



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

MODELOS CONTINUOS Y DISCRETOS PARA EL
ESTUDIO DE LA RED GENÉTICA DE DESARROLLO
FLORAL DE ARABIDOPSIS THALIANA

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

FÍSICO

P R E S E N T A :

DIEGO ANDRÉS HARTASÁNCHEZ FRENK

TUTOR:

DR. MAXIMINO ALDANA GONZÁLEZ



FACULTAD DE CIENCIAS

UNAM

2008



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

Índice general

1. Introducción	2
2. Redes de regulación genética	6
2.1. Métodos	8
2.2. Redes Booleanas	9
2.3. Redes de Kauffman	10
3. <i>Arabidopsis thaliana</i>	12
3.1. Modelo ABC	12
3.2. Red de regulación genética de desarrollo floral de <i>Arabidopsis thaliana</i>	13
3.3. Resultados del análisis sincrónico de la red	15
4. Modelos	17
4.1. Modelos asíncronos	17
4.1.1. Modelo asíncrono de orden aleatorio secuencial	17
4.1.2. Modelo asíncrono de orden aleatorio grupal	18
4.2. Modelo de <i>Glass</i>	18
4.3. Determinación de atractores	20
4.3.1. Atractores en modelos asíncronos	20
4.3.2. Atractores en modelos de <i>Glass</i>	21
5. Resultados	23
5.1. Resultados del modelo asíncrono de orden aleatorio secuencial	23
5.1.1. Cuencas de atracción	24
5.1.2. Tiempos transitorios	26
5.2. Resultados del modelo asíncrono de orden aleatorio grupal	26
5.2.1. Atractores	27
5.2.2. Cuencas de atracción	27
5.3. Resultados del modelo de <i>Glass</i>	28
5.3.1. Variación de α_i	29
5.3.2. Variación de θ_i	30
6. Conclusiones	32

Capítulo 1

Introducción

A lo largo de la historia, la física ha probado ser una disciplina capaz de describir de manera muy precisa ciertos fenómenos de nuestro Universo y en los últimos años los físicos han decidido aplicar sus métodos al fenómeno más fascinante de todos: la vida. Esta incursión ha significado el nacimiento de áreas completamente novedosas en la investigación científica, donde la colaboración entre físicos, biólogos y matemáticos se ha convertido en una necesidad.

La evolución es un proceso continuo mediante el cual, las especies se han ido modificando a lo largo de sucesivas generaciones. Hasta el día de hoy, la explicación más aceptada del mecanismo de la evolución es la “teoría sintética de la evolución” que está basada en la genética mendeliana y en las ideas de Charles Darwin. Darwin propuso que el proceso evolutivo se da a partir de cambios aleatorios en los organismos y que estos cambios son seleccionados por un proceso llamado “selección natural”. Darwin fue un visionario, un gran genio y sus ideas han padecido una falta de comprensión por parte de muchos evolucionistas postdarwinianos que han tenido la tendencia de propagar una visión sumamente empobrecida de la selección natural centrada en la idea de la “lucha por la vida” y la “supervivencia del más fuerte” [27]. En mi opinión, las ideas de Darwin deben servir como una base de la cual partir para estudiar la vida y la evolución, y no un pedestal en el cual estemos subidos que nos impida descubrir nuevos terrenos.

En este sentido, cualquier incursión hacia nuevos terrenos debe ser bien recibida si está bien fundamentada. Por ejemplo, el concepto de auto-organización, de acuerdo con la propuesta de Stuart Kauffman [20], es un mecanismo a través del cual los elementos de un sistema se organizan por sí mismos sin la necesidad de un poder central que regule y dirija la actividad de todos los elementos. Este mecanismo actúa como una fuerza evolutiva desde etapas tempranas del desarrollo restringiendo la variabilidad (producida por cambios aleatorios) sobre la cual pueden actuar otros procesos evolutivos como la selección natural o la deriva génica.

Como dijera Jacques Monod:

“Muchos espíritus distinguidos, aún hoy, parecen no poder aceptar ni incluso comprender, que de una fuente de ruido la selección haya podido, ella sola, sacar todas las músicas de la biósfera. La selección opera, en efecto, sobre los productos del azar y no puede alimentarse de otra forma; pero opera en un dominio de exigencias rigurosas donde el azar es desterrado. Es de estas exigencias y no del azar, que la evolución ha sacado sus orientaciones generalmente ascendentes, sus conquistas sucesivas, el crecimiento ordenado del que ella parece dar la imagen.” [27]

Fue precisamente Monod quien en 1961, junto con François Jacob, descubrió que el control de la concentración de enzimas en células procariontes se da mediante la retroalimentación en la transcripción de secuencias de DNA. Sus experimentos y sus ideas representaron el surgimiento de lo que hoy son grandes áreas de estudio relacionadas con la regulación de la transcripción entre las cuales se encuentra la biología del desarrollo (desde el punto de vista molecular).

Con el avance de la biología molecular, en años recientes se ha logrado descifrar el genoma completo de muchos organismos y mucho trabajo se está haciendo para identificar y clasificar a las proteínas codificadas por estos genes. El siguiente reto es determinar las interacciones que existen entre los genes, las proteínas y otros componentes celulares e integrar este conocimiento para llevar nuestro entendimiento de los procesos de la naturaleza a otro nivel, que involucra, entre otras cosas, relacionar las características estructurales y dinámicas de las redes genéticas con el fenotipo de los organismos. Sabemos que las redes de interacción y regulación entre las diferentes entidades celulares es altamente compleja y su estudio requiere de mucho trabajo teórico, experimental y de modelaje [3].

Los genes y sus productos interactúan a distintos niveles formando redes. En particular, a nivel genómico, proteínas de un cierto tipo, llamadas factores de transcripción, activan o inhiben la transcripción de genes a mRNA. Debido a que estos factores de transcripción son, a su vez, productos de genes, el resultado es un grupo de genes que regulan mutuamente su expresión formando una red de regulación genética (una explicación más detallada se da en el capítulo 2).

Las redes de regulación genética juegan un papel sumamente importante en el desarrollo de los seres vivos. Sabemos que cada célula de un organismo contiene exactamente el mismo DNA. Sin embargo, en ningún momento de su existencia habrá una célula en particular en la que se expresen todos sus genes. Se debe precisamente a las diferencias en la expresión de los genes que una célula de riñón es diferente a una célula de piel, por ejemplo, en un ser humano. La red que se analizará en este trabajo es la red genética de desarrollo floral de *Arabidopsis thaliana*, una planta angiosperma empleada como organismo modelo en la biología.

Para poder comprender el funcionamiento de una red de regulación genética, es necesario conocer, en primer lugar, los componentes de la red y sus interac-

ciones. Esta información estructural debe complementarse necesariamente con la información cinética de las interacciones. Como sabemos, el desarrollo (de la flor, en este caso) es un proceso dinámico y al ser así, para estudiarlo se requiere definir una variable que indique el estado de cada componente de la red y analizar cómo se modifican estos estados a partir de las interacciones en la red.

El mapeo de redes de regulación genética es bastante reciente. Sin embargo, el modelaje de aspectos generales de las redes de regulación comenzó a finales de los años sesenta con el trabajo de Stuart Kauffman, en particular, y de otros pioneros en el área como René Thomas.

En ausencia de datos experimentales, Kauffman estudió redes genéticas construidas aleatoriamente. En sus redes considera a todos los genes equivalentes y le asigna aleatoriamente un número K de inputs de manera aleatoria. Las variables que describen el estado de los genes son variables binarias (encendido/apagado) y la dinámica de cada gen depende de una función Booleana cuyos argumentos son precisamente los inputs de cada gen (para una descripción detallada de las redes de Kauffman, ver la sección 2.3). Kauffman estudió la dinámica de estas redes Booleanas aleatorias concentrándose en los atractores encontrados en el espacio fase. Entre otros hallazgos, Kauffman propuso que se podría identificar a los atractores de una red de regulación genética con los diferentes tipos o destinos celulares, o más generalmente, con el estado funcional de la célula [18]. Además notó que el número de tipos celulares parecía aumentar aproximadamente con la raíz cuadrada del número de genes por célula, sugiriendo que las redes de regulación genética se encuentran en un estado crítico, en la frontera entre el orden y el caos [19].

Las redes Booleanas utilizadas por Kauffman se han aplicado al estudio de redes de regulación genética encontradas en la naturaleza. La red de genes de polaridad de segmentos en la mosca de la fruta *Drosophila melanogaster* [1, 3, 10] y la red de desarrollo floral de *Arabidopsis thaliana* [13, 24, 25] son ejemplos, entre otros [16, 26], en los que los atractores dinámicos del modelo se han identificado con tipos celulares.

Sin embargo, el método de Kauffman ha sido criticado pues se pensaba que el modelo era una descripción muy lejana de la realidad. En esta tesis se proponen 3 modelos (dos asíncronos y uno continuo) que introducen diferencias respecto al modelo de Kauffman con el propósito de demostrar que estas críticas no están bien fundamentadas. Mediante el estudio de la red de desarrollo floral de *Arabidopsis thaliana* con estos 3 modelos, se pretende mostrar que alteraciones en el modelaje de la red no acarrearán resultados tan diferentes. Estos análisis pretenden dar una evidencia más de que es la estructura básica de la red y no los detalles finos los que determinan la dinámica de la red.

Así pues, me embarco en esta tarea no sin antes recordar a uno de los físicos que me inspiró a adentrarme en el terreno de la biología a pesar de haber estudiado física. En el primer capítulo de “*What is Life?*” [30], seguramente uno de los libros que más ha influenciado a la ciencia desde mediados del siglo XX, Erwin Schrödinger escribió:

“Propongo, en primer lugar, desarrollar lo que podríamos llamar las ideas inocentes de un físico acerca de los organismos, es decir, las ideas que pueden surgir en la mente de un físico quien, después de haber aprendido su física y, particularmente, los fundamentos estadísticos de su ciencia, comienza a pensar en los organismos y en la forma en la que se comportan y funcionan y que llega a preguntarse de una manera concienzuda si él, a partir de lo que ha aprendido, desde el punto de vista de su comparativamente simple, clara y humilde ciencia, puede contribuir de forma relevante a resolver la pregunta.”¹

¹Traducción del autor

Capítulo 2

Redes de regulación genética

El dogma central de la biología molecular establece que la información almacenada en el DNA se transcribe a RNA mensajero (mRNA) y luego se traduce a proteínas. Repasemos rápida y simplificada cómo se lleva a cabo este proceso. El DNA (siglas en inglés del ácido desoxirribonucleico) es la molécula en la cual está almacenada la información genética de todo ser vivo. En las células eucariontes, el DNA se encuentra en el núcleo y cada célula tiene una sola copia de DNA. Al dividirse una célula, el DNA se duplica de manera que ambas células resultantes mantienen una copia del DNA de la célula original. El RNA (ácido ribonucleico) es una molécula muy parecida al DNA que tiene distintas formas y funciones celulares. Una forma en la que encontramos al RNA se conoce como RNA mensajero (mRNA) y su nombre proviene de su función de intermediario entre el DNA y la síntesis de proteínas. Un gen es un fragmento de DNA, y al expresarse su información se transcribe ensamblándose un fragmento de mRNA. Este mRNA se recorta de manera que sólo queda en él la información relevante. El mRNA se acopla a un ribosoma y la información que contiene se utiliza para sintetizar una proteína, en un proceso que se conoce como traducción. Decimos entonces que cada gen codifica para una proteína, o para un pedazo de proteína.

Las proteínas son la maquinaria de la célula. Podríamos decir que cada proteína es una máquina especializada en una función celular específica. En la célula existen proteínas que actúan ya sea como enzimas catalizando reacciones, como receptores de membrana respondiendo a estímulos externos, como canales iónicos regulando el flujo de partículas cargadas entre el interior y el exterior de la célula, como medios de transporte dentro de la célula o como componentes en rutas metabólicas [12]. Pero existe otra función de las proteínas que resulta sumamente interesante: algunas proteínas se unen directa o indirectamente al DNA para regular su transcripción. Es decir, que DNA, RNA y proteína forman un ciclo de retroalimentación cuando las proteínas sintetizadas a partir de la

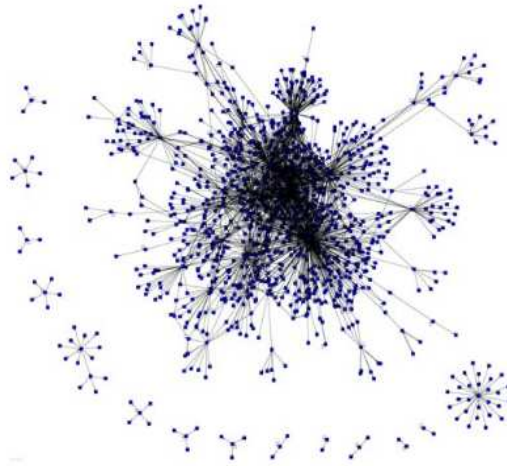


Figura 2.1: Representación gráfica de la red de transcripción genética de *E. Coli* K-12. Para esta red, la probabilidad de que un gen cualquiera tenga K reguladores es $P(K) = e^{-z} z^K / K!$ (topología de Poisson en inputs), mientras que la probabilidad de sea un regulador de otros n genes es $P(n) = Cn^{-\gamma}$ (topología libre de escala en outputs). (Imagen reproducida con el consentimiento de M.A. [5].)

expresión de ciertos genes, regulan a su vez cuándo, cuánto y en qué células se van a expresar esos u otros genes. A estas proteínas se les conoce con el nombre de *factores de transcripción*.

La regulación de la expresión genética no sólo se da en la transcripción del DNA (por la presencia o ausencia de factores de transcripción); también se da en otras etapas de la síntesis de proteínas: mediante regulación de la traducción de RNA, durante el transporte de RNA o mediante la modificación de proteínas [12]. En cualquier caso, los factores responsables de estas regulaciones son, a su vez, proteínas sintetizadas a partir de la expresión de otros genes. De esta manera, tenemos un conjunto de factores (DNA, RNA y proteínas) que están relacionados por medio de una *red de regulación*.

Las redes de regulación genética pueden ser sumamente complejas ya que involucran interacciones entre un gran número de factores a distintos niveles. Para poder reconstruir una red de regulación genética se requieren técnicas de experimentación avanzadas. En los últimos años se han desarrollado técnicas experimentales como los microarreglos de cDNA que han hecho posible medir de manera rápida y en forma paralela los niveles de expresión de varios genes. De esta manera, se han podido obtener datos reales de redes de regulación en organismos como *Escherichia coli* [29], *Bacillus subtilis* [22] y *Saccharomyces cerevisiae* [21]. A partir de estos datos, se han llevado a cabo análisis para conocer la estructura (topología) de estas redes [4]. Los análisis muestran que

estas redes tienen una topología de Poisson en inputs y una topología libre de escala en outputs (ver Figura 2.1) (para una definición precisa de inputs y outputs ver la sección 2.2).

Por otro lado, se puede modelar, simular y estudiar la dinámica de las redes de regulación genética mediante métodos computacionales.

2.1. Métodos

Monod y Jacob fueron los primeros en proponer que el genoma debía consistir en circuitos complejos que conformasen programas para estabilizar el metabolismo y regular la diferenciación celular [28]. Propusieron también que para estudiar estos circuitos genéticos se podían utilizar modelos *continuos* basados en ecuaciones diferenciales.

Los modelos continuos son la herramienta más utilizada cuando se intenta modelar procesos regulatorios. Asumiendo ciertos detalles (en ocasiones, mucho más trascendentes de lo que estamos dispuestos a admitir), se puede transformar prácticamente cualquier sistema de reacciones químicas y restricciones físicas en un sistema de ecuaciones diferenciales ordinarias no lineales, cuyas variables son las concentraciones de proteínas, mRNA, complejos químicos, etc. Conocer todas las variables cinéticas necesarias para describir un sistema determinado resulta cada vez más complicado conforme aumenta el tamaño del sistema. Además, otro aspecto que resulta cada vez más relevante al crecer el tamaño del sistema es el tiempo de cómputo necesario para encontrar las soluciones [14].

En los modelos *Booleanos*, más sencillos que los continuos, los genes sólo tienen dos estados posibles: “completamente expresados” o “no expresados”. Estos modelos tienen ciertas ventajas: dado un estado al tiempo t , el estado al tiempo $t + 1$ está dado por una sencilla función Booleana; además, dado que el espacio-fase es finito, en principio se pueden explorar todos los estados posibles (en ocasiones habrá que aplicar “fuerza bruta computacional”), a diferencia de lo que ocurre en un espacio-fase infinito. Los modelos Booleanos consideran que todos los genes cambian de estado al mismo tiempo (sincronía) [14].

También existen diversos modelos *híbridos*. En algunos, se combinan funciones Booleanas y continuas para determinar el estado de distintos genes [31], y en otros, los estados de expresión genética son discretos y las transiciones entre estados se determinan por ecuaciones diferenciales lineales con coeficientes constantes. Estos modelos permiten la introducción de distintos umbrales de expresión y tiempos de relajación para cada gen [14].

En este trabajo, se realiza un análisis de robustez de la red de regulación genética de desarrollo floral de *Arabidopsis thaliana*. Para llevarlo a cabo, se comparan los resultados obtenidos con el modelo Booleano con los resultados de otros modelos. En particular, se analiza la red con tres diferentes modelos: “asíncrono de orden aleatorio secuencial”, “asíncrono de orden aleatorio grupal” y “de *Glass*” (para una descripción detallada de cada uno de estos modelos, ver el capítulo 4).

Mediante este análisis se demuestra que, en el caso de la red de desarrollo

floral de *Arabidopsis thaliana*, las críticas al modelo de Kauffman no están bien fundamentadas. El modelo de Kauffman ha sido criticado tanto por su carácter síncrono, como por su carácter discreto, y los modelos utilizados tienen modificaciones respecto al modelo Booleano precisamente en estos dos aspectos. La similitud entre los resultados de los distintos modelos apoya a la idea de que a pesar de ser modelos sumamente simplificados, los modelos Booleanos son una excelente herramienta para modelar redes de regulación genética sin necesidad de recurrir a modelos más complejos [2, 7].

2.2. Redes Booleanas

Una red está compuesta por un determinado número de nodos y reglas de interacción. En el caso de las redes genéticas, se identifica cada nodo de la red con un gen. El estado de expresión de cada gen está representado por una variable Booleana X_i . Cuando el gen i está expresado (o “activado”), $X_i = 1$, cuando no esté expresado, $X_i = 0$. Dada una red con N genes, el estado de todo el sistema se representa por un conjunto (de tamaño N) de estas variables: $\{X_1, X_2, \dots, X_N\}$. El número total de estados posibles de la red será 2^N .

Como se trata de una red genética, las reglas de interacción son formas de *regulación genética*. Existen esencialmente dos tipos de regulación: positiva y negativa. Cada gen puede estar regulado positiva o negativamente por cualquier gen de la red, incluido él mismo. El número de genes que regulen a un determinado gen, serán los *inputs* de dicho gen, mientras que el número de genes que son regulados por un gen determinado serán sus *outputs*.

Cada variable X_i tiene asociado un conjunto de k_i inputs (o reguladores), que determinan el estado de expresión de X_i .

$$X_i \leftarrow \{X_{i_1}, X_{i_2}, \dots, X_{i_{k_i}}\} \quad (2.1)$$

El estado de expresión de los genes de la red cambia en el tiempo, de manera síncrona, de acuerdo a la ecuación:

$$X_i(t + \tau) = F_i(X_{i_1}(t), X_{i_2}(t), \dots, X_{i_{k_i}}(t)), \quad i = 1, 2, \dots, N \quad (2.2)$$

donde F_i es una función Booleana que se construye de acuerdo con la naturaleza activadora (positiva) o inhibidora (negativa) de los reguladores.

Dado un estado del sistema al tiempo t , el sistema cambia determinísticamente a otro estado al tiempo $t+1$ y a este estado le llamaremos su *sucesor*.

Un análisis del espacio de atractores de la red nos brinda un conocimiento amplio de la red misma. Debido a que existen un número finito de estados en la red, partiendo de un *estado inicial*, el sistema pasará por un secuencia de estados y eventualmente deberá alcanzar un estado por el cual ya haya pasado. A partir de este momento el sistema atravesará una misma secuencia de estados *ad infinitum* y le llamaremos a esta secuencia *atractor* de la red. La *longitud del atractor* será el número de estados que lo conforman, a saber, el número de estados que el sistema debe atravesar antes de alcanzar un estado previamente

alcanzado. A los atractores de longitud 1 (un estado cuyo sucesor es él mismo), les llamaremos *atactores puntuales* o *estados de equilibrio*, mientras que a todos los demás les llamaremos *atractores cíclicos*. A los estados que conforman un atractor les llamaremos *componentes del atractor*.

Otras definiciones serán necesarias para describir más fácilmente la red. El *tiempo transitorio* es el número de estados entre un determinado estado inicial y la primer componente del atractor que sea alcanzada. A cada estado corresponde un tiempo transitorio. Nótese que aquí estamos equiparando un tiempo con un número de estados. Esto no representa problema alguno pues en el modelo Booleano, en cada unidad de tiempo se cambia de un estado a otro. Por otro lado, una *cuenca de atracción* estará conformada por todos los estados que eventualmente alcancen el mismo atractor, incluyendo los componentes del atractor mismo. El *tamaño de la cuenca* será el número de estados que la conformen.

2.3. Redes de Kauffman

En 1969, Stuart A. Kauffman publicó un artículo reportando el comportamiento de grandes redes de “genes” interconectados de manera aleatoria [18]. Cada uno de estos “genes” era considerado binario, es decir, que sólo tenía dos posibles estados (encendido y apagado). En este artículo, Kauffman presentó evidencia de que redes de genes (binarios) conectados aleatoriamente eran altamente ordenados y estables, como los seres vivos; que estos sistemas presentaban modos o ciclos de comportamiento y que las longitudes de estos ciclos predecían el tiempo de replicación de las células en muchos *phyla*; que el número de diferentes modos de comportamiento distinguibles en una determinada red genética con conexiones aleatorias predecían el número de tipos celulares en un organismo con una red genética del mismo tamaño; que en la presencia de ruido, un determinado estado podía diferenciarse en tan solo unos cuantos modos de comportamiento y que estas limitadas posibilidades de transición entre distintos modos de comportamiento permitían establecer una teoría de diferenciación para redes genéticas con conexiones aleatorias.

El carácter visionario de Kauffman le hizo percatarse, desde un principio, de las inmensas posibilidades que significaba ahondar en este tema. El trabajo de Kauffman se ha basado en la hipótesis de que las principales propiedades biológicas de los seres vivos dependen de su estructura bruta y no de los detalles específicos de la misma. En su libro titulado “*At Home in the Universe*” [20], Kauffman escribió:

“Los organismos no son sistemas aleatorios sencillos, sino sistemas heterogéneos y altamente complejos que han evolucionado durante casi cuatro mil millones de años. La esperanza de que se pueda construir una teoría del orden biológico radica en la existencia de propiedades clave de estos organismos que no dependan de los detalles. Si las propiedades de los sistemas vivos dependen de los detalles en su estructura y su lógica, si los organismos

*son mecanismos arbitrarios contruidos a partir de propiedades arbitrarias hasta el último detalle, entonces los problemas epistemológicos que enfrentaremos al intentar entender la maravilla de la biósfera serán vastos. Si, en cambio, los principales fenómenos, aquéllos de la mayor importancia, no dependen de los detalles, entonces podremos soñar con encontrar teorías profundas y hermosas.”*¹

Esta tesis es una pequeña contribución a la evidencia experimental que apoya la hipótesis de Kauffman. En la red de regulación del desarrollo floral de *Arabidopsis thaliana*, el modelo de Kauffman ha resultado ser un modelo excelente para reproducir resultados experimentales a pesar de no hacer énfasis en los detalles de la red sino únicamente en su estructura bruta. Los detalles finos de la red no se conocen y por lo tanto sería imposible en este momento analizar la red mediante modelos más precisos. Sin embargo, se pueden simular estos detalles introduciendo modificaciones aleatorias al modelo de Kauffman. Si la introducción de variables aleatorias al modelo no afectan los resultados del análisis, la conclusión obvia es que los resultados son independientes de los detalles, y en esto consiste precisamente esta tesis.

¹Traducción del autor

Capítulo 3

Arabidopsis thaliana

El objeto de estudio de esta tesis es un grupo de genes pertenecientes a *Arabidopsis thaliana* (L.) Heynh. (a partir de ahora, *Arabidopsis*). Esta planta angiosperma se emplea como organismo modelo para el estudio de la biología molecular, la fisiología y la genética de las plantas. Fue la primera planta cuyo genoma fue completamente secuenciado. La secuenciación se terminó en 2000 por la “Iniciativa para el Genoma de *Arabidopsis*” [32] (AGI, por sus siglas en inglés). *Arabidopsis* presenta ventajas muy importantes para la investigación básica:

- Se tiene un mapeo genético completo de sus 5 cromosomas.
- Tiene un genoma pequeño que consta de 157 millones de pares de bases [6] y alrededor de 27,000 genes que codifican 35,000 proteínas [17].
- Tiene un ciclo de vida muy rápido: alrededor de 6 semanas entre la germinación y la aparición de semillas maduras.
- Su pequeño tamaño es ideal para cultivarla en condiciones de laboratorio.
- Cada planta produce varios miles de semillas.
- Tiene un gran número de líneas mutantes.
- Su transformación utilizando *Agrobacterium tumefaciens* para transferir DNA a su genoma es muy eficiente.

3.1. Modelo ABC

Las flores maduras de *Arabidopsis thaliana* presentan una morfología que consiste en cuatro capas concéntricas de órganos. De afuera hacia adentro, éstas son: cuatro sépalos (capa 1), cuatro pétalos (capa 2), seis estambres (capa 3) y dos carpelos fusionados (capa 4). Mutaciones en distintos genes pueden alterar el desarrollo normal de la flor y así, modificar la arquitectura final de la misma.

A partir del análisis de estas mutaciones se propuso un modelo que se ha utilizado ampliamente para describir la morfología de las flores de *Arabidopsis*, tanto silvestres como mutantes. El modelo **ABC** de desarrollo floral de *Arabidopsis* [11] postula la existencia de tres diferentes funciones o actividades genéticas (**A**, **B** y **C**) cuya combinación determina la identidad de los órganos que se desarrollan en las flores. De acuerdo a este modelo, la presencia de **A** determinará la diferenciación de los sépalos, una combinación de **A** y **B** dará como resultado pétalos, mientras que la actividad de tanto **B** como **C** resultará en estambres, y finalmente, la sola presencia de **C** resultará en la formación de carpelos [24] (ver figura 3.1).

Existen varios genes relacionados con las tres actividades antes mencionadas. *APETALA1* (*AP1*) y *APETALA2* (*AP2*) son considerados genes del tipo **A** pues mutaciones inducidas en cada uno de estos genes dieron como resultado flores sin sépalos ni pétalos. *APETALA3* (*AP3*) y *PISTILLATA* (*PI*) son genes del tipo **B** y *AGAMOUS* (*AG*) es el único gen relacionado con la función **C** [24].

El modelo **ABC** resulta sumamente robusto al describir la morfología de plantas tanto mutantes como transgénicas [24]. Además, la estructura estereotípica de las flores de más de 250 000 especies angiospermas (plantas con flor) y los datos moleculares y genéticos disponibles sugieren que el modelo **ABC** se conserva en prácticamente todas las plantas con flor [8] y más particularmente en el grupo de las eudicotiledóneas. La única angiosperma que no conserva la morfología floral típica es una especie endémica de México: *Lacandonia schismatica* [23].

A pesar de que el modelo **ABC** ha servido de guía para numerosos experimentos en diversas especies de plantas, no nos brinda una explicación acerca de la dinámica mediante la cual se alcanzan los patrones estables de expresión genética característicos de los órganos florales ni de las interacciones regulatorias entre genes incluidos en el modelo y otros genes relacionados con el desarrollo floral de *Arabidopsis*.

3.2. Red de regulación genética de desarrollo floral de *Arabidopsis thaliana*

En 1998, Luis Mendoza y Elena Álvarez-Buylla, del Instituto de Ecología de la UNAM, publicaron una red de regulación genética de desarrollo floral de *Arabidopsis thaliana* conformada por 11 genes, cuatro de los cuales habían sido descritos por el modelo **ABC** [24]. Se llevó a cabo un análisis dinámico utilizando un modelo semisíncrono y se obtuvieron 6 atractores. Cuatro de ellos correspondían a patrones de expresión genética observados en las células primordiales de los órganos florales de *Arabidopsis* (sépalos, pétalos, estambres y carpelos) predichos por el modelo **ABC**. El quinto correspondía a células no relacionadas con la floración y el sexto no se encontraba en plantas silvestres pero pudo ser inducido experimentalmente. Análisis posteriores brindaron re-

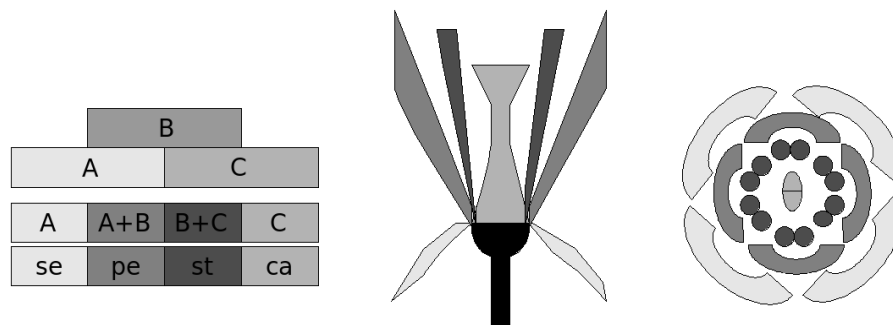


Figura 3.1: El ABC de *Arabidopsis*. Los diferentes tonos de gris representan las actividades combinadas de los genes ABC identificados con los cuatro órganos florales: sépalos (se), pétalos (pe), estambres (st) y carpelos (ca). Los diagramas longitudinal y radial de la flor silvestre de *Arabidopsis thaliana* muestran las cuatro capas concéntricas de órganos florales con los tonos de gris correspondientes.

sultados similares [25].

En 2004, habiendo adquirido un conocimiento más profundo de los componentes genéticos de la red y sus interacciones, se introdujeron nuevos genes conformándose una red de 15 genes [13]. Las reglas lógicas se establecieron a partir de datos experimentales de genética molecular y capturaban información que determinó la base de esta red de regulación. Ocho de los quince nodos de la red se consideraron Booleanos con dos posibles estados: 0 (apagado) y 1 (encendido); mientras que para los otros siete se consideraron tres posibles niveles de expresión: 0, 1 y 2, donde 1 representaba un nivel intermedio de expresión.

Para el análisis dinámico de la red se implementó un modelo discreto síncrono como el descrito en la sección 2.3. Al analizar esta red se obtuvieron atractores coherentes con datos experimentales. En particular, se encontraron diez atractores puntuales: seis correspondían a células primordiales de los cuatro órganos florales y cuatro a células meristémicas que dan lugar a inflorescencias. A partir del análisis dinámico, se pudo incluso hipotetizar acerca de otras interacciones entre los genes de la red aún sin descubrir.

En 2006, se le hicieron modificaciones aleatorias a las funciones lógicas que determinaban las interacciones entre los genes y se demostró que los atractores de la red eran estables ante este tipo de perturbación [9]. Posteriormente, se redujo la red que contenía genes con tres posibles estados a una red Booleana en donde cada gen sólo tenía dos estados posibles (0 y 1). Curiosamente, esta red modificada alcanzó el mismo número y tipo de atractores que la red original y respondió ante perturbaciones de una forma cualitativamente idéntica a la red anterior. Además, algunas de las interacciones predichas en el 2004 fueron corroboradas experimentalmente e incluidas en esta nueva red. También se corrigieron ciertos errores a partir de nuevos descubrimientos y los atractores y

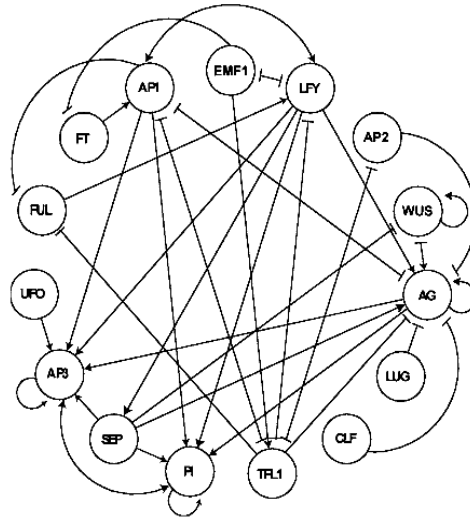


Figura 3.2: Red de regulación de desarrollo floral de *Arabidopsis thaliana*. Los círculos representan a los genes que conforman la red; las flechas indican una regulación positiva y las líneas chatas, una regulación negativa. Reproducido con el consentimiento de E.A-B. y M.A. [9]

sus correspondientes cuencas de atracción se modificaron mínimamente. La red de 15 genes y sus interacciones se muestra en la figura 3.2.

Todos los resultados encontrados a largo de estos años sugieren que la base del modelo **ABC** no se encuentra en las señales específicas entre genes sino en la estructura básica y la dinámica de la red de regulación genética. El objetivo de esta tesis es encontrar más evidencias de este hecho y comprobar que los resultados obtenidos mediante un análisis síncrono se mantienen si se analiza la red con otros modelos.

3.3. Resultados del análisis síncrono de la red

Como se mencionó en la sección anterior, con un análisis síncrono del modelo Booleano, se obtuvieron diez atractores puntuales que corresponden a patrones de expresión genética observados en los primordios de los 4 órganos florales y a células meristémicas de inflorescencias. La configuración de la red, es decir, el estado de cada uno de los 15 genes en los 10 atractores se muestra a continuación. Las columnas corresponden a los 10 atractores puntuales encontrados con el análisis síncrono de la red Booleana del desarrollo floral de *Arabidopsis thaliana*. Los renglones corresponden a los 15 genes que conforman la red y los unos y ceros indican si el gen está activado (1) o desactivado (0) en cada atractor. *TamCA* indica el tamaño de la cuenca de atracción de cada atractor.

	Infl1	Infl2	Infl3	Infl4	Sep	Pe1	Pe2	St1	St2	Car
TFL1	1	1	1	1	0	0	0	0	0	0
LFY	0	0	0	0	1	1	1	1	1	1
FUL	0	0	0	0	0	0	0	1	1	1
AP1	0	0	0	0	1	1	1	0	0	0
AP3	0	0	0	0	0	1	1	1	1	0
PI	0	0	0	0	0	1	1	1	1	1
AG	0	0	0	0	0	0	0	1	1	1
FT	0	0	0	0	1	1	1	1	1	1
EMF1	1	1	1	1	0	0	0	0	0	0
WUS	0	0	1	1	0	0	0	0	0	0
AP2	0	0	0	0	1	1	1	1	1	1
SEP	0	0	0	0	1	1	1	1	1	1
UFO	0	1	1	0	0	1	0	1	0	0
LUG	1	1	1	1	1	1	1	1	1	1
CLF	1	1	1	1	1	1	1	1	1	1
<i>TamCA</i>	128	128	64	64	152	160	8	3744	136	3608

En este marco, podría parecer que haber encontrado atractores de la red que concuerdan con datos experimentales es, digamos, normal. Recordemos que Kauffman propuso en 1969 que los atractores dinámicos de una red de regulación genética podrían identificarse con tipos o destinos celulares [18]. Kauffman recibió un sinnúmero de críticas pues se consideraba que su modelo resulta una aproximación demasiado alejada de la realidad; en primer lugar, por el carácter binario de los genes y la nula consideración en el modelo de la concentración de las proteínas codificadas por los genes. Sí se puede hablar de genes activados o desactivados, pero la acción de la proteína codificada por cada gen depende de la concentración de la primera y ésta aumenta de manera continua. Además, el umbral por encima del cual se considera que la proteína está activada depende de cada gen y el modelo de Kauffman no lo considera. Otra característica del modelo de Kauffman que fue ampliamente criticada es el carácter totalmente síncrono de la dinámica de la red. En los organismos existen diversos factores, como las constantes cinéticas de cada gen y los tiempos de difusión de las proteínas en la célula, que determinan el ritmo de la dinámica. Sin embargo, en el modelo de Kauffman todos los genes se actualizan al mismo tiempo, una vez por cada unidad de tiempo, sin que se consideren tiempos de relajación.

A pesar de todo lo antedicho, en el caso de la red de regulación del desarrollo floral de *Arabidopsis*, los resultados hablan por sí mismos. El modelo de Kauffman es una simplificación de lo que se observa en el fenómeno de regulación genética, pero captura los aspectos esenciales de dicho proceso. Además, la hipótesis de Kauffman de identificar atractores dinámicos con tipos celulares se cumple.

Capítulo 4

Modelos

4.1. Modelos asíncronos

Consideremos una red de N genes donde el estado de expresión de cada gen está representado por una variable Booleana X_i . Generalmente, la dinámica de un modelo Booleano se estudia actualizando simultáneamente el estado de todos los nodos de la red de acuerdo a:

$$X_i(t + \tau) = F_i(X_{i_1}(t), X_{i_2}(t), \dots, X_{i_{k_i}}(t)), \quad i = 1, 2, \dots, N \quad (4.1)$$

donde F_i es la función regulatoria del nodo X_i , como se describe en la sección 2.2.

Una hipótesis subyacente a esta actualización simultánea consiste en la existencia de una sincronización perfecta entre los distintos procesos regulatorios. Sin embargo, se sabe que las escalas de tiempo de los procesos de transcripción, traducción y degradación varían mucho (desde minutos hasta horas) de un gen a otro.

Existen diversos métodos que introducen diferentes escalas de tiempo para los distintos procesos regulatorios propios de la red. Se puede emplear, por ejemplo, un algoritmo completamente asíncrono, en el cual el próximo tiempo de actualización para cada nodo se elige aleatoriamente a cada instante. Otra posibilidad es un algoritmo un tanto más intuitivo que le asigna un tiempo de actualización constante a cada nodo. En este trabajo se utilizarán dos algoritmos asíncronos que llamaremos “algoritmo asíncrono de orden aleatorio secuencial” y “algoritmo asíncrono de orden aleatorio grupal” y que se describen a continuación.

4.1.1. Modelo asíncrono de orden aleatorio secuencial

El algoritmo de orden aleatorio secuencial garantiza que cada nodo se actualice exactamente una sola vez en cada tiempo t . Al inicio del tiempo t y previo a la actualización de los nodos, se elige aleatoriamente una permutación ϕ^t de

$\{1, 2, \dots, N\}$ entre las $N!$ posibles. El orden de las entradas de la ϕ^t elegida determina el orden en que los N genes se van a actualizar en el tiempo t .

A pesar de que en este algoritmo no estamos introduciendo constantes cinéticas para cada nodo, estamos haciendo variar aleatoriamente el intervalo entre dos actualizaciones para cada nodo en cada intervalo de tiempo. El valor de este intervalo puede ir desde $\frac{1}{N}$ hasta $1 + \frac{N-1}{N}$, donde N es el número de nodos de la red y 1 es el intervalo de tiempo en el cual se actualizan todos y cada uno de los nodos una sola vez. El algoritmo garantiza que un mismo nodo no se actualice más de dos veces seguidas antes de que todos los demás nodos se hayan actualizado por lo menos una vez.

4.1.2. Modelo asíncrono de orden aleatorio grupal

El algoritmo de orden aleatorio grupal se deriva del algoritmo de orden aleatorio secuencial y la diferencia radica, como su nombre lo indica, en que la actualización en el algoritmo de orden aleatorio grupal se lleva a cabo en grupos. Todos los nodos que formen parte de un mismo grupo se van a actualizar al mismo tiempo. Al igual que en el modelo anterior, se elige aleatoriamente un orden de actualización ϕ^t al inicio de cada tiempo t . Además de esto se eligen aleatoriamente los tamaños de los grupos de actualización mediante un algoritmo que opera de la siguiente manera: para cada tiempo t se elige aleatoriamente un número $\beta_1^t \in [1, N]$, posteriormente se elige un nuevo número entero $\beta_2^t \in [1, N - \beta_1^t]$ y así sucesivamente se eligen las $\beta_n^t \in [1, N - \beta_1^t - \beta_2^t - \dots - \beta_{n-1}^t]$ hasta que la suma total de los tamaños de los grupos sea N .

De esta manera obtenemos un arreglo $\{\beta_1^t, \beta_2^t, \dots, \beta_n^t\}$ que determina el tamaño de los grupos de actualización simultánea. Los nodos que pertenecen a cada uno de los grupos quedarán determinados por la permutación ϕ^t , es decir que para el tiempo t se van a actualizar simultáneamente, primero los nodos $\{\phi^t(1), \dots, \phi^t(\beta_1^t)\}$, luego los nodos $\{\phi^t(\beta_1^t + 1), \dots, \phi^t(\beta_1^t + \beta_2^t)\}$ y así hasta que se actualicen todos los nodos.

Claramente, la manera en la que se eligen los tamaños de los grupos provoca que sea más probable tener grupos grandes al inicio de cada tiempo t que hacia el final. Sin embargo, este hecho no es relevante pues al elegirse aleatoriamente (para cada tiempo t) los nodos que conformarán estos grupos se evita cualquier recurrencia que pudiera surgir.

4.2. Modelo de *Glass*

Un modelo que nos brinda un puente entre los modelos continuos y los modelos discretos fue propuesto por Leon Glass en 1975. Glass [15] propuso un sistema de ecuaciones diferenciales lineales por pedazos (*“piecewise linear differential equations”*) que combinan las reglas lógicas para la síntesis de proteínas con el decaimiento lineal de las mismas mediante la descripción de cada nodo mediante dos variables: una continua y otra discreta. Sean \widehat{X}_i la variable continua asociada al nodo i , X_i la variable discreta y F_i la función Booleana asociada

al nodo i .

Si desarrollamos $\widehat{X}_i(t + \tau_i)$ hasta el primer orden en potencias de τ_i , obtenemos:

$$\widehat{X}_i(t + \tau_i) \approx \widehat{X}_i(t) + \frac{d\widehat{X}_i(t)}{dt} \tau_i \quad (4.2)$$

Si ahora combinamos la ecuación 4.2 con la ecuación 4.1 del modelo Booleano, se obtiene inmediatamente:

$$\frac{d\widehat{X}_i(t)}{dt} = \alpha_i \left[F_i(X_{i_1}(t), X_{i_2}(t), \dots, X_{i_{k_i}}(t)) - \widehat{X}_i(t) \right], \quad i = 1, 2, \dots, N \quad (4.3)$$

donde $\alpha_i = \frac{1}{\tau_i}$ representa un tiempo de relajación asociado a cada gen, que está relacionado con las vidas medias de las proteínas codificadas por cada gen [10]. Si no se conocen estos tiempos de relajación, se puede modelar la red eligiendo aleatoriamente las α_i y analizando su dinámica.

Por otro lado, sabemos que la función F_i debe estar evaluada en argumentos discretos, pues se trata de la misma F_i de la ecuación 4.1. Por lo tanto, para determinar el valor de la función F_i en la ecuación 4.3, debemos asignarle, en cada tiempo t , un valor discreto a la variable X_i a partir de la variable continua según la ecuación:

$$X_i = H(\widehat{X}_i - \theta_i) \quad (4.4)$$

donde $\theta_i \in (0, 1)$ es el *valor umbral* (“threshold”) asociado al nodo i y H la función de Heaviside (también conocida como función escalón).

Las variables discretas X_i representan el estado discreto de los nodos del modelo Booleano, es decir, si $X_i = 0$ significa que el nodo está apagado y si $X_i = 1$ significa que el nodo está encendido. Como $F_i \in \{0, 1\}$ y si elegimos $\widehat{X}_i(0)$ entre 0 y 1, es evidente que los valores límite para las \widehat{X}_i serán también 0 y 1. En este caso, $\widehat{X}_i = 0$ significa “ausencia del producto del nodo i ” y $\widehat{X}_i = 1$ significa “concentración máxima del producto del nodo i ”.

En cada instante de tiempo y para cada nodo, la dinámica continua se traduce en un estado Booleano de encendido/apagado dependiendo del valor de θ_i : en el instante en que $\widehat{X}_i > \theta_i$, se considerará que el nodo i estará encendido; mientras que esto no suceda, el nodo i se considerará apagado. De esta manera, el parámetro θ_i define la concentración mínima necesaria para que una proteína realice su función biológica. Llamaremos entonces a θ_i el “parámetro de concentración mínima asociado al nodo i ”.

Este método nos permite estudiar la dinámica continua de una red genética simplemente especificando θ_i , sin la necesidad de conocer los parámetros cinéticos específicos para cada gen. En este trabajo se analizarán la dinámica de la red y el espacio de atractores de la misma, variando el parámetro θ_i . Cuando se utilice el mismo valor de θ_i para todos los genes, le llamaremos simplemente θ .

Nótese que los atractores puntuales del sistema dado por (4.3) deben ser iguales a los del modelo síncrono pues

$$\frac{d\widehat{X}_i}{dt} = 0 \iff \widehat{X}_i = F_i(X_{i_1}(t), X_{i_2}(t), \dots, X_{i_{k_i}}(t)), \quad i = 1, 2, \dots, N \quad (4.5)$$

Es decir, que en modelo de *Glass* se deben obtener al menos los mismos atractores puntuales que en el modelo Booleano. Los atractores cíclicos del modelo Booleano (inexistentes para nuestra red) se perderán, en principio, y se podrán obtener nuevos atractores con en el modelo de *Glass*.

Para modelar la red, se resolvió el sistema de ecuaciones 4.3 numéricamente eligiendo los parámetros α_i y θ_i como se describe en la sección 5.3.

4.3. Determinación de atractores

Las formas en que se van a determinar los atractores de la red difieren entre los distintos modelos que vamos a utilizar. Es importante recalcar la manera en que se determinan los atractores en el modelo síncrono. Como se trata de un modelo completamente determinista, se puede conocer el sucesor de cualquier estado inicial. Si conocemos el sucesor de todos los estados iniciales posibles podremos identificar a los atractores, que se definen como estados iniciales cuyos sucesores son ellos mismos (atractores puntuales) o grupos de estados que después de pasar por uno o varios estados intermedios, regresan a ellos mismos (atractores cíclicos).

4.3.1. Atractores en modelos asíncronos

La forma que se eligió para determinar a los atractores en los dos modelos asíncronos fue utilizando el mismo principio de que a cada estado inicial corresponde un estado sucesor, que sirvió como base para encontrar los atractores en el modelo síncrono. Sin embargo, en los modelos asíncronos, el sucesor de cada estado inicial puede ser distinto en cada realización. Para determinar estos sucesores se hizo lo siguiente: en cada realización se le asoció aleatoriamente un orden de actualización ϕ a cada estado inicial. Dada una ϕ , el sucesor de cada estado inicial queda completamente determinado para una realización específica. Una vez conocido el estado sucesor de todos los estados iniciales, las trayectorias y atractores quedan definidos para una realización específica.

Cabe señalar que elegir una ϕ para cada estado inicial y mantenerla fija para la realización en cuestión conlleva ciertas implicaciones. En particular, los atractores cíclicos encontrados serán, en cierta medida, *atractores espurios*. A juzgar por los modelos presentados en las secciones 4.1.1 y 4.1.2, que una trayectoria atraviese dos veces por el mismo estado, no implica necesariamente la existencia de un atractor cíclico, pues al modificarse el orden de actualización ϕ^t para cada tiempo t , la trayectoria saldría del ciclo probablemente al cabo de poco

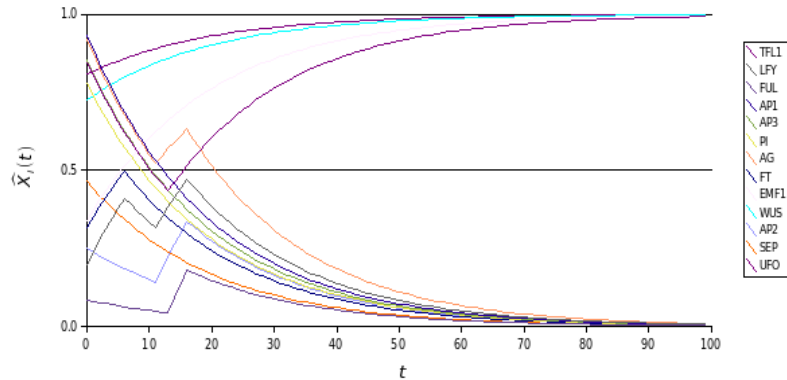


Figura 4.1: Trayectorias de 13 genes partiendo de un estado inicial, con $\theta_i = 0.5$ y $\alpha_i = 0.05$, $\forall i$. En la gráfica se muestra la trayectoria hasta las 100 iteraciones.

tiempo. Es decir, que únicamente deberían encontrarse atractores puntuales en los modelos asíncronos.

Siendo así, tendría sentido no fijarse en estos atractores espurios. Sin embargo, el carácter espurio de estos atractores pasó desapercibido durante la mayor parte de la realización de esta tesis y el análisis que se llevó a cabo, arrojó resultados curiosos que se mencionan en la sección 5.2.1. Por otro lado, el análisis principal de los modelos es en torno a las cuencas de atracción y para este análisis sólo se tomaron en cuenta las realizaciones en las que se obtuvieron únicamente atractores puntuales (ver secciones 5.2.2) siendo irrelevantes los atractores espurios en este caso.

4.3.2. Atractores en modelos de *Glass*

Para determinar los atractores en el modelo de *Glass* se requiere de un proceso completamente distinto pues en este caso no se puede determinar el sucesor de un estado inicial al cabo de un tiempo unitario, ya que la dinámica se lleva a cabo de manera continua. En este modelo se eligen primeramente el número de iteraciones n y el valor de los parámetros α_i y θ_i . Dada un estado inicial $\vec{X}(0)$, se le asigna aleatoriamente un valor $\widehat{X}_i(0)$ a cada nodo, dependiendo del valor del umbral θ_i , es decir, si el estado inicial del nodo i es 0, a $\widehat{X}_i(0)$ se le asignará un valor aleatorio entre 0 y θ_i y si el estado inicial del nodo i es 1, a $\widehat{X}_i(0)$ se le asignará un valor aleatorio entre θ_i y 1. En cada iteración, el valor de cada uno de los nodos cambiará de acuerdo a las ecuaciones 4.3 y 4.4.

En el modelo de *Glass* existen diferencias sustanciales entre los atractores puntuales y los atractores cíclicos. Los atractores puntuales se pueden determinar muy fácilmente si las trayectorias de todos los nodos convergen a 0 o a 1. En la figura 4.1 se muestra un ejemplo típico de cómo se ven las trayectorias de

cada uno de los nodos (genes) en una realización. Las trayectorias muestran el nivel de expresión de cada gen. El estado \widehat{X}_i del gen i cambia, incluso en varias ocasiones, de dirección pero eventualmente se acerca de manera asintótica a 0 o a 1.

Podría también darse el caso de que la trayectoria de uno o más de los nodos adquiriera un carácter ondulatorio periódico. En este caso, deberá verificarse si esta oscilación significa o no un atractor cíclico. Esto dependerá por supuesto del valor de θ_i , pues si al oscilar la trayectoria se modifica el valor de X_i (ver ecuación 4.4), querrá decir que el atractor alcanzado por el estado inicial $\vec{X}(0)$ es un atractor cíclico. Por otro lado, si la oscilación ocurre siempre por arriba o por abajo del valor de θ_i , se tratará de un atractor puntual.

En el modelo de *Glass*, en todas las realizaciones que se llevaron a cabo e independientemente de los parámetros escogidos, nunca se dio el caso de que la trayectoria de un nodo en particular adquiriera un carácter ondulatorio periódico y tampoco se encontraron atractores puntuales diferentes a los del modelo síncrono.

Capítulo 5

Resultados

5.1. Resultados del modelo asíncrono de orden aleatorio secuencial

Para poder realizar un análisis de robustez, se implementó, en primer lugar, el algoritmo asíncrono de orden aleatorio secuencial descrito en la sección 4.1.1. Por simplicidad nos referiremos a este algoritmo o modelo como “asíncrono secuencial” a partir de ahora. Este algoritmo se asegura que cada gen se actualice únicamente una vez durante cada intervalo (unitario) de tiempo. Sin embargo, esta actualización se lleva a cabo gen por gen. Con fines de claridad, pongamos un ejemplo con una red de 3 genes en donde la regla de actualización para cada uno de ellos depende de los otros dos. Al inicio de cada intervalo de tiempo se elige un orden de actualización aleatoriamente. Supongamos que el orden elegido para el tiempo t es $2,0,1$. Esto implica que primero se va a actualizar el gen 2 , tomando en cuenta los estados de los genes 0 y 1 al tiempo $t-1$. El segundo en actualizarse será el gen 0 , sin embargo, éste se actualizará tomando en cuenta el estado del 2 al tiempo t y el estado del gen 1 al tiempo $t-1$. Por último, se actualizará el gen 1 , tomando en cuenta los estados de los genes 0 y 2 al tiempo t . De esta manera, a pesar de que para fines de análisis consideraremos que durante el proceso descrito sólo transcurrió un intervalo de tiempo, la actualización se llevó a cabo en 3 pasos. En el tiempo $t+1$, el orden de actualización se elige de nuevo aleatoriamente y por lo tanto, puede ser distinto al anterior. El estado de la red al tiempo t claramente depende del orden de actualización elegido para ese determinado tiempo.

El análisis de la red del desarrollo floral de *Arabidopsis thaliana* utilizando el algoritmo asíncrono secuencial, arrojó varios resultados que apoyan la tesis de que la red es robusta. Se llevaron a cabo 35000 realizaciones, y los resultados obtenidos de se muestran a continuación.

En primer lugar, el hecho más relevante es que en el 100% de los casos (35000 realizaciones), se obtuvieron todos y cada uno de los 10 atractores obtenidos con el modelo síncrono. A partir de ahora, nos referiremos a estos 10 atractores

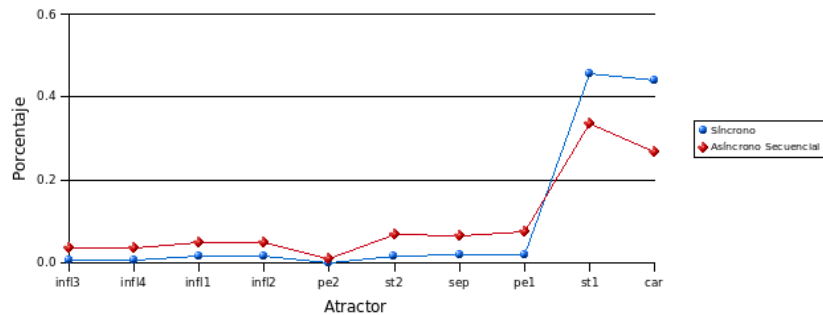


Figura 5.1: Promedio de la fracción del espacio de atractores que ocupa cada cuenca de atracción en 35000 realizaciones.

como los “atractores originales”. Éste hecho forzó a redirigir el análisis hacia otros aspectos.

5.1.1. Cuenca de atracción

A pesar de que se obtuvieron los 10 atractores originales en las 35000 realizaciones, las cuencas de atracción variaban entre realizaciones. Se calculó el tamaño promedio de las cuencas de atracción entre todas las realizaciones y esto evidenció que el tamaño de las cuencas de atracción de Car y St1 (las dos mayores) disminuyó considerablemente entre el modelo síncrono y el asíncrono secuencial. En promedio, la cuenca de St1 disminuyó en un 26.38% y la cuenca de Car disminuyó en un 39.45%. Obviamente, si estas cuencas disminuyeron, debe haber otras que en promedio hayan aumentado de tamaño. Resulta que son todas las demás. En la figura 5.1 se muestra el porcentaje que representa cada cuenca de atracción en ambos modelos.

Para poder comparar más visual y cualitativamente los cambios en las cuencas de atracción entre el modelo síncrono y el asíncrono se graficó un *mapa comparativo de cuencas de atracción* (ver Figura 5.2). Este mapa es un cuadrado de 10x10 y cada cuadro está sombreado en escala de grises. En ambos ejes están los atractores originales de la red. El eje de las abscisas corresponde a los atractores en el modelo síncrono y el eje de las ordenadas a los atractores en el modelo asíncrono secuencial. La escala de grises denota la probabilidad de que un estado inicial que alcanza el atractor X en el modelo síncrono, alcance el atractor Y en el modelo asíncrono secuencial, donde blanco representa 1 y negro, 0. Cada probabilidad está calculada en relación al tamaño de las cuencas de atracción originales, con lo cual, cada columna suma 1. Por ejemplo, el 100% de los estados iniciales que alcanzan el atractor Pe2 en el modelo síncrono, también lo alcanzan en el asíncrono secuencial y por lo tanto, el cuadrado que corresponde a Pe2-Pe2 es completamente blanco. Por otro lado, el 11.33% de los estados que forman parte de la cuenca de atracción de Car en el modelo

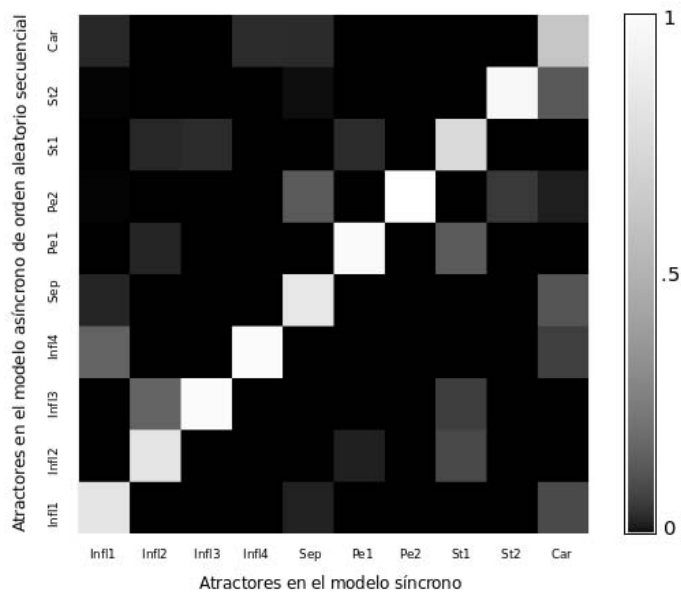


Figura 5.2: Mapa comparativo de cuencas de atracción entre el modelo asíncrono de orden aleatorio secuencial y el modelo síncrono. La escala de grises denota la probabilidad de que un estado inicial que pertenece a una determinada cuenca de atracción en el modelo síncrono, pertenezca a otra (o a la misma) cuenca de atracción en el modelo asíncrono secuencial. Las probabilidades son el promedio de 35000 realizaciones.

síncrono pertenecen a la cuenca de atracción de Sep en el asíncrono secuencial y por lo tanto, el color del cuadrado Car-Sep tiene un tono grisáceo.

En el mapa (Figura 5.2) se puede notar como, en general, las cuencas de atracción no se modificaron drásticamente al modelar la red con el modelo asíncrono secuencial. Si las cuencas de atracción entre ambos modelos fueran idénticas, el mapa mostraría una diagonal completamente blanca y todo lo demás sería negro. En este mapa se puede apreciar perfectamente la diagonal (entre blanca y gris claro) y fuera de la diagonal hay tan sólo algunos cuadros color gris muy oscuro.

Otro aspecto del análisis se puede hacer desde el punto de vista de los estados pertenecientes a las cuencas de atracción en lugar de los atractores. Por ejemplo, más del 27% del total de estados iniciales (2224 de 8192) alcanzan siempre el mismo atractor que alcanzan en el modelo síncrono. Por otro lado, sólo 8 de los 8192 estados iniciales pertenecen, en el 100% de las realizaciones, a una cuenca de atracción diferente a la que pertenecen en el modelo síncrono.

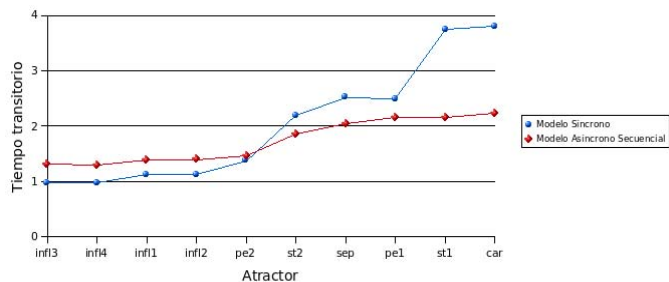


Figura 5.3: Promedio de los tiempos transitorios de todos los estados pertenecientes a cada cuenca de atracción en 35000 realizaciones.

5.1.2. Tiempos transitorios

Se calculó el promedio, entre las 35000 realizaciones, de los tiempos transitorios de todos los estados iniciales pertenecientes a cada cuenca de atracción y de nuevo se hizo la comparación entre el modelo asíncrono secuencial y el modelo síncrono. Los resultados se pueden ver en la Figura 5.3. El promedio de los tiempos transitorios de los estados iniciales pertenecientes a las cuencas St1 y Car es significativamente menor en el modelo asíncrono secuencial. Haciendo un análisis más detallado, resulta que estados pertenecientes a la cuenca de St1 y Car tienen tiempos transitorios de hasta 9 en el modelo síncrono. Sin embargo, en el modelo asíncrono, el tiempo transitorio máximo para estos mismos estados es 5. En el modelo síncrono, la sincronía es una restricción que impone rigidez a la dinámica, mientras que en el modelo asíncrono secuencial la dinámica es menos rígida y probablemente sea ésta la razón por la cual los estados iniciales pueden explorar caminos más cortos para llegar al atractor.

El promedio de los tiempos transitorios de todos los estados en el modelo síncrono es ~ 3.57 , mientras que en el asíncrono secuencial es ~ 1.55 . Estos datos parecerían no concordar con la Figura 5.3, pues de la gráfica podría deducirse que ambos modelos tienen promedios cercanos, sin embargo, esto no es cierto pues cabe resaltar que las cuencas de St1 y Car son mucho más grandes que las demás y por lo tanto, pesan más en el promedio.

5.2. Resultados del modelo asíncrono de orden aleatorio grupal

El modelo asíncrono de orden aleatorio grupal, a partir de ahora, “asíncrono grupal”, descrito en la sección 4.1.2 se propuso como un modelo posible para analizar la red de *Arabidopsis* pues introduce una dinámica que no es ni completamente síncrona ni totalmente asíncrona.

Para analizar la red con el modelo asíncrono grupal se llevaron a cabo 35000 realizaciones. Los resultados son alentadores pues apoyan la tesis de que la red de

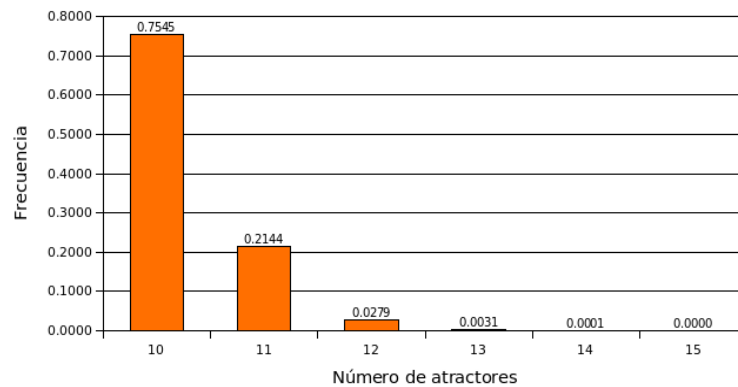


Figura 5.4: Frecuencia de distintos números de atractores encontrados, resultado de 35000 realizaciones con el modelo asíncrono grupal

Arabidopsis es robusta ante alteraciones en los modelos utilizados para estudiar su dinámica. Los resultados obtenidos se presentan a continuación.

5.2.1. Atractores

En el 100 % de las realizaciones se obtuvieron al menos los 10 atractores puntuales originales de la red. De hecho, en poco más del 75 % de las realizaciones se obtuvieron únicamente los 10 atractores originales, pero en el restante 25 % se obtuvieron entre 11 y 14 atractores (ver figura 5.4), habiéndose obtenido 14 atractores en tan sólo 3 de las 35000 realizaciones.

Curiosamente, ninguno de los 9795 atractores nuevos encontrados fue puntual: 9664 fueron de período 2, 121 de período 4, 6 de período 5, 3 de período 6 y 1 de período 7. Resulta aún más extraño que ninguno fue de período 3, pero sí de período 5 y 7. Como se comentó en la sección 4.3.1, estos atractores cíclicos son atractores espurios, sin embargo, esto de ninguna manera justifica que ninguno de los atractores haya sido de período 3.

5.2.2. Cuencas de atracción

El análisis de las cuencas de atracción se hizo únicamente para aquellas realizaciones en que se obtuvieron los 10 atractores originales, es decir en 26409 realizaciones. El análisis que se hizo es exactamente el mismo que el realizado para el modelo asíncrono secuencial descrito en la sección 5.1.1. El resultado se muestra en el mapa comparativo de cuencas de atracción de la figura 5.5. Al igual que en el modelo asíncrono secuencial, los resultados del modelo asíncrono grupal muestran una congruencia fuerte con los resultados del modelo síncrono. La imagen muestra una clara diagonal blanca que indica que la gran mayoría de los estados iniciales alcanzan el mismo atractor en ambos modelos.

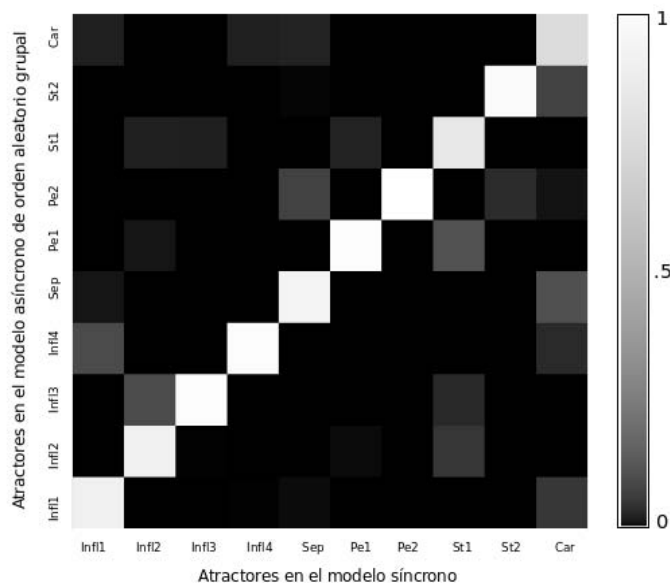


Figura 5.5: Mapa comparativo de cuencas de atracción entre el modelo asíncrono grupal (35000 realizaciones) y el modelo síncrono.

5.3. Resultados del modelo de *Glass*

El modelo de *Glass* difiere sustancialmente de los otros modelos en la determinación de los atractores de la red. Un análisis preliminar de la red utilizando el modelo de *Glass* consiste en asegurarse de que los atractores alcanzados por el sistema sean los mismos que los alcanzados con el modelo síncrono. Como se mencionó en la sección 4.3.2, no se encontraron atractores cíclicos y los atractores puntuales encontrados fueron únicamente los mismos que en el modelo síncrono.

Una vez comprobado que no se alcanzaron atractores nuevos, se analizaron, al igual que en los modelos anteriores, las cuencas de atracción. En el modelo de *Glass*, existen dos parámetros que se pueden variar obteniendo resultados diversos. Los parámetros en cuestión son: θ_i y α_i . Estos parámetros pueden variarse de manera controlada o aleatoriamente para cada gen y/o en cada realización, con lo cual, tenemos una gran gama de análisis posibles.

El análisis de las cuencas de atracción se hizo esencialmente desde dos frentes: variando α_i manteniendo θ_i fija y variando θ_i manteniendo α_i fija. Se utilizarán los términos α y θ cuando α_i y θ_i sean iguales para toda i en una misma simulación.

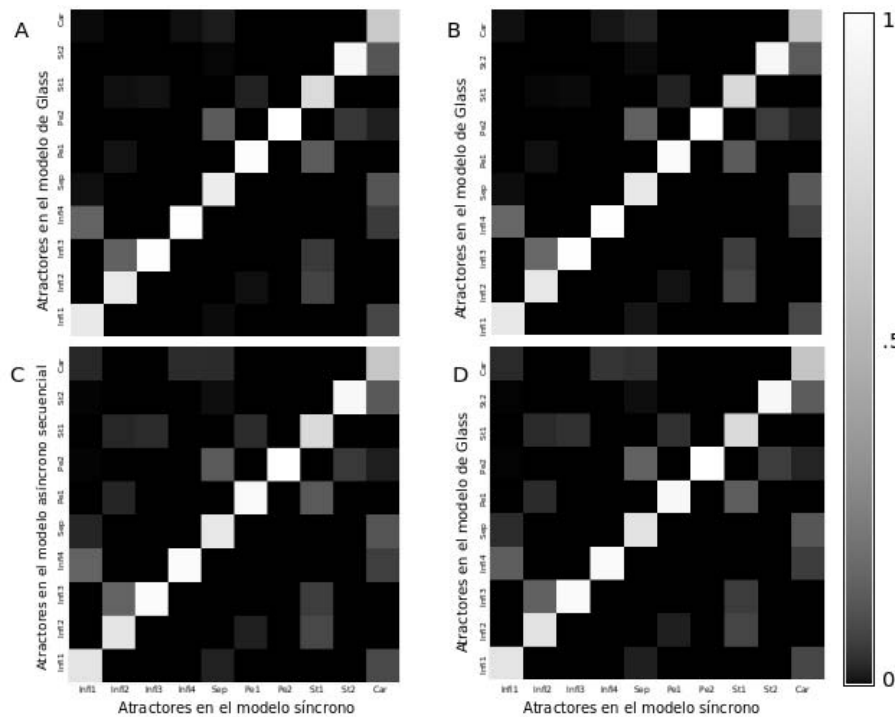


Figura 5.6: Mapas comparativos de las cuencas de atracción obtenidas con el modelo síncrono y las obtenidas con cada uno de los siguientes modelos: A) modelo de *Glass* con $\alpha = 0.05$; B) modelo de *Glass* con $\alpha = 0.005$; C) modelo asíncrono secuencial; D) modelo de *Glass* con tiempos de relajación para cada gen aleatoriamente distribuidos en el intervalo $\alpha \in [0,001, 0,1]$. Nótese la similitud entre todos los casos.

5.3.1. Variación de α_i

Una de las críticas al modelo de Kauffman era que cada gen tiene su propio tiempo de respuesta y de relajación y que éstos no se consideran en el modelo síncrono. Recordemos que $\alpha_i = \frac{1}{\tau_i}$ representa precisamente un tiempo de relajación. En esta sección se comparan los resultados de simulaciones con distintos valores de α (iguales para todos los genes) y una simulación con distintos α_i para cada gen. Para todas las simulaciones se mantuvo $\theta = 0.5$ fija en todas las realizaciones.

Se llevaron a cabo 5 simulaciones con distintos valores de α . Cada simulación constó de 100 realizaciones y cada trayectoria se obtuvo con 5000 iteraciones. Se utilizó el mismo método en todas las simulaciones para determinar los atractores alcanzados (ver sección 4.3.2). Los valores de α elegidos fueron: 0.005, 0.01, 0.025, 0.05, 0.1. No se eligieron valores de α mayores a 0.1 pues en estos casos

las trayectorias decaían inmediatamente. En la figura 5.6 se muestran las mapas comparativos de las cuencas de atracción para $\alpha = 0.5$ (imagen A) y $\alpha = 0.005$ (imagen B). Las diferencias entre estos dos mapas son mínimas, casi indistinguibles, y lo mismo ocurre para las otras 3 simulaciones (mapas comparativos de éstas no se muestran). La conclusión que se obtiene a partir de este análisis es que los resultados obtenidos con el modelo de *Glass* son independientes del valor de α , si éste es menor a 0.1. Además, las diferencias en las cuencas de atracción entre el modelo síncrono y el modelo de *Glass* son muy pequeñas (la diagonal en los mapas comparativos es sumamente clara).

Por otro lado, si comparamos los mapas del modelo de *Glass* con el mapa del modelo asíncrono secuencial (imagen C de la figura 5.6) observamos algo inesperado: las diferencias en las cuencas de atracción entre el modelo síncrono y el modelo asíncrono secuencial son prácticamente idénticas a las diferencias entre el modelo síncrono y el modelo de *Glass*. Esto es un resultado que implica una robustez aún mayor en el modelo, pues a pesar de que las diferencias de ambos modelos respecto al síncrono son completamente distintas, los resultados son los mismos.

Para que quede demostrada la invariancia del modelo frente a cambios en α_i es necesario hacer una simulación en donde se elijan valores de α_i diferentes para cada gen en cada realización. Se le asignó un valor de α_i entre 0.001 y 0.1 aleatoriamente para cada gen al inicio de cada una de las 100 realizaciones. El mapa comparativo de cuencas de atracción se muestra en la imagen D de la figura 5.6. El análisis confirma que los resultados del modelo de *Glass* son independientes del valor de α_i (si son menores a 0.1). Al variar α entre 0.001 y 0.1 se están variando los tiempos de relajación en un factor de 100. Es verdaderamente sorprendente que los resultados no cambien ante tiempos de relajación tan disímiles.

5.3.2. Variación de θ_i

Una vez demostrada la invariancia de los resultados respecto al valor de α se llevaron a cabo análisis variando el valor de θ_i . Para todas las simulaciones presentadas en esta sección se mantuvo fijo el valor de $\alpha = 0.05$ para todos los genes de la red. Cada simulación constó de 100 realizaciones y cada trayectoria se obtuvo con 8000 iteraciones. Se hicieron 4 simulaciones y en cada una se eligieron los θ_i dentro de los siguientes intervalos: $\theta_i \in [0.4, 0.6]$, $\theta_i \in [0.3, 0.7]$, $\theta_i \in [0.2, 0.8]$ y por último $\theta_i \in [0.1, 0.9]$. En cada realización se eligió aleatoriamente un valor de θ_i dentro del intervalo establecido, distinto para cada gen. En la figura 5.7 se muestran los mapas comparativos de cuencas de atracción entre las 4 simulaciones y el modelo síncrono. Las imágenes muestran como, conforme el rango de θ_i aumenta, la diagonal (desde la esquina inferior izquierda hasta la esquina superior derecha) se vuelve cada vez más difusa. Esto implica que la elección de valores de θ_i alejados de 0.5 provoca mayores diferencias entre el modelo de *Glass* y el modelo síncrono. A partir de las imágenes se puede afirmar que el modelo de *Glass* es robusto ante variaciones aleatorias de ± 0.1 respecto a 0.5.

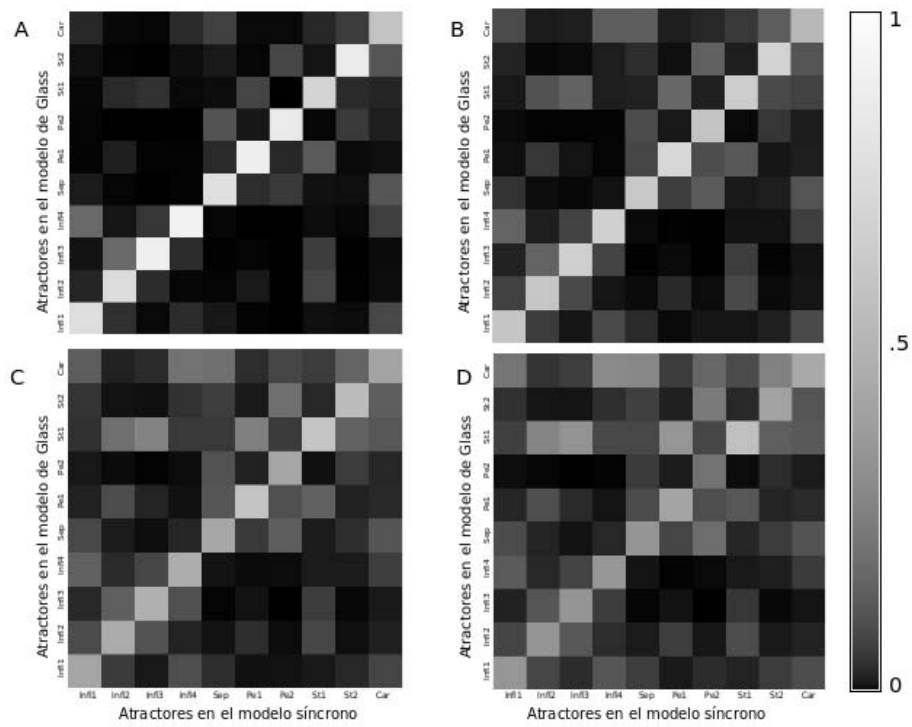


Figura 5.7: Mapas comparativos de las cuencas de atracción obtenidas con el modelo síncrono y las obtenidas con el modelo de *Glass* con θ_i elegidas aleatoriamente para cada gen en los siguientes intervalos: A) $\theta_i \in [0,4,0,6]$; B) $\theta_i \in [0,3,0,7]$; C) $\theta_i \in [0,2,0,8]$; D) $\theta_i \in [0,1,0,9]$.

Capítulo 6

Conclusiones

El modelo Booleano síncrono propuesto por Stuart Kauffman para modelar redes de regulación genética ha probado ser una herramienta útil, pues en varios casos [1, 13, 16, 26], los resultados de las simulaciones han reproducido de manera precisa las observaciones experimentales. En el caso de la red de regulación genética de desarrollo floral de *Arabidopsis thaliana*, con la simulación se obtienen 10 atractores puntuales y éstos corresponden a 10 patrones de expresión observados experimentalmente. Este resultado refuerza la tesis expuesta por Kauffman de que los atractores de la red deben corresponder a destinos o tipos celulares.

Sin embargo, han habido algunas críticas al modelo de Kauffman. Éstas consideran que dicho modelo es muy poco realista desde el punto de vista biológico y se centran en tres aspectos principales: en primer lugar, la sincronía en la actualización de los genes; en segundo lugar, el carácter discreto del estado de los genes; y por último, la nula consideración de que distintos genes tienen distintos tiempos de relajación. La riqueza del modelo de Kauffman radica precisamente en su simplificación. Kauffman [18], al igual que muchos otros autores [15, 28, 2, 7], considera que la dinámica de una red depende de las propiedades generales de la red y no de sus detalles específicos.

En esta tesis se abordaron los tres aspectos principales de crítica al modelo de Kauffman mencionados anteriormente. Se modeló la red de desarrollo floral de *Arabidopsis thaliana*, cuya estructura y reglas lógicas de regulación se obtuvieron a partir de datos experimentales, utilizando 3 diferentes modelos: el “modelo asíncrono de orden aleatorio secuencial”, el “modelo asíncrono de orden aleatorio grupal” y el “modelo de *Glass*”. Los primeros son modelos discretos que introducen dos algoritmos distintos de actualización asíncrona (una secuencial y otra grupal) y el último es un modelo continuo en el cual se introducen distintos tiempos de relajación.

El modelo asíncrono secuencial tiene una dinámica completamente asíncrona en la cual los genes se actualizan uno por uno en orden aleatorio. Con este modelo se obtuvieron los 10 atractores originales de la red en el 100% de las 35000 realizaciones. Las diferencias entre el modelo asíncrono secuencial y el

modelo síncrono radican en el tamaño de las cuencas de atracción y en los tiempos transitorios. Las cuencas de atracción de mayor tamaño y con mayores tiempos transitorios en el modelo síncrono corresponden a los atractores St1 y Car (identificados experimentalmente como los órganos sexuales de la flor). En el modelo asíncrono secuencial, tanto el tamaño de las cuencas de atracción como los tiempos transitorios de ambos atractores fueron menores. Este hecho parece indicar que las trayectorias que siguen muchos de los estados para alcanzar un atractor son más sencillas en el modelo asíncrono secuencial en comparación al modelo síncrono. Sin embargo, en general las cuencas de atracción como tales son muy parecidas entre ambos modelos.

El modelo asíncrono grupal introduce una dinámica semisíncrona, en la cual, se eligen grupos de genes de tamaño y componentes aleatorios que se actualizan síncronamente. Con este modelo se obtuvieron los 10 atractores originales de la red en el 75 % de las 35000 realizaciones. Al comparar las cuencas de atracción de estas realizaciones con las cuencas obtenidas con el modelo asíncrono se observa una concordancia sorprendente. En el restante 25 % también se obtuvieron los 10 atractores, pero además se encontraron otros atractores. De éstos ninguno es un atractor puntual ni de período 3 cuando *a priori* no hay ninguna razón por lo que esto deba ocurrir. Por otro lado, ninguno de los nuevos atractores encontrados se consideró relevante pues ninguno apareció con la suficiente frecuencia como para destacar entre los demás.

Por último, el modelo de *Glass* introduce una dinámica continua, utilizando el método de Euler para encontrar el valor continuo de una variable que denota la concentración de la proteína codificada por cada gen. Sin embargo, la función de actualización es la misma función Booleana utilizada en el modelo síncrono y por lo tanto, depende de los estados discretos de cada gen. Éstos se determinan al comprobar si la variable continua se encuentra por arriba (1) o por abajo (0) de un umbral determinado en cada instante de tiempo. De manera que este modelo permite la variación de dos principales parámetros que hemos denominado α_i y θ_i y que corresponden a los tiempos de relajación y al valor de los umbrales para cada gen.

Por construcción, con el modelo de *Glass* se deben obtener los mismos atractores puntuales que en modelo síncrono. Sin embargo, no deben ser forzosamente los únicos atractores ni debe necesariamente haber una semejanza en las cuencas de atracción. El análisis mostró que no se obtiene ningún atractor cíclico sino únicamente los 10 atractores puntuales originales y que las cuencas de atracción resultan ser similares a las del modelo síncrono y sumamente parecidas a las del modelo asíncrono secuencial. Los resultados obtenidos con el modelo de *Glass* resultaron ser invariantes ante diferencias en los tiempos de relajación hasta en un factor de 100 entre distintos genes y robustos ante variaciones aleatorias de ± 0.1 respecto a 0.5 en el valor de los umbrales de cada gen.

En general, el análisis de la red mediante los 3 modelos brindó resultados muy similares a los obtenidos con el modelo síncrono propuesto por Kauffman. Estos resultados refuerzan la hipótesis de que es la estructura básica de la red de regulación la que determina la dinámica y contribuyen a anular las críticas al modelo de Kauffman en relación a la sincronía, los estados discretos de los genes

y los tiempos de relajación. Un método de modelaje puede ser elegido en lugar de otro dependiendo de las ventajas o desventajas que ofrezca, pero a partir de los resultados encontrados en esta tesis, se puede afirmar que en particular para la red de desarrollo floral de *Arabidopsis thaliana* el modelo síncrono Booleano propuesto por Kauffman es una excelente herramienta. Es de esperarse, claro está, que estos resultados sean similares para otras redes de regulación genética encontradas en la naturaleza.

Es importante entender que la labor de los científicos es llevar nuestro entendimiento de la naturaleza lo más profundo posible no sólo para nuestra propia satisfacción, sino para poder entablar una relación mucho más íntima y fortalecedora con ella y así poder contribuir a que prevalezca. La física, las matemáticas y la computación han resultado ser herramientas enormemente útiles para describir ciertos aspectos de la naturaleza, pero en particular la vida ha probado ser un terreno mucho más complejo de lo que podemos describir a través de puras ecuaciones.

Kauffman ha sido criticado porque muchos consideran que su acercamiento al estudio de la vida, la denigra, pues la sobresimplifica y le roba parte de su riqueza. En mi opinión, es todo lo contrario: la riqueza de la vida radica justamente en la inmensidad de posibilidades que surgen de relaciones que en principio son sencillas. Nuestro acercamiento al estudio de la vida debe ser siempre teniendo en cuenta que nuestras herramientas servirán para estudiarla y comprenderla hasta cierto punto. Lejos de intentar explicar fenómenos biológicos a través de fundamentos físicos, esta tesis representa una expresión del diálogo entre ambas disciplinas; un diálogo encaminado a mejorar el entendimiento del orden y el desarrollo en los seres vivos.

Bibliografía

- [1] Albert, R. & Othmer H.G. (2003) The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *J. Theor. Biol.* **223**, 1-18.
- [2] Albert, R. & Othmer H.G. (2003) ...But No Kinetic Details Needed. *SIAM News* **36**.
- [3] Albert, R. (2004) Boolean Modeling of Genetic Regulatory Networks. *Lect. Notes Phys.* **650**, 459-481.
- [4] Aldana, M., Balleza, E., Kauffman, S. & Resendiz, O. Robustness and evolvability in genetic regulatory networks. *J. Theor. Biol.* **245**, 433-448.
- [5] Balleza, E., Alvarez-Buylla, E., Chaos, A., Kauffman, S., Shmulevich, I. & Aldana, M. (2008) Critical Dynamics in Genetic Regulatory Networks: Examples from Four Kingdoms. *PLoS ONE* **3**, e2456.
- [6] Bennett, M. D., Leitch, I. J., Price, H. J., & Johnston, J. S. (2003) Comparisons with *Caenorhabditis* (100 Mb) and *Drosophila* (175 Mb) Using Flow Cytometry Show Genome Size in *Arabidopsis* to be 157 Mb and thus 25% Larger than the *Arabidopsis* Genome Initiative Estimate of 125 Mb. *Annals of Botany* **91**, 547-557.
- [7] Bornholdt, S. (2005) Less Is More in Modeling Large Genetic Networks. *Science* **310**, 449-451.
- [8] Bowman, J.L. (1997) Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. *J. Biosci.* **22**, 515-527.
- [9] Chaos, A., Aldana, M., Espinosa-Soto, C., García Ponce de León, B., Garay Arroyo, A. & Alvarez-Buylla, E.R. (2006) From Genes to Flower Patterns and Evolution: Dynamic Models of Gene Regulatory Networks. *J. Plant Growth Regul.* **25**, 278-289.
- [10] Chaves, M., Sontag, E.D. & Albert, R. (2006) Methods of robustness analysis for Boolean models of gene control networks. *IEE Proc.-Syst. Biol.* **153**, 154-167.

- [11] Coen, E.S. & Meyerowitz, E.M. (1991) The war of the whorls: genetic interactions controlling flower development. *Nature* **353**, 3-37.
- [12] De Jong, H. (2002) Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. *Jour. Comput. Biol.* **9**, 67-103.
- [13] Espinosa-Soto, C., Padilla-Longoria, P. & Alvarez-Buylla, E.R. (2004) A Gene Regulatory Network Model for Cell-Fate Determination during *Arabidopsis thaliana* Flower Development That Is Robust and Recovers Experimental Gene Expression Profiles. *The Plant Cell* **16**, 2923-2939.
- [14] Gibson, M.A. & Mjolsness, E. (2001) Modeling the Activity of Single Genes. *Computational Modeling of Genetic and Biochemical Networks*. MIT Press, Cambridge, USA, 3-48.
- [15] Glass, L. (1975) Classification of biological networks by their qualitative dynamics. *J. Theor. Biol.* **54**, 85-107.
- [16] Huang, S. & Ingber, D.E. (2000) Shape-Dependent Control of Cell Growth, Differentiation, and Apoptosis: Switching between Attractors in Cell Regulatory Networks. *Exp. Cell Res.* **261**, 91-103.
- [17] Integr8 - A. thaliana Genome Statistics: "<http://www.ebi.ac.uk/integr8/OrganismStatsAction.do;jsessionid=08E9058B5B688A4F7FF7D161CB9E36A4?orgProteomeId=3>", 13 de mayo de 2008.
- [18] Kauffman, S.A. (1969) Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets. *J. Theor. Biol.* **22**, 437-467.
- [19] Kauffman, S.A. (1993) *The Origins of Order. Self-Organization and Selection in Evolution*. New York, Oxford University Press.
- [20] Kauffman, S.A. (1995) *At Home in the Universe. The Search for Laws of Self-Organization and Complexity*. New York, Oxford University Press.
- [21] Luscombe, N.M., Babu, M.M., Yu, H., Snyder, M., Teichmann, S.A. & Gerstein, M. (2004) Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* **431**, 308-312.
- [22] Makita, Y., Nakao, M., Ogasawara, N. & Nakai, K. (2004) DBTBS: database of transcriptional regulation in *Bacillus subtilis* and its contribution to comparative genomics. *Nucleic Acids Res.* **32**, Database issue D75-D77.
- [23] Martinez, E. & Ramos, C.H. (1989) Lacandoniaceae (Triuridales): Una nueva familia de México. *Ann. M. Bot. Gard.* **76**, 128-135.
- [24] Mendoza, L. & Alvarez-Buylla, E.R. (1998) Dynamics of the Genetic Regulatory Network for *Arabidopsis thaliana* Flower Morphogenesis. *J. Theor. Biol.* **193**, 307-319.

- [25] Mendoza, L., Thieffry, D. & Alvarez-Buylla, E.R. (1999) Genetic control of flower morphogenesis in *Arabidopsis thaliana*: A logical analysis. *Bioinformatics* **15**, 593-606.
- [26] Mendoza, L. & Alvarez-Buylla, E.R. (2000) Genetic Regulation of Root Hair Development in *Arabidopsis thaliana*: A Network Model. *J. Theor. Biol.* **204**, 311-326.
- [27] Monod, J. (1975) *El azar y la necesidad. Ensayo sobre la filosofía natural de la biología moderna.* 7 edición. Barral Editores. Barcelona.
- [28] Monod, J. & Jacob, F. (1961) General conclusions: Telenomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 389-401.
- [29] Salgado, H., Gama-Castro, S., Peralta-Gil, M., Díaz-Peredo, E., Sánchez-Solano, F., Santos-Zavaleta, A., Martínez-Flores, I., Jiménez-Jacinto, V., Bonavides-Martínez, C., Segura-Salazar, J., *et al.* (2006) RegulonDB (version 5.0): *Escherichia coli* K-12 transcriptional regulatory network, operon organization, and growth conditions. *Nucleic Acids Res.* **34**, Database issue D394-D397.
- [30] Schrödinger, E. (2006) *What is Life?* Cambridge, UK, Cambridge University Press.
- [31] Smolen, P., Baxter, D.A. & Byrne, J.H. (2000) Modeling Transcriptