_1^{n/m} + E^{n/m}} \times \frac{z_1^n}{z_1^n + P^n} + \beta_E - E \quad (3.6a)$$

$$\frac{dP}{dt} = \alpha_2 \frac{E^{nm}}{z_2^{nm} + E^{nm}} + \alpha_3 \frac{E^{nm}}{z_2^{nm} + E^{nm}} \frac{P^n}{c_2^n + P^n} + \beta_P - P \quad (3.6b)$$

De manera similar a casos anteriores, se usan E y P como referencia a los genes *ESE2* y *p16*, respectivamente.

3.4. Triestabilidad del sistema dinámico y equivalencia con el sistema booleano

La reducción del sistema a dos dimensiones permite utilizar métodos gráficos para identificar los equilibrios. Para esto, se utilizó el método de las ceroclinas.

Una ceroclina es una curva en el espacio de estados en la cual una de las derivadas es cero. Por ejemplo, la ceroclina de la ecuación (3.6a) representa todos los valores de E y P para los cuales $\frac{dE}{dt} = 0$. De igual manera, la ceroclina de la ecuación (3.6b) representa a todos los valores de E y P para los cuales $\frac{dP}{dt} = 0$. Las intersecciones de las ceroclinas dan los puntos de equilibrio del sistema. Utilizando este método, se buscó un conjunto de parámetros que permitiera la existencia de tres equilibrios estables cualitativamente equivalentes a los obtenidos en el sistema booleano. Entonces, se graficaron las ceroclinas para diferentes valores de parámetros utilizando la función `fimplicit` del lenguaje de programación Matlab (figs. 3.3, 3.4). Dada la forma de las ceroclinas, es evidente que sí existe un subespacio en el conjunto de parámetros tal que los equilibrios del sistema continuo se parecen a los equilibrios del sistema booleano (fig. 3.5, cuadro 3.3). Adicionalmente, utilizamos esta misma estrategia para demostrar que una sobresimplificación del sistema, presentada en el apéndice C, no es capaz de ser triestable. Consideramos que estas diferencias son debidas a que la sobresimplificación no incorpora la regulación de *Snai2* sobre *ESE2* y *p16*, la cual se considera en la ecuación (3.6) con el parámetro m y la modulación de la autorregulación de *p16*.

El análisis de ceroclinas únicamente indica la posición de los puntos de equilibrio. Aunque su estabilidad puede ser inferida por la forma del plano fase, se requiere de una caracterización más formal para determinar la (in)estabilidad de los puntos de equilibrio. Para esto se hizo uso del primer teorema de Lyapunov (ver apéndice B), el cual dice que un punto de equilibrio de un sistema no lineal es exponencialmente estable si y solo si lo es también en la linealización del sistema al rededor de ese punto. Para verificar esto, se buscaron los puntos de equilibrio utilizando métodos numéricos del módulo de cómputo científico SciPy del lenguaje de programación Python. Como punto de partida para la identificación de los equilibrios, se utilizaron las intersecciones de las ceroclinas mostradas en la figura 3.5. Una vez identificados los puntos de equilibrio, se calculó el Jacobiano del sistema (3.6) de manera simbólica utilizando SymPy, y se evaluó en cada uno de los equilibrios. Finalmente, se calcularon los valores propios de cada uno de los Jacobianos utilizando la paquetería NumPy.

Se puede ver que los tres puntos de equilibrio equivalentes a los del sistema booleano, marcados como *epi*, *sen* y *mes* en el cuadro 3.4 y la figura 3.5, corresponden a los fenotipos epitelial, senescente y mesenquimal, respectivamente, y que además son estables (ambos valores propios son negativos). Adicionalmente, se obtienen dos puntos de equilibrio inestables (al menos un valor propio positivo) que no aparecen en el sistema booleano. Esto es natural, ya que los sistemas booleanos únicamente presentan equilibrios estables. Debido a su posición relativa a los atractores, se llamó a estos s_{SE} y s_{ME} , ya que separan las cuencas de atracción de los equilibrios *sen* y *epi*, y *mes* y *epi*, respectivamente (fig. 3.5). Además, se hicieron simulaciones del sistema empezando en diferentes puntos del plano fase, para demostrar gráficamente la (in)estabilidad de los puntos de equilibrio. Finalmente, se hicieron simulaciones hacia atrás en el tiempo comen-

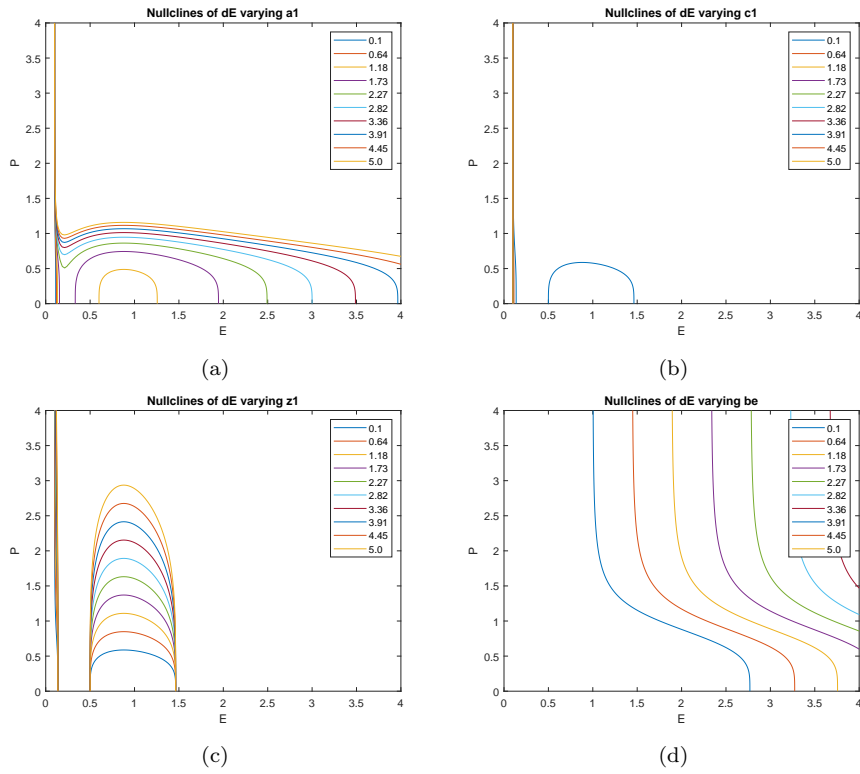


Figura 3.3: Representación de las ceroclinas de dE al variar los parámetros. Variación sobre los parámetros α_1 (fig. 3.3a), c_1 (fig. 3.3b), z_1 (fig. 3.3c) y β_E (fig. 3.3d) de la ecuación (3.6a).

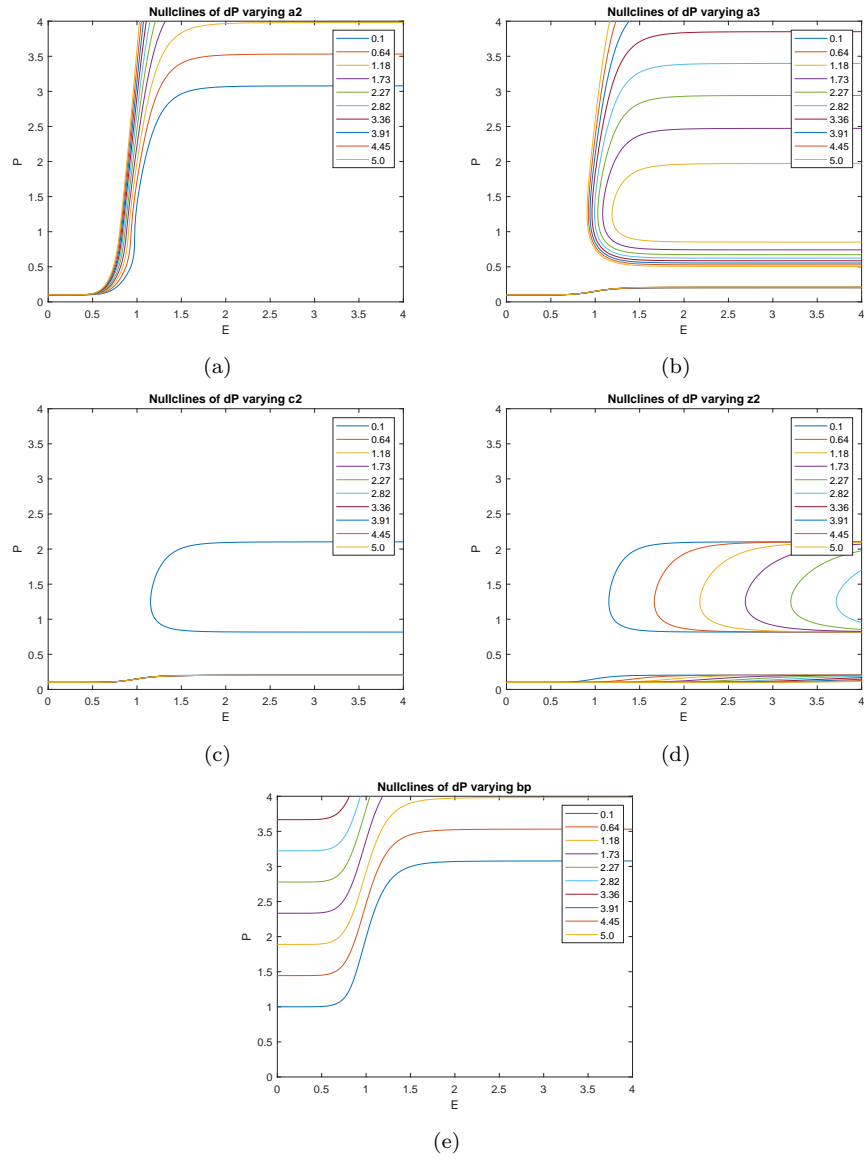


Figura 3.4: Representación de las ceroclinas de dP al variar los parámetros. Variación sobre los parámetros α_2 (fig. 3.4a), α_3 (fig. 3.4b), c_2 (fig. 3.4c), z_2 (fig. 3.4d) y β_P (fig. 3.4e) de la ecuación (3.6b).

Cuadro 3.3: *Parámetros utilizados para la obtención de tres puntos de equilibrio equivalentes a los booleanos en el sistema continuo (3.6).*

Parámetro	Valor
z_1	5
z_2	1
c_1	1
c_2	1
α_1	2
α_2	0.1
α_3	2
β_E	0.1
β_P	0.1
n	4
m	2

Cuadro 3.4: *Equilibrios del sistema dinámico (3.6) utilizando los parámetros mostrados en el cuadro 3.3.*

ID	(E, P)	Equivalente Booleano	Valores Propios	Fenotipo
mes	$\begin{bmatrix} 0.14 \\ 0.10 \end{bmatrix}$	$[0, 0]$	$\{-0.47, -1\}$	Mesenquimal
s_{ME}	$\begin{bmatrix} 0.50 \\ 0.1 \end{bmatrix}$	—	$\{\mathbf{0.28}, -1\}$	—
epi	$\begin{bmatrix} 1.46 \\ 0.19 \end{bmatrix}$	$[1, 0]$	$\{-0.41, -0.94\}$	Epitelial
s_{SE}	$\begin{bmatrix} 1.46 \\ 0.85 \end{bmatrix}$	—	$\{-0.40, \mathbf{1.02}\}$	—
sen	$\begin{bmatrix} 1.39 \\ 1.93 \end{bmatrix}$	$[1, 1]$	$\{-0.28, -0.85\}$	Senescente

zando desde los equilibrios inestables, para estimar la forma de su *separatriz*: subespacio $(n - 1)$ -dimensional para el cual el equilibrio inestable *es* estable. Este subespacio sirve para separar las cuencas de atracción de los equilibrios estables, y está anclado a la posición del equilibrio inestable.

Los resultados presentados en esta sección nos indican que el modelo continuo generado a partir del modelo booleano se puede utilizar para estudiar estrategias de control suave, con variación pequeña de concentración y parámetros de interacción en lugar de *KI* o *KO* de genes como se hizo en el trabajo original. El efecto de la variación de parámetros sobre los puntos de equilibrio, así como el análisis del sistema tetradimensional, se muestran en el siguiente capítulo.

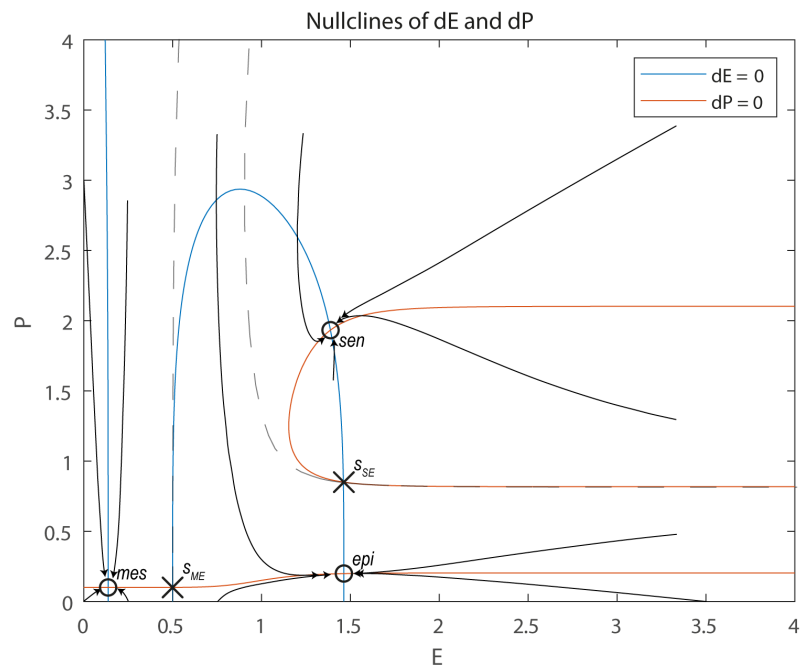


Figura 3.5: Plano fase del sistema descrito por las ecuaciones (3.6a) y (3.6b) con las ceroquinas (líneas continuas azul y roja), puntos de equilibrio estables (o) e inestables (x), un aproximado de las fronteras de las cuencas de atracción (líneas punteadas) y trayectorias ejemplo comenzando en diferentes regiones del plano fase (flechas negras).

Capítulo 4

Análisis de sensibilidad de los equilibrios inestables

4.1. Análisis de sensibilidad en el sistema bidimensional

El tamaño de las cuencas de atracción depende de la ubicación de la separatriz, la cual a su vez está anclada al equilibrio inestable. Entonces, una primera aproximación para manipular el tamaño de la cuenca de atracción, es mover el equilibrio inestable en el espacio de estados. Para no realizar una búsqueda exhaustiva del espacio de parámetros hasta encontrar una dirección de cambio que mueva al equilibrio en la dirección deseada, se puede utilizar el análisis de sensibilidad para determinar el cambio de la posición de los puntos de equilibrio al mover los parámetros.

El análisis de sensibilidad es un método utilizado para calcular el cambio de una trayectoria del sistema debido a una variación en los parámetros. Se considera un sistema dinámico de la forma

$$\dot{x} = F(x, \lambda) \tag{4.1}$$

donde F es una función no lineal dependiente del estado x y de un conjunto de parámetros λ . Para hacer este análisis en una escala global del sistema, para cualquier solución de la ecuación diferencial, se requiere del uso de métodos de fuerza bruta computacionalmente pesados. Sin embargo, en este trabajo nos interesan los efectos sobre un conjunto particular de soluciones que cumplen con la condición $F(x, \lambda) = 0$, lo cual reduce significativamente la complejidad del análisis. Como nos interesa el cambio en los equilibrios del sistema dependiendo

del valor de los parámetros, busquemos de manera implícita una función

$$\phi(\lambda) = x_{eq}$$

con la cual se pueda obtener el valor de los equilibrios en función de los parámetros. Así, la ecuación (4.1) se puede expresar como

$$F(\phi(\lambda), \lambda) \equiv 0 \quad (4.2)$$

Sabemos a partir del teorema de la función implícita (ver apéndice B) y el cuadro 3.4, que una función $\phi_i(\lambda) = x_{eq}$ existe en una vecindad de cada punto de equilibrio.

Como se busca conocer el cambio de la posición de los equilibrios con respecto a cada uno de los parámetros, nos interesa encontrar $\frac{\partial \phi}{\partial \lambda_i}$. Sin embargo, el teorema de la función implícita indica únicamente la existencia de una función, mas no una expresión de la misma. Por ende, dicha derivada debe ser encontrada utilizando métodos alternativos. Sabemos por la regla de la cadena que al derivar (4.2) con respecto a un parámetro λ_i se obtiene

$$\frac{\partial}{\partial \lambda_i} (F(\phi(\lambda), \lambda) \equiv 0) \Rightarrow J(F) \frac{\partial \phi}{\partial \lambda_i} + \frac{\partial F}{\partial \lambda_i} \equiv 0.$$

De esta expresión se puede despejar $\frac{\partial \phi}{\partial \lambda_i}$ para obtener

$$\frac{\partial \phi}{\partial \lambda_i} = -J(F)^{-1} \cdot \frac{\partial F}{\partial \lambda_i} \quad (4.3)$$

Como sabemos que el jacobiano de la función es invertible en cada uno de los puntos de equilibrio, la igualdad tiene una solución y se puede calcular la sensibilidad de cada punto de equilibrio a la variación de los parámetros. Es importante recordar que el valor de esta derivada es local con respecto a un conjunto de parámetros λ^* , que puede cambiar al cambiar los parámetros, o volverse incalculable por este método, por ejemplo al llegar a un punto de bifurcación. Sin embargo, el análisis nos permite obtener el cambio en los puntos de equilibrio utilizando un método analítico, como se ejemplifica en el apéndice D para el punto de equilibrio *mes* y el parámetro α_2 .

Este método para el cálculo de las sensibilidades es sumamente sencillo de realizar una vez se restringen las soluciones que se están estudiando. Por otro lado, métodos como el de Sobol, más generalmente utilizado en la literatura, requieren de mucho poder de cómputo, ya que utilizan métodos estocásticos como simulaciones por Montecarlo. Son métodos muy aptos para determinar el porcentaje de varianza global por la variación de uno o varios parámetros. Sin embargo, indican únicamente una magnitud de variación. Ya que en este análisis nos interesa estudiar no solo la magnitud sino también la dirección de variación, la determinación de la sensibilidad por el método propuesto es ideal.

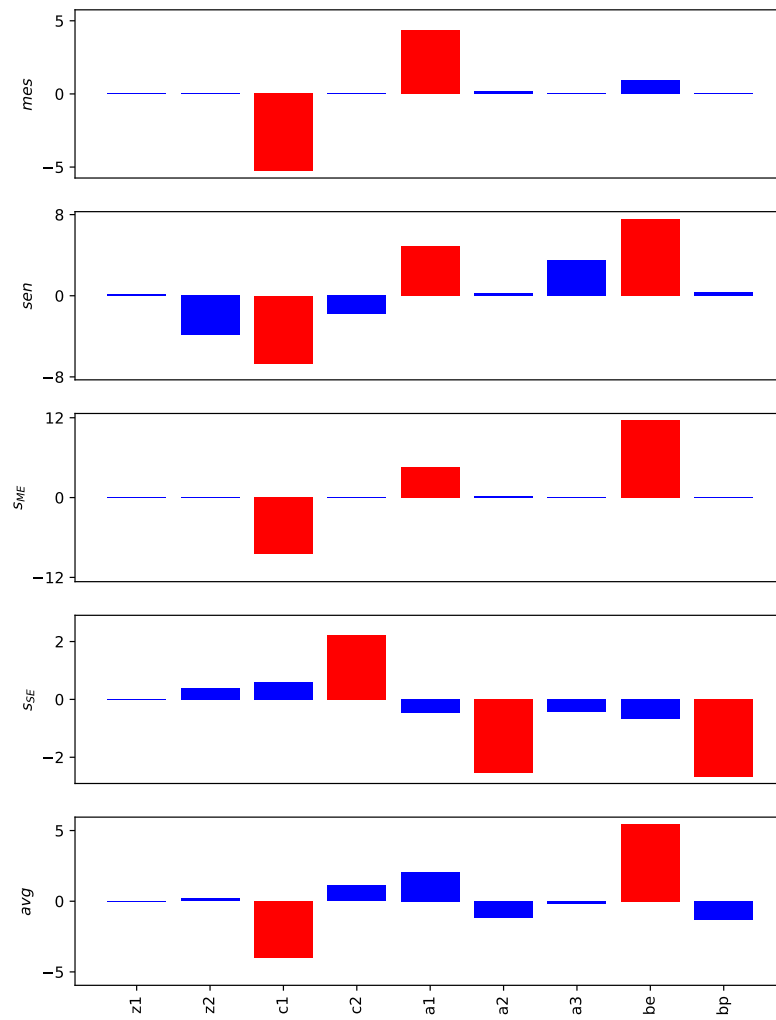
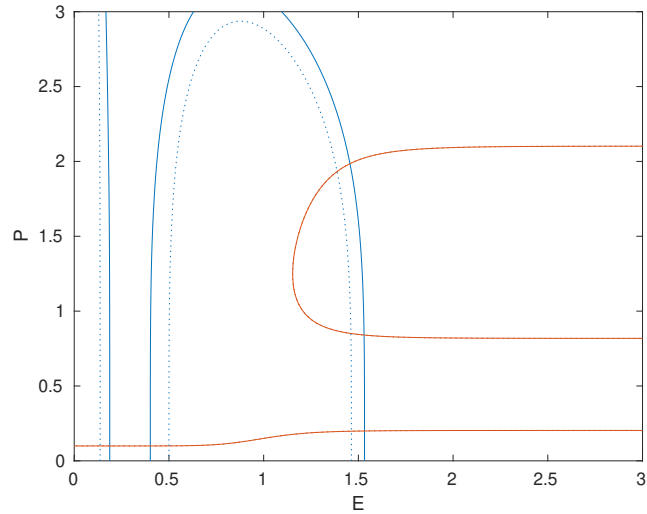
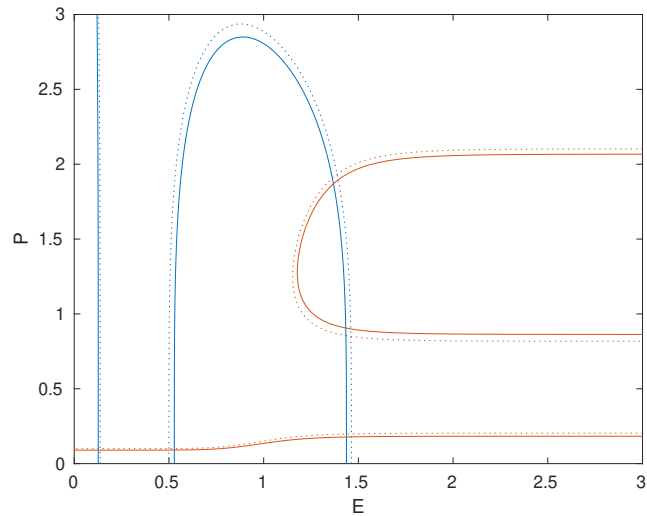


Figura 4.2: Representación de la sensibilidad de la posición de los equilibrios a la variación de parámetros. Se marcan en rojo los parámetros cuya modificación induzca un incremento en la distancia mayor al percentil 90.

4.1. ANÁLISIS DE SENSIBILIDAD EN EL SISTEMA BIDIMENSIONAL 47



(a)



(b)

Figura 4.3: *Modificación de la posición de las ceroclinas por la manipulación de los parámetros, como propuesto por el análisis de sensibilidad. En la figura 4.3a se cambió los parámetros de tal forma que el equilibrio inestable s_{ME} se moviera a la izquierda. En la 4.3b se modificaron para que el equilibrio inestable s_{SE} se moviera hacia arriba. Las ceroclinas con los parámetros originales se muestran con líneas punteadas.*

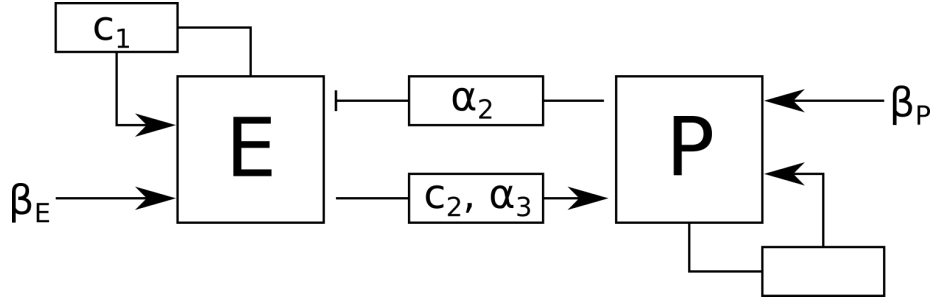


Figura 4.4: Diagrama de las interacciones del sistema de ecuaciones (3.6). Los elementos E y P representan los genes $ESE2$ y $p16$, respectivamente. Las flechas con terminación \rightarrow indican activación, mientras que aquellas con terminación \dashv indican inhibición. Se muestran los parámetros con alta sensibilidad junto a las interacciones a las que afectan.

de los parámetros, indicando que solo se puede llegar hasta cierto punto antes de llegar a una bifurcación. El efecto de la variación de parámetros sobre la posición del equilibrio s_{SE} se puede ver en la figura 4.3b.

Los parámetros propuestos, c_1 , c_2 , β_E , α_2 y α_3 , corresponden a las autorregulaciones de los elementos, y a la síntesis basal de E (fig. 4.4). Debido a las hipótesis tomadas para la simplificación del modelo, estos parámetros pueden corresponder a la interacción de E con S (c_1 y β_E) o a la regulación de P por sí mismo (c_2) o por S (α_2 y α_3). Los parámetros correspondientes a estas interacciones en el sistema tetradimensional se proponen como variables de control para el incremento de la cuenca de atracción epitelial.

Los resultados obtenidos en esta sección demuestran que se puede utilizar el análisis de sensibilidad para mover los puntos de equilibrio del sistema. Además, los parámetros obtenidos en el análisis del sistema bidimensional sirven como punto de partida para el análisis del sistema en cuatro dimensiones, el cual se discute en la siguiente sección.

4.2. Análisis de Sensibilidad del Sistema Tetradimensional

4.2.1. Parametrización del Sistema

El sistema (3.5) se muestra nuevamente a continuación:

$$\begin{aligned}\frac{dS}{dt} &= \alpha_1 \frac{k_1^n}{E^n + k_1^n} + \alpha_2 \frac{S^n}{S^n + k_2^n} \times \frac{E^n}{E^n + k_3^n} \times \frac{N^n}{N^n + k_4^n} + \beta_S - S \\ \frac{dE}{dt} &= \alpha_3 \frac{k_5^n}{S^n + k_5^n} + \alpha_4 \frac{E^n}{E^n + k_6^n} \times \frac{S^n}{S^n + k_7^n} \times \frac{k_8^n}{N^n + k_8^n} + \beta_E - E \\ \frac{dN}{dt} &= \alpha_5 \frac{S^n}{S^n + k_9^n} + \alpha_6 \frac{E^n}{E^n + k_{10}^n} + \alpha_7 \frac{N^n}{N^n + k_{11}^n} + \alpha_8 \frac{P^n}{P^n + k_{12}^n} + \beta_N - N \\ \frac{dP}{dt} &= \alpha_9 \frac{k_{13}^n}{S^n + k_{13}^n} \left[\left(\alpha_{10} \frac{E^n}{E^n + k_{14}^n} \times \frac{N^n}{N^n + k_{15}^n} \right) + \alpha_{11} \frac{P^n}{P^n + k_{16}^n} \right] + \beta_P - P.\end{aligned}$$

Al igual que en el sistema bidimensional (3.6), la variedad de comportamientos posibles dependiendo de parámetros es muy vasta. Entonces, la primera tarea es encontrar un conjunto de parámetros tal que el sistema tetradimensional se comporta como el sistema booleano; *i.e.*, que existan tres equilibrios estables los cuales sean cualitativamente similares a los equilibrios del sistema booleano (cuadro 3.2). Para lograr esto, se partió desde un conjunto de parámetros biológicamente irreal, el cual divide al sistema en dos partes con dinámica independiente. La primera considera únicamente a los componentes S y E , y se logra haciendo k_4 muy pequeño y k_8 muy grande. De esta manera, la dinámica del subsistema se vuelve insensible a la variación de N , y se puede buscar un conjunto de parámetros que resulte en un *switch* biestable entre S y E . Por otro lado, la segunda parte del sistema considera únicamente la dinámica de P , y se aísla haciendo k_{14} y k_{15} muy pequeños y k_{13} muy grande, de tal forma que P se vuelve insensible a la variación del resto del sistema. Con esto, se buscan valores de α_{11} , k_{16} y β_P que resulten en un sistema biestable para P cuando E está en concentración alta. Finalmente, se pueden relajar las condiciones sobre los parámetros de aislamiento. Se simuló el sistema para diferentes valores de parámetros hasta obtener un conjunto de tres equilibrios estables similares a los booleanos.

Ya que simulando el sistema se encuentran únicamente los equilibrios estables, y que el análisis de sensibilidad requiere de la identificación de los equilibrios inestables, se utilizaron métodos numéricos para identificar la totalidad de los equilibrios del sistema. De esta forma se identificaron siete equilibrios, de los cuales 3 corresponden a los equilibrios estables encontrados con simulaciones, y 4 son inestables (cuadro 4.1). A diferencia del caso bidimensional, en el que se encontró un equilibrio inestable para separar cada par de equilibrios estables, en este hay un aparente exceso de equilibrios inestables. Esto se debe a la existencia

de un equilibrio con variedad estable de dimensión 2 (ei_{2D} en el cuadro 4.1), el cual no puede dividir el espacio tetradimensional en dos partes, y por ende no forma una separatriz. Además, la existencia de este equilibrio inestable permite la existencia de una separatriz adicional cuya variedad inestable corresponde a una de las direcciones de la variedad estable de ei_{2D} .

Se realizó una caracterización de los equilibrios inestables para determinar qué equilibrios forman qué separatrices. Con este fin, se hicieron proyecciones de los equilibrios sobre un estimado de la variedad inestable (v_i) de cada equilibrio inestable. El resultado p de la proyección es un vector en la misma dirección de v_i que está lo más cerca posible del equilibrio (fig. 4.5). Como ambos vectores cambian únicamente en su magnitud, se puede decir que $p = cv_i$, donde c es un escalar. Entonces se puede comparar el signo de c de ambas proyecciones para determinar si los equilibrios se encuentran en la misma dirección de la variedad inestable, y por ende si están del mismo lado de la separatriz. Haciendo esto, se determina que

1. s_M y s_{M2} separan al equilibrio epitelial del mesenquimal y al mesenquimal del senescente.
2. s_S separa al equilibrio epitelial del senescente y al mesenquimal del senescente.
3. ei_{2D} se encuentra «entre» los equilibrios s_M y s_{M2} , con los cuales comparte variedad inestable.

Los resultados expuestos en los puntos 1 y 2 se corroboraron haciendo simulaciones del sistema al rededor de cada uno de los candidatos de separatriz, y evaluando a qué puntos de equilibrio llega el sistema. Para corroborar el punto 3 se realizaron simulaciones del sistema comenzando en una vecindad de ei_{2D} y viendo a qué equilibrios inestables se acercan las trayectorias, antes de llegar al equilibrio. Estos resultados indican que para incrementar la cuenca de atracción de epi , se debe mover a los equilibrios s_M , s_{M2} y ei_{2D} en dirección de mes , y s_S en dirección de sen .

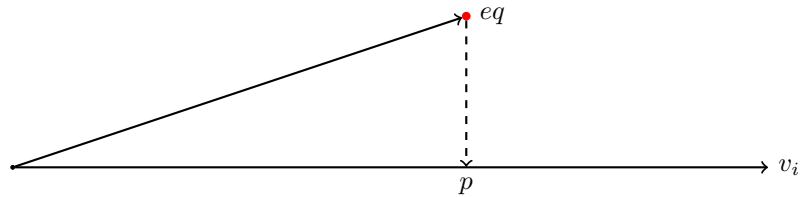


Figura 4.5: Representación de la proyección de un equilibrio sobre un vector v_i . El resultado de la proyección es un vector p que tiene la misma dirección que v_i . Como ambos tienen la misma dirección, se puede decir que $p = cv_i$.

Cuadro 4.1: *Equilibrios del sistema tetradimensional, restringidos al hipercubo de longitud 2. Los vectores de estado se representan como $[Snai2, ESE2, NFκB, p16]^T$.*

ID	Coordenadas	Valores Propios	Fenotipo
s_M	$\begin{bmatrix} 0.298 \\ 0.303 \\ 1.762 \\ 0.134 \end{bmatrix}$	$\{-2.180, \mathbf{0.189}, -0.962, -0.522\}$	—
ei_{2D}	$\begin{bmatrix} 0.298 \\ 0.303 \\ 2.056 \\ 0.663 \end{bmatrix}$	$\{-2.181, -0.974, \mathbf{0.187}, \mathbf{0.177}\}$	—
s_{M2}	$\begin{bmatrix} 0.297 \\ 0.303 \\ 2.159 \\ 1.141 \end{bmatrix}$	$\{-2.182, -0.978, \mathbf{0.189}, -0.209\}$	—
s_S	$\begin{bmatrix} 0.141 \\ 0.767 \\ 2.101 \\ 0.479 \end{bmatrix}$	$\{-1.616, -0.978, -0.333, \mathbf{0.244}\}$	—
epi	$\begin{bmatrix} 0.142 \\ 0.768 \\ 1.904 \\ 0.162 \end{bmatrix}$	$\{-1.615, -0.974, -0.335, -0.389\}$	Epitelial
mes	$\begin{bmatrix} 0.872 \\ 0.126 \\ 1.899 \\ 0.115 \end{bmatrix}$	$\{-1.662, -0.334, -0.745, -0.968\}$	Mesenquimal
sen	$\begin{bmatrix} 0.142 \\ 0.767 \\ 2.311 \\ 1.432 \end{bmatrix}$	$\{-1.618, -0.982, -0.333 - 0.396\}$	Senescente

4.2.2. Análisis de Sensibilidad

El análisis de sensibilidad se utilizó para determinar si los parámetros para maximizar la cuenca de atracción propuestos en la sección anterior sirven en el modelo tetradimensional. En otras palabras, si el modelo bidimensional se puede utilizar para obtener conclusiones sobre el sistema tetradimensional.

Recordando los resultados presentados al final de la sección 4.1, los parámetros candidato para la manipulación de la cuenca de atracción del fenotipo epitelial son c_1 , c_2 , α_2 , α_3 y β_E . Estos factores corresponden en el sistema bidimensional a factores que modulan la interacción de E con E (c_1 y β_E), la interacción de P con P (c_2), o la interacción de E con P (α_2 , α_3). Considerando los efectos de S en el sistema (3.5), estos parámetros se pueden asociar a k_1 , k_3 , k_5 , k_7 y β_E para la regulación de E , y α_9 , α_{10} , α_{11} , k_{13} , k_{14} , k_{15} y k_{16} para la regulación de P .

Ya que los resultados en cuatro dimensiones no se pueden graficar como se hizo con el sistema bidimensional, se recurrió a una estrategia distinta para el análisis de los resultados. Se definió una nueva función

$$f(\lambda) = (\phi_1(\lambda) - \phi_2(\lambda))^2 \quad (4.4)$$

donde $f : \mathbb{R}^m \rightarrow \mathbb{R}$ es una función que mide la distancia entre dos equilibrios dado un conjunto de parámetros. Igual que en el capítulo anterior, las funciones $\phi_i : \mathbb{R}^m \rightarrow \mathbb{R}^n$ son funciones que llevan del espacio de parámetros a los equilibrios del sistema y sabemos sobre su existencia debido al teorema de la función implícita.

Como $f(\lambda)$ es una función escalar de variable vectorial, su derivada es un vector de m componentes que apunta en la dirección del espacio de parámetros que lleva al mayor incremento de $f(\cdot)$. Esta se puede calcular como

$$\begin{aligned} \nabla f(\lambda) &= \nabla \left((\phi_1 - \phi_2)^2 \right) \\ &= 2(\phi_1 - \phi_2)^T \cdot J(\phi_1 - \phi_2) \\ &= 2(\phi_1 - \phi_2)^T \cdot (J(\phi_1) - J(\phi_2)) \end{aligned}$$

donde $J(\cdot)$ es la matriz jacobiana, y T es la operación de transposición. Al igual que en el caso bidimensional, no se cuenta con expresiones simbólicas de las funciones $\phi_i(\cdot)$ ni sus derivadas. Sin embargo, éstas no se requieren ya que el valor de la función corresponde a los valores de los puntos de equilibrio, y sus derivadas se pueden obtener resolviendo la igualdad (4.3).

Al evaluar el gradiente de la distancia entre *epi* y el resto de los equilibrios del sistema, se puede observar claramente una sobrerrepresentación de los parámetros k_1 , k_5 , β_E y β_S (fig. 4.6), lo cual corresponde satisfactoriamente con las hipótesis generadas con el sistema bidimensional. Sin embargo, el cambio en

la dirección de los parámetros es muy variable, aún entre las dos separatrices «equivalentes» s_M y s_{M2} . Para identificar una dirección que considere estas diferencias, se calculó el gradiente de una función adicional, que calcula el promedio de los gradientes de las distancias entre *epi* y los equilibrios inestables (ver *avg* en fig 4.6). Ya que los equilibrios inestables son los que modifican el tamaño de las cuencas de atracción, los otros equilibrios estables se pueden ignorar. En este caso, el vector gradiente apunta en una dirección de decremento de k_1 y β_S , e incremento en k_5 y β_E . Estos cambios son comprensibles y casi intuitivos, ya que k_1 y k_5 modulan el ciclo de inhibiciones entre *S* y *E*, y β_S y β_E modifican sus síntesis basales.

El análisis arroja también cambios de los parámetros β_P , α_2 y α_{11} . Estos parámetros están relacionados principalmente a la estabilización del fenotipo senescente (β_P y α_{11}) y la coactivación de *S* por *S*, *E* y *N* (α_2). Los cambios obtenidos para estos parámetros indican un incremento en los valores de β_P y α_{11} , y un decremento en α_2 . De estos, la reducción de β_P es un resultado obvio, ya que reduce la síntesis basal de *p16*, cuya alta concentración caracteriza el fenotipo senescente, y por ende reduce su cuenca de atracción. De manera similar, el decremento de α_2 reduce la magnitud de la activación de *S* por *S*, *E* y *N*. Es razonable entonces que al reducir el valor de este parámetro se reduzca también la cuenca de atracción de *mes*. Sin embargo, los cambios obtenidos para α_{11} estabilizan en vez de desestabilizar el equilibrio senescente; un valor mayor de α_{11} incrementa la fuerza de autoactivación de *P*. Aunque este cambio incrementa la cuenca de atracción de *sen*, también reduce la de *mes* en mayor medida. Entonces, el cambio neto en la cuenca de atracción de *epi* es positivo. Finalmente, aunque no tiene un efecto tan importante en el *avg*, es interesante resaltar el efecto de k_{16} , para el cual se propone un decremento. De manera similar a α_{11} , este cambio desestabiliza *mes* en mayor medida de lo que estabiliza a *sen*, y tiene un efecto neto positivo sobre la cuenca de atracción de *epi*.

Para corroborar que estos cambios realmente causan un incremento en la cuenca de atracción de *epi*, se hicieron simulaciones del sistema empezando en diferentes puntos del espacio de estados y se registró el atractor en el que finaliza la simulación. Éstas se realizaron con los parámetros originales, y un conjunto de parámetros modificado (cuadro 4.2) en la dirección propuesta por el análisis de sensibilidad (cuadro 4.3). Se puede ver que la cuenca de atracción de *epi* sí incrementa, junto con la cuenca de *sen* en menor medida. La única cuenca de atracción cuyo tamaño se reduce es la de *mes*.

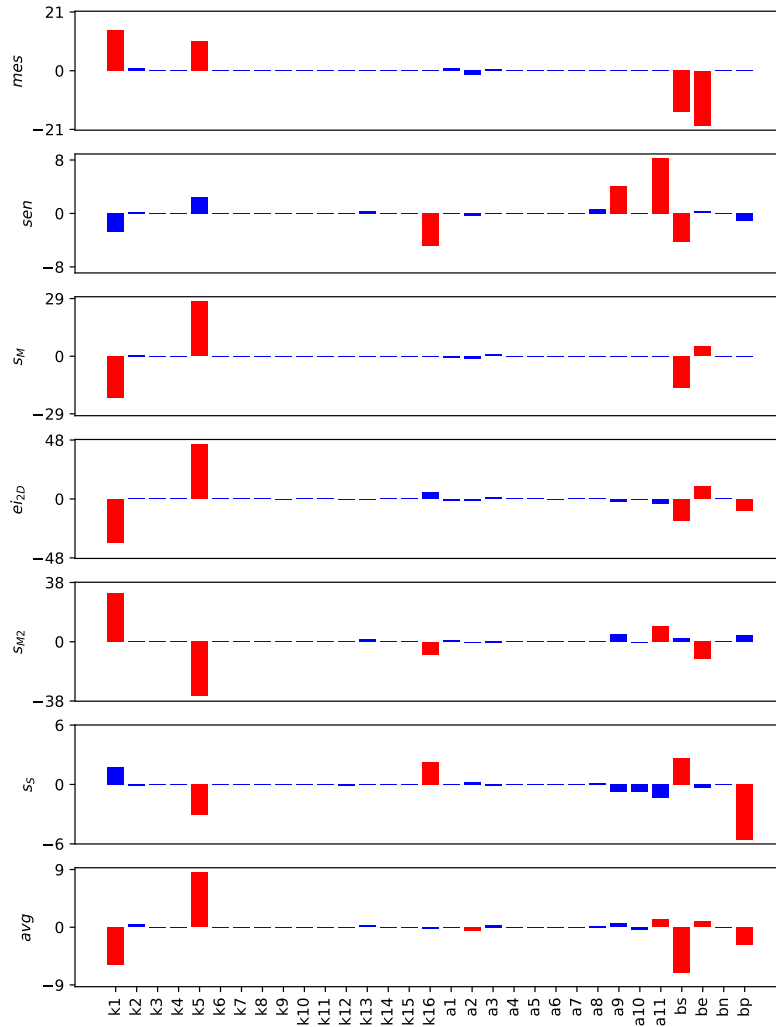


Figura 4.6: Dirección de cambio de los parámetros del sistema para el mayor incremento de la distancia entre el equilibrio epi y el resto de los equilibrios del sistema. Se marcan en rojo aquellos parámetros cuya variación induce un incremento en la distancia mayor al percentil 95.

Cuadro 4.2: Valores originales y modificados de los parámetros para la estimación del tamaño de las cuencas de atracción.

Parámetro	Valor Original	Valor Nuevo
n	2.0	2.0
k_1	0.1	0.109
k_2	0.1	0.099
k_3	2.0	2.0
k_4	0.1	0.1
k_5	0.1	0.086
k_6	2.0	2.0
k_7	0.1	0.1
k_8	2.0	2.0
k_9	0.5	0.5
k_{10}	0.5	0.5
k_{11}	0.5	0.5
k_{12}	0.5	0.5
k_{13}	1.0	1.0
k_{14}	2.0	2.0
k_{15}	2.0	2.0
k_{16}	1.0	1.0
α_1	2.0	2.0
α_2	0.1	0.101
α_3	2.0	2.001
α_4	0.1	0.099
α_5	0.5	0.5
α_6	0.5	0.5
α_7	0.5	0.5
α_8	0.5	0.5
α_9	2.0	2.0
α_{10}	0.1	0.099
α_{11}	1.0	1.001
β_S	0.1	0.098
β_E	0.1	0.111
β_N	1.0	0.999
β_P	0.1	0.1

Cuadro 4.3: Tamaño de las cuencas de atracción de los atractores del sistema con los parámetros originales y modificados en la dirección de máximo incremento del promedio de las distancias.

Equilibrio	Original	Modificada	Tasa de Cambio
epi	80	140	1.75
mes	375	250	0.6
sen	170	235	1.38

Capítulo 5

Discusión

5.1. Resumen de los resultados del trabajo

En el trabajo, se pudo construir un sistema de ecuaciones diferenciales que reproduce de manera satisfactoria los resultados del modelo booleano original. Por consiguiente, se espera que reproduzca de forma satisfactoria el fenómeno biológico.

Con el modelo obtenido se calculó la dirección óptima de cambio de parámetros para incrementar el tamaño de la cuenca de atracción de *epi*. Las modificaciones propuestas a los parámetros son razonables, ya que su relación con la dinámica del sistema y sus equilibrios es clara. Aunque muchos de los parámetros propuestos parecieran obvios, como las modificaciones sobre k_1 y k_5 , se obtuvieron resultados novedosos que proponen cambios de parámetros en direcciones poco esperadas, como el incremento de α_{11} y el decremento de α_2 y k_{16} . Estos parámetros representan concentraciones críticas para las interacciones, así como la magnitud del efecto de las interacciones. Estos no se podrían haber obtenido de una simple observación del sistema ya que su efecto positivo sobre la cuenca de *epi* es debido a la suma de diversos factores. Sería sumamente interesante evaluar su relevancia en experimentos *in vitro* o *in vivo*, o con resultados presentados en bases de interacción como KEGG, Human Reference Interactome Mapping Project, o Reactome.

En los resultados presentados, se observa una gran disparidad en la sensibilidad a ciertos parámetros. Sin embargo, los resultados del análisis son locales tanto para el espacio de parámetros como para el de estados. Sería interesante entonces, evaluar si existen condiciones biológicamente relevantes en las que la importancia de los parámetros cambie. Por ejemplo, una condición en la cual

alguna de las cuencas de atracción sea mínima.

Aunque el promedio de la distancia entre las sillas y el atractor *epi* sirvió para determinar una dirección de cambio en el espacio de estados, esta es una aproximación sumamente simplista. Ya que el sistema dinámico es altamente no lineal, la forma de las separatrices se puede alterar de maneras que no se pueden predecir por la posición del equilibrio. Además, ésta y otras alternativas sencillas ponen un peso mayor sobre las separatrices entre *mes* y *epi* debido a que existe una mayor cantidad de equilibrios entre ambos. Es posible que este sesgo analítico altere los resultados, y que la dirección de cambio obtenida no sea la ideal. Por esto, sería interesante estudiar la cantidad de puntos de equilibrio inestables del sistema. Aunque encontrar estos de manera experimental es sumamente complicado debido a su inestabilidad, se puede tener idea de su ubicación en el espacio realizando haciendo experimentos de ralentamiento crítico: el decremento en la velocidad de cambio del sistema al aproximarse a un punto crítico, como una separatriz, un equilibrio inestable, o un punto de bifurcación [27, 28, 29].

Aunque la aproximación con el primer teorema de Lyapunov y el análisis de sensibilidad es prometedora y sencilla, existen alternativas que permiten evaluar la cuenca de atracción de manera exacta utilizando el segundo teorema de Lyapunov. En este, se dice que el punto de equilibrio es estable si existe una función positiva que describa la energía interna del sistema, llamada función de Lyapunov, cuya derivada a lo largo de las trayectorias sea negativa [30]. En un corolario de este teorema, se indica que si el equilibrio es estable, existe una función de Lyapunov que describe exactamente la forma de la cuenca de atracción. Aunque esta alternativa es mucho más precisa que el estudio con el primer teorema de Lyapunov, encontrar funciones de Lyapunov es sumamente complicado, y quizás de poco valor para un sistema tan cualitativo como este. Sin embargo, si se quisiera utilizar el segundo teorema de Lyapunov, se podrían proponer funciones de forma cuadrática para estimar la cuenca de atracción por sus intersecciones, y ver cómo cambian estas con la variación de parámetros.

Es importante recordar que, ya que los modelos propuestos son de naturaleza cualitativa, sus parámetros son únicamente parámetros virtuales, que no reflejan de manera precisa la dinámica del sistema. Debido a esto, sería sumamente interesante profundizar en las interacciones propuestas con modelos enfocados en ellas. Por ejemplo, un modelo dinámico enfocado en la interacción únicamente entre los genes de los fenotipos epitelial y mesenquimal permitiría elucidar qué interacciones reflejan un cambio sobre las constantes k_1 y k_5 de este sistema.

5.2. Comparación con otras aproximaciones

Una alternativa a la metodología de funciones de Hill no acotadas es el cambio a funciones de Hill acotadas entre 0 y 1, que reflejan de manera más precisa lo que se obtiene con modelos booleanos. Esto lleva a concentraciones saturantes más claras y operaciones de conversión distintas. Sin embargo, impiden el uso de las constantes α_i , las cuales indican la magnitud del efecto de la interacción de un gen sobre otro. Esto quita mucha riqueza al análisis de sensibilidad, ya que este efecto podría ser también un blanco interesante para manipulación farmacológica.

El trabajo original utiliza métodos estocásticos booleanos para evaluar las probabilidades de transición entre atractores [13]. Observan que hay un camino preferente que toman las transiciones, en el cual se transita de un estado epitelial a uno senescente, y de ahí a uno mesenquimal. La transición por un estado senescente intermediario se ha observado *in vivo*, por lo cual este resultado da validez al modelo. El uso de sistemas estocásticos en ecuaciones diferenciales también se podrían utilizar para validar el modelo de manera similar, y encontrar un conjunto de parámetros que lo replique. Además, estos resultados nos hablan de los cambios de energía interna del sistema necesarios para causar una u otra transición. También sería interesante ver qué debe pasar a los parámetros para inducir el cambio opuesto.

Otra alternativa para hacer este trabajo implica el uso de modelos mecanísticos. Aunque el análisis de estos es relativamente similar a lo que se hizo en este trabajo, la construcción de los modelos es considerablemente más complicada. Se requiere de un conocimiento preciso de la cinética de interacción entre los elementos del sistema. La obtención de esta información depende de la realización de experimentos apropiados entre todos los elementos del sistema y sus intermediarios. Ya que obtenerla es muy complicado, y que las dinámicas se parecen a las ecuaciones de Hill [18], utilizar estos modelos no aporta mucha información nueva sobre el comportamiento del sistema. Sin embargo, ya que los parámetros sí reflejan comportamientos reales, la información obtenida del análisis es más aplicable a la realidad. Cabe resaltar que si se hacen simplificaciones como las que se hicieron en este modelo, los parámetros van a englobar nuevamente un conjunto de fenómenos en lugar de ser específicos. Entonces, la ventaja de un modelo cinético se puede perder.

5.3. Relación con el fenómeno biológico

El modelo obtenido refleja de manera satisfactoria el fenómeno biológico, ya que se obtienen los tres equilibrios estables reportados en la literatura [13]. Se podría alegar que la cantidad de equilibrios obtenida es insuficiente debido a que

se han reportado estados metaestables entre el fenotipo epitelial y mesenquimal durante la transición [11]. Sin embargo, es importante recordar que el propósito de este trabajo (y del trabajo que le precede) no es describir exactamente el fenómeno de la TEM, sino dar una idea de su regulación y su asociación a otros fenómenos como inflamación y senescencia. Además, es posible que la aparición de estados metaestables se pueda ver como el movimiento de los equilibrios en el espacio de estados por la variación de parámetros, lo cual permitiría entender a la transición como un paso por un continuo, como se sugiere en [11]. Para corroborar esto con el sistema presentado, se tendría que hacer un estudio más fino de las bifurcaciones del sistema, así como un análisis más global de las sensibilidades.

Ya que el modelo propuesto es una vista global del fenómeno de TEM, hay muchos detalles que salen del alcance del trabajo. Por ejemplo, se ha reportado que existen muchas vías alternativas para iniciar la TEM, y que éstas dependen del tejido o evento de transición particular [11]. Este hecho, además de los mencionados en secciones anteriores, incita a estudiar las diferentes formas en las que interactúan los módulos epitelial y mesenquimal de la RRG original. De esta manera, se podría entender cómo el contexto celular influye en las interacciones entre elementos regulatorios, y cómo puede alterar el resultado del proceso. Por ejemplo, se ha reportado que se necesita un decremento de *p16* para la transición entre el estado epitelial y mesenquimal en córneas [31]. Aunque esto va en contra a lo propuesto en [13], es posible que se deba a un fenómeno contexto-dependiente, y que esto permita un cambio en los parámetros apropiados del sistema para causar la transición.

5.4. Control del proceso biológico

La metodología utilizada permite hacer predicciones sobre la manipulación del sistema. Es importante recordar que las interacciones modeladas son únicamente interacciones virtuales, y que las interacciones «reales» son vías de señalización con muchos intermediarios y puntos de control. Entonces, la manipulación de un parámetro α_i se puede dar por mecanismos simples como un incremento en la fuerza de un promotor, hasta mecanismos más complicados como el tiempo de vida media de alguna molécula intermediaria. Por otro lado, la manipulación de los parámetros k_i se puede dar por algún efecto alostérico entre proteínas, el cual incrementa la afinidad entre diferentes moléculas de las vías de señalización y reduce la concentración crítica. La proposición de estos métodos depende de las vías de señalización involucradas, y deben considerar no solo los elementos que las componen, sino su interacción con otros procesos: la modificación permanente de un sitio promotor podría llevar a efectos imprevistos en otras funciones celulares debido a la interacción entre módulos.

Otras alternativas para el control de sistemas como el presentado implican

utilizar circuitos genéticos para hacer control en lazo abierto o lazo cerrado. De estos, el control en lazo cerrado permite realizar un control fino de la dinámica del sistema, ya que brinda al controlador robustez al ruido e incertidumbres [32]. Aunque aparentemente más sencillo, el control en lazo abierto implica utilizar un controlador que es función del tiempo en lugar del estado. Por ende, cualquier variación del sistema, ya sea por variación espontánea o por incertidumbre, causa que el resultado final obtenido no sea el deseado. Dependiendo de la sensibilidad del sistema a las condiciones iniciales, un error de este tipo puede tener conclusiones catastróficas.

Bibliografía

- [1] Francois Jacob and Jacques Monod. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 1961.
- [2] Thomas Schlitt and Alvis Brazma. Current approaches to gene regulatory network modeling. *BMC Bioinformatics*, 2007.
- [3] Robert May. Simple mathematical models with very complicated dynamics. *Nature*, 1976.
- [4] John J. Tyson and Béla Novák. Functional motifs in biochemical reaction networks. *Annual Review of Physical Chemistry*, 2010.
- [5] Dirk M. Lorenz, Alice Jeng, and Michael W. Deem. The emergence of modularity in biological systems. *Physics of Life Reviews*, 2011.
- [6] Leland H. Hartwell, John J. Hopfield, Stanislas Leibler, and Andrew W. Murray. From molecular to modular cell biology. *Nature*, 1999.
- [7] Berta Verd, Nicholas AM Monk, and Johannes Jaeger. Modularity, criticality, and evolvability of a developmental gene regulatory network. *eLife*, 2019.
- [8] Chaya Kalcheim. Epithelial-mesenchymal transitions during neural crest and somite development. *Journal of Clinical Medicine*, 2015.
- [9] Samy Lamouille, Jian Xu, and Rik Derynck. Molecular mechanisms of epithelial-mesenchymal transition. *Nature*, 2014.
- [10] Tsai-Tsen Liao and Muh-Hwa Yang. Revisiting epithelial-mesenchymal transition in cancer metastasis: the connection between epithelial plasticity and stemness. *Molecular Oncology*, 2017.
- [11] M. Angela Nieto, Ruby Yun-Ju Hwang, Rebecca A. Jackson, and Jean Paul Thiery. Emt: 2016. *Cell*, 2016.

- [12] Claudia Umbreit, Julia Flanjak, Christel Weiss, Philipp Erben, Christoph Aderhold, Anne Faber, Jens Stern-Straeter, Karl Hoermann 1, and Johannes David Schultz. Incomplete epithelial–mesenchymal transition in p16-positive squamous cell carcinoma cells correlates with β -catenin expression. *Anticancer Research*, 2014.
- [13] Luis Fernando Méndez-López, Jose Davila-Velderrain, Elisa Domínguez-Hüttinger, Christian Enríquez-Olguín, Juan Carlos Martínez-García, and Elena R. Alvarez-Buylla. Gene regulatory network underlying the immortalization of epithelial cells. *BMC Systems Biology*, 2017.
- [14] Bram De Craene and Gert Berx. Regulatory networks defining emt during cancer initiation and progression. *Nature Reviews Cancer*, 2013.
- [15] Mingyang Lu, Mohit Kumar Jolly, Herbert Levine, and José N. Onuchic Eshel Ben-Jacob. MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. *PNAS*, 2013.
- [16] Xiao-Jun Tian, Hang Zhang, and Jianhua Xing. Coupled reversible and irreversible bistable switches underlying $\text{tgf}\beta$ -induced epithelial to mesenchymal transition. *Biophysical Journal*, 2013.
- [17] Mihaela T. Matache and Valentin Matache. Logical reduction of biological networks to their most determinative components. *Bulletin of Mathematical Biology*, 2016.
- [18] Assieh Saadatpour, Réka Albert, and Timothy C. Reluga. A reduction method for boolean network models proven to conserve attractors. *Journal on Applied Dynamical Systems*, 2013.
- [19] Alan Veliz-Cuba. Reduction of boolean network models. *Journal of Theoretical Biology*, 2011.
- [20] Stuart Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 1969.
- [21] Nicolas Le Novère. Quantitative and logic modelling of molecular and gene networks. *Nature Reviews*, 2015.
- [22] Regina Samaga and Steffen Klamt. Modeling approaches for qualitative and semi-quantitative analysis of cellular signaling networks. *Cell Communication and Signalling*, 2013.
- [23] Stalin Muñoz, Miguel Carrillo, Eugenio Azpeitia, and David A. Rosenblueth. Griffin: A tool for symbolic inference of synchronous boolean molecular networks. *Frontiers in Genetics*, 2018.
- [24] Aurélien Naldi, Denis Thieffry, and Claudine Chaouiya. Decision diagrams for the representation and analysis of logical models of genetic networks. *Computational Methods in Systems Biology*, 2007.

- [25] Luis Mendoza and Ioannis Xenarios. A method for the generation of standardized qualitative dynamical systems of regulatory networks. *Theoretical Biology and Medical Modeling*, 2006.
- [26] Dominik M Wittmann, Jan Krumsiek, Julio Saez-Rodriguez, Douglas A Lauffenburger, Steffen Klamt, and Fabian J Theis. Transforming boolean models to continuous models: methodology and application to t-cell receptor signaling. *BMC Systems Biology*, 2009.
- [27] Martin Scheffer, Jordi Bascompte, William A. Brock, Victor Brovkin, Stephen R. Carpenter, Vasilis Dakos, Hermann Held, Egber H. van Nes, Max Rietkerk, and George Sugihara. Early-warning signal for critical transitions. *Nature Reviews*, 2009.
- [28] R. Wiebe and L. N. Virgin. On the experimental identification of unstable static equilibria. *Proceedings in Mathematical, Physical and Engineering Sciences*, 2016.
- [29] Robin M. Neville, Rainer M. J. Groh, Alberto Pirrera, and Mark Schenk. Beyond the fold: experimentally traversing limit points in nonlinear structures. *Proceedings in Mathematical, Physical and Engineering Sciences*, 2020.
- [30] Hassan K. Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, New Jersey, 2002.
- [31] Anil Tiwari, Sudha Swamynathan, Nicholas Alexander, John Gnalian, Shenghe Tian, Paul R. Kinchington, and Shivalingappa K. Swamynathan. Klf4 regulates corneal epithelial cell cycle progression by suppressing canonical $\text{tgf-}\beta$ signaling and upregulating cdk inhibitors p16 and p27. *Investigative Ophthalmology and Visual Science*, 2019.
- [32] Charles A. Desoer and Yung-Terng Wang. Foundations of feedback theory for nonlinear dynamical systems. *IEEE Transactions on Circuits and Systems*, 1980.
- [33] <http://www.math.toronto.edu/courses/mat237y1/20189/notes/Chapter3/S3.1.html>. Revisado: 19 de agosto de 2020.

Apéndice A

Reglas de Actualización de la RRG booleana de la TEM

$Snai2$

Original

$$\begin{aligned} Snai2(t+1) = & (\neg ESE2 \wedge \neg NF\kappa B \wedge \neg Snai2) \vee (\neg ESE2 \wedge \neg NF\kappa B \wedge Snai2) \\ & \vee (\neg ESE2 \wedge NF\kappa B \wedge \neg Snai2) \vee (ESE2 \wedge NF\kappa B \wedge Snai2) \end{aligned}$$

Simplificada

$$Snai2(t+1) = \neg ESE2 \vee (Snai2 \wedge ESE2 \wedge NF\kappa B)$$

ESE2**Original**

$$ESE2(t+1) = (\neg NF\kappa B \wedge \neg Snai2 \wedge \neg ESE2) \vee (\neg NF\kappa B \wedge \neg Snai2 \wedge ESE2) \\ \vee (\neg NF\kappa B \wedge Snai2 \wedge ESE2) \vee (NF\kappa B \wedge \neg Snai2 \wedge ESE2)$$

Simplificada

$$ESE2(t+1) = \neg Snai2 \vee (Snai2 \wedge ESE2 \wedge \neg NF\kappa B)$$

NF κ B**Original**

$$NF\kappa B(t+1) = \neg(\neg ESE2 \wedge \neg p16 \wedge \neg NF\kappa B)$$

Simplificada

$$NF\kappa B(t+1) = Snai2 \vee ESE2 \vee NF\kappa B \vee p16$$

p16

Original

$$\begin{aligned}
 p16(t+1) = & (\neg p16 \wedge \neg E2F \wedge p53 \wedge \neg TELasa \wedge \neg Snai2) \\
 & \vee (\neg p16 \wedge \neg E2F \wedge p53 \wedge \neg TELasa \wedge Snai2) \\
 & \vee (\neg p16 \wedge \neg E2F \wedge p53 \wedge TELasa \wedge \neg Snai2) \\
 & \vee (\neg p16 \wedge E2F \wedge p53 \wedge \neg TELasa \wedge \neg Snai2) \\
 & \vee (\neg p16 \wedge E2F \wedge p53 \wedge \neg TELasa \wedge Snai2) \\
 & \vee (\neg p16 \wedge E2F \wedge p53 \wedge TELasa \wedge \neg Snai2) \\
 & \vee (p16 \wedge \neg E2F \wedge \neg p53 \wedge \neg TELasa \wedge \neg Snai2) \\
 & \vee (p16 \wedge \neg E2F \wedge p53 \wedge \neg TELasa \wedge \neg Snai2) \\
 & \vee (p16 \wedge \neg E2F \wedge p53 \wedge \neg TELasa \wedge Snai2) \\
 & \vee (p16 \wedge \neg E2F \wedge p53 \wedge TELasa \wedge \neg Snai2) \\
 & \vee (p16 \wedge E2F \wedge \neg p53 \wedge \neg TELasa \wedge \neg Snai2) \\
 & \vee (p16 \wedge E2F \wedge \neg p53 \wedge \neg TELasa \wedge Snai2) \\
 & \vee (p16 \wedge E2F \wedge \neg p53 \wedge TELasa \wedge \neg Snai2) \\
 & \vee (p16 \wedge E2F \wedge \neg p53 \wedge TELasa \wedge Snai2) \\
 & \vee (p16 \wedge E2F \wedge p53 \wedge \neg TELasa \wedge \neg Snai2) \\
 & \vee (p16 \wedge E2F \wedge p53 \wedge \neg TELasa \wedge Snai2) \\
 & \vee (p16 \wedge E2F \wedge p53 \wedge TELasa \wedge \neg Snai2) \\
 & \vee (p16 \wedge E2F \wedge p53 \wedge TELasa \wedge Snai2) \\
 & \vee (p16 \wedge \neg E2F \wedge \neg p53 \wedge TELasa \wedge \neg Snai2)
 \end{aligned}$$

Simplificada

$$p16(t+1) = (p16 \wedge E2F) \vee (p16 \wedge \neg Snai2) \vee (p53 \wedge \neg TELasa) \vee (p53 \wedge \neg Snai2)$$

p53**Original**

$$\begin{aligned}
p53(t+1) = & (\neg p53 \wedge \neg NFkB \wedge \neg TELasa \wedge \neg p16 \wedge \neg Snai2) \\
& \vee (\neg p53 \wedge \neg NFkB \wedge \neg TELasa \wedge p16 \wedge \neg Snai2) \\
& \vee (\neg p53 \wedge NFkB \wedge \neg TELasa \wedge p16 \wedge \neg Snai2) \\
& \vee (p53 \wedge \neg NFkB \wedge \neg TELasa \wedge \neg p16 \wedge \neg Snai2) \\
& \vee (p53 \wedge \neg NFkB \wedge \neg TELasa \wedge p16 \wedge \neg Snai2) \\
& \vee (p53 \wedge NFkB \wedge \neg TELasa \wedge p16 \wedge \neg Snai2)
\end{aligned}$$

Simplificada

$$p53(t+1) = (\neg NFkB \vee p16) \wedge \neg TELasa \wedge \neg Snai2$$

Ciclina**Original**

$$\begin{aligned}
Ciclina(t+1) = & (\neg ESE2 \wedge \neg E2F \wedge \neg p16 \wedge \neg NFkB \wedge \neg Snai2) \\
& \vee (\neg ESE2 \wedge \neg E2F \wedge \neg p16 \wedge NFkB \wedge \neg Snai2) \\
& \vee (\neg ESE2 \wedge \neg E2F \wedge \neg p16 \wedge NFkB \wedge Snai2) \\
& \vee (\neg ESE2 \wedge E2F \wedge \neg p16 \wedge \neg NFkB \wedge \neg Snai2) \\
& \vee (\neg ESE2 \wedge E2F \wedge \neg p16 \wedge NFkB \wedge \neg Snai2) \\
& \vee (\neg ESE2 \wedge E2F \wedge \neg p16 \wedge NFkB \wedge Snai2) \\
& \vee (ESE2 \wedge \neg E2F \wedge \neg p16 \wedge \neg NFkB \wedge \neg Snai2) \\
& \vee (ESE2 \wedge \neg E2F \wedge \neg p16 \wedge NFkB \wedge \neg Snai2) \\
& \vee (ESE2 \wedge \neg E2F \wedge \neg p16 \wedge NFkB \wedge Snai2) \\
& \vee (ESE2 \wedge E2F \wedge \neg p16 \wedge \neg NFkB \wedge \neg Snai2) \\
& \vee (ESE2 \wedge E2F \wedge \neg p16 \wedge NFkB \wedge \neg Snai2)
\end{aligned}$$

Simplificada

$$Ciclina(t + 1) = \neg p16 \wedge ((\neg ESE2 \wedge NF\kappa B) \vee \neg Snai2)$$

TELasa**Original**

$$TELasa(t + 1) = (\neg Snai2 \wedge \neg ESE2) \vee (Snai2 \wedge \neg ESE2)$$

Simplificada

$$TELasa(t + 1) = \neg ESE2$$

Rb**Original**

$$\begin{aligned} Rb(t + 1) = & (\neg Cyclin \wedge \neg p16 \wedge p53) \vee (\neg Cyclin \wedge p16 \wedge \neg p53) \\ & \vee (\neg Cyclin \wedge p16 \wedge p53) \vee (Cyclin \wedge \neg p16 \wedge p53) \\ & \vee (Cyclin \wedge p16 \wedge \neg p53) \vee (Cyclin \wedge p16 \wedge p53) \end{aligned}$$

Simplificada

$$Rb(t + 1) = p16 \vee p53$$

E2F**Original**

$$E2F(t+1) = (\neg Rb \wedge \neg p53 \wedge \neg Snai2 \wedge \neg Cyclin) \\ \vee (\neg Rb \wedge \neg p53 \wedge \neg Snai2 \wedge Cyclin)$$

Simplificada

$$E2F(t+1) = \neg(Rb \vee p53 \vee Snai2)$$

Apéndice B

Definiciones

Todas las definiciones y el primer teorema fueron tomadas de [30]. El último teorema fue tomado de [33].

Definición 1. Equilibrio Estable

Sea $\hat{x}(t; x_0, t_0)$ una solución al sistema de ecuaciones diferenciales

$$\dot{x} = F(x)$$

con condición inicial $\hat{x}(0) = x_0$ y tiempo inicial t_0 . La solución \hat{x} es estable, si para toda $\epsilon > 0$, existe una $\delta > 0$ tal que

$$\hat{x}(t; x_0, t_0) - x(t; x_0 + \zeta, t_0) < \epsilon \Rightarrow \hat{x}(t; x_0, t) - x(t; x_0 + \zeta, t) < \delta; \forall t \geq t_0$$

Si no es estable, es inestable.

□

Definición 2. Equilibrio Asintóticamente Estable

Sea $\hat{x}(t; x_0, t_0)$ una solución al sistema de ecuaciones diferenciales

$$\dot{x} = F(x)$$

con condición inicial $\hat{x}(0) = x_0$ y tiempo inicial t_0 . La solución \hat{x} es asintóticamente estable si es estable y

$$\lim_{t \rightarrow \infty} x(t; x_0 + \zeta, t) - \hat{x}(t; x_0, t) = 0$$

□

Definición 3. Equilibrio Exponencialmente Estable

Sea $\hat{x}(t; x_0, t_0)$ una solución al sistema de ecuaciones diferenciales

$$\dot{x} = F(x)$$

con condición inicial $\hat{x}(0) = x_0$ y tiempo inicial t_0 . La solución \hat{x} es exponencialmente estable si es estable, y

$$\hat{x}(t; x_0, t) \leq x_0 e^{-kt}$$

□

Teorema 1. Primer teorema de Lyapunov

Una solución de la ecuación diferencial no lineal

$$\dot{x} = F(x)$$

es:

- estable, si el Jacobiano evaluado en ese punto es Hurwitz
- inestable, si el Jacobiano evaluado en ese punto tiene al menos un valor propio con parte real positiva

Si alguno de los valores propios del Jacobiano evaluado en ese punto tiene parte real nula, no se puede concluir nada sobre la estabilidad.

□

Teorema 2. Teorema de la Función Implícita

Asuma que S es un subconjunto abierto de \mathbb{R}^{n+k} y que $F : S \rightarrow \mathbb{R}^k$ es una función de clase C^1 . Asuma también que (a, b) es un punto en S tal que

$$F(a, b) = 0 \quad \text{y} \quad \det(D_y F(a, b)) \neq 0.$$

donde $\det(\cdot)$ es la operación matricial determinante, y $D_y F(a, b)$ es la derivada de F con respecto a y evaluada en el punto (a, b) . Entonces,

- i* existen $r_0, r_1 > 0$ tales que para toda $x \in \mathbb{R}^n$ tal que $|x - a| < r_0$, existe una única $y \in \mathbb{R}^k$ tal que $|y - b| < r_1$ y $F(x, y) = 0$. En otras palabras, si la derivada de F con respecto a y es no singular, puedo expresar $y = f(x)$ en una vecindad de x .

- ii* La función $f(x) = y$ es de tipo C^1 .

□

Apéndice C

Sistema sin la influencia de *Snai2*

El sistema sobresimplificado se presenta en la ecuación C.1.

$$\begin{aligned}\frac{dE}{dt} &= \alpha_1 \frac{E^n}{c_1^n + E^n} \times \frac{z_1^n}{z_1^n + P^n} + \beta_E - E \\ \frac{dP}{dt} &= \alpha_3 \frac{E^n}{z_2^n + E^n} + \alpha_3 \frac{P^n}{c_2^n + P^n} + \beta_P - P\end{aligned}\tag{C.1}$$

A diferencia de la ecuación (3.6), ésta no muestra un efecto claro de la dinámica de *Snai2* sobre *ESE2* y *p16*. Esta diferencia está presente en la falta de un factor m , y la falta de la «asistencia» de *ESE2* en la autoactivación de *p16*. Como se puede ver en las figuras C.1 y C.2, estas diferencias impiden la existencia de un conjunto de parámetros que haga al sistema triestable. Esto se concluye porque la forma de las ceroclinas hacen que haya una falta o exceso de puntos de equilibrio. El parámetro más ilustrativo de esto es α_2 , cuyo efecto sobre las ceroclinas se muestra en la figura C.2a. A diferencia de lo que se ve en el sistema (3.6), la variación de este parámetro no hace que se pierda la ceroclina de la mitad de abajo de la figura. Esto, dado un conjunto particular de parámetros, causa un exceso de puntos de equilibrio.

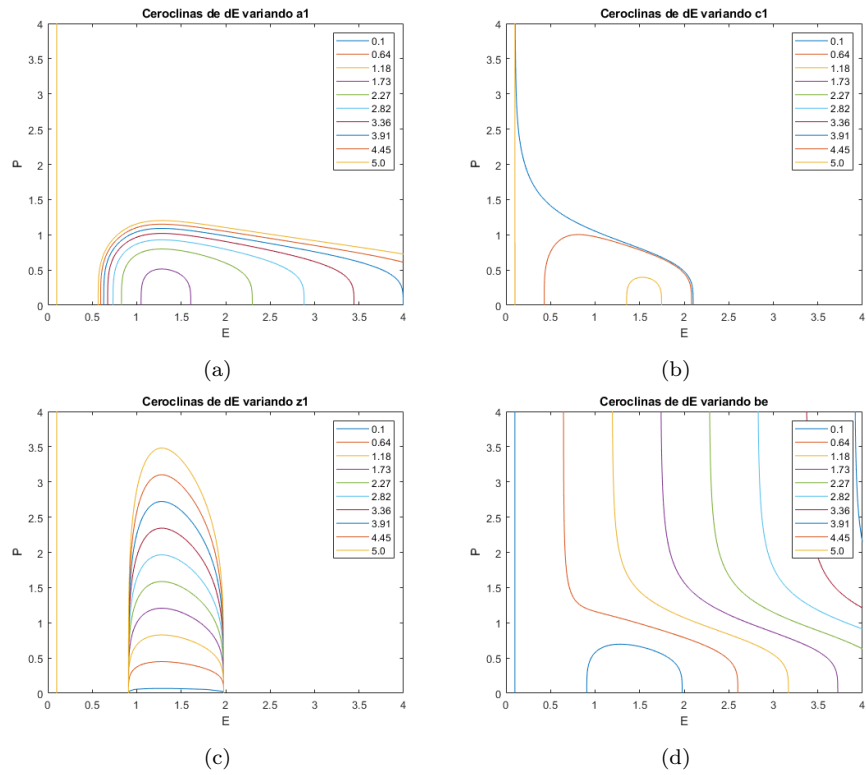


Figura C.1: Ceroclinas del sistema sin S_{nai2} para la ecuación $\frac{dE}{dt}$ en el sistema de ecuaciones diferenciales (C.1).

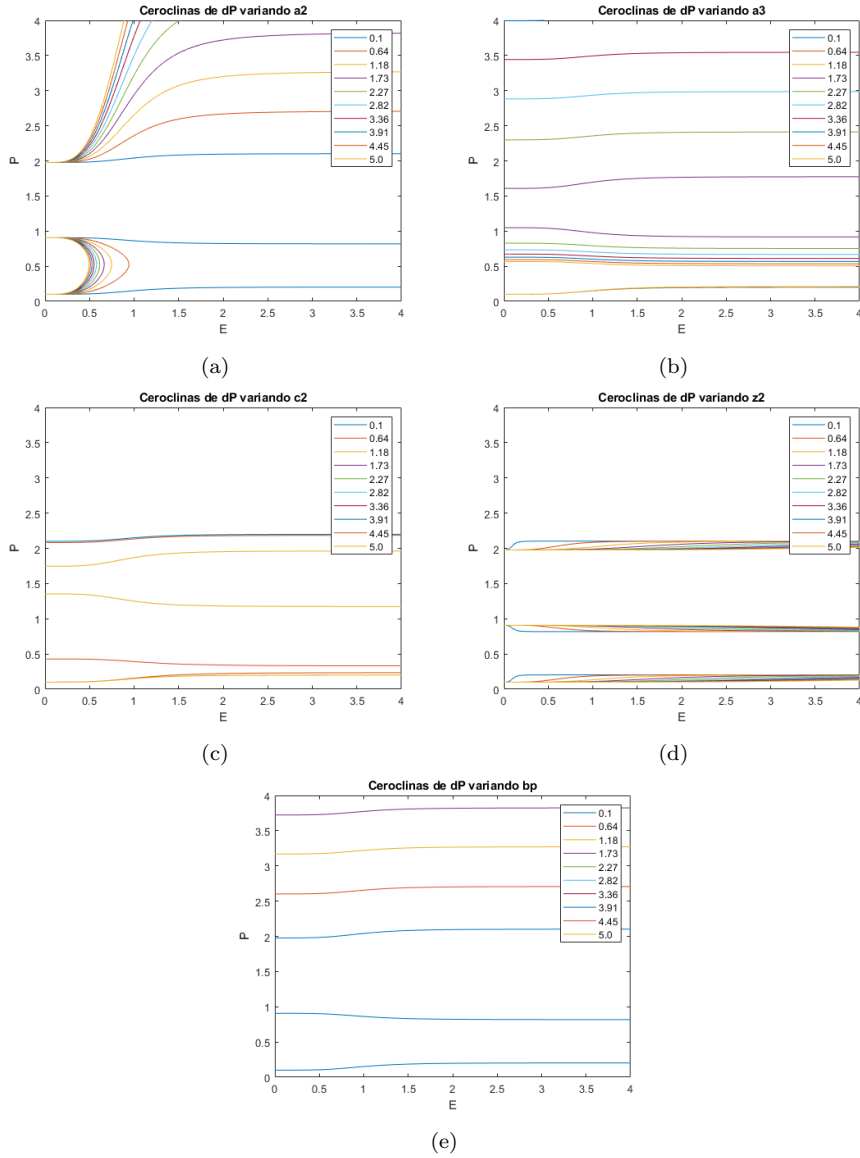


Figura C.2: Ceroclinas del sistema sin S_{nai2} para la ecuación $\frac{dP}{dt}$ en el sistema de ecuaciones diferenciales (C.1).

Apéndice D

Análisis de Sensibilidad a variación de parámetros del equilibrio *mes*

El equilibrio *mes* está ubicado en las coordenadas $[0.14, 0.1]^T$. Al evaluar el Jacobiano del sistema en el punto de equilibrio *mes*, se obtiene la matriz

$$J(F) |_{mes} = \begin{bmatrix} \frac{\partial F_1}{\partial E} & \frac{\partial F_1}{\partial P} \\ \frac{\partial F_2}{\partial E} & \frac{\partial F_2}{\partial P} \end{bmatrix} = \begin{bmatrix} -4.73e-1 & -2.34e-7 \\ 7.14e-7 & -9.99e-1 \end{bmatrix}$$

Sabemos del cuadro 3.4 que los valores propios del Jacobiano son -0.47 y -1, los cuales son ambos distintos entre sí y distintos de cero. Por ende, el Jacobiano evaluado en este punto es no-singular, y existe una función $\phi(\lambda) = mes$ en una vecindad de *mes*. Siguiendo la ecuación (4.3), podemos calcular la derivada de esta función si conocemos $J(F)^{-1} |_{mes}$ y $\frac{\partial F}{\partial \lambda_i}$. Ya que el determinante del Jacobiano es distinto de cero en *mes*, entonces su inverso existe. Podemos además obtener la derivada de $F(\cdot)$ con respecto a alguno de los parámetros de manera explícita, ya que tenemos una expresión de la función. Como ejemplo, se muestra el cálculo de la sensibilidad con respecto a la variación del parámetro α_2 .

$$\frac{\partial F}{\partial \alpha_2} = \begin{bmatrix} 0 \\ \frac{E^{nm}}{z_2^{nm} + E^{nm}} \end{bmatrix}$$

Al evaluar esta expresión en mes , se obtiene la matriz

$$\begin{bmatrix} 0 \\ 1.21e-7 \end{bmatrix}$$

Con esta matriz, y el inverso del Jacobiano, se obtiene

$$\begin{aligned} \frac{\partial \phi}{\partial \alpha_2} &= -J(F)^{-1} \Big|_{mes} \cdot \frac{\partial F}{\partial \alpha_2} \\ &= - \begin{bmatrix} -2.11 & 4.96e-7 \\ -1.51e-6 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1.21e-7 \end{bmatrix} \\ &= \begin{bmatrix} -6.04e-14 \\ 1.22e-7 \end{bmatrix} \end{aligned} \quad (D.1)$$

Esto nos indica que el punto de equilibrio mes se mueve hacia menores concentraciones de E y mayores concentraciones de P al hacer un cambio positivo e infinitesimal en α_2 . Adicionalmente, nos dice que el incremento en P es considerablemente mayor que el decremento en E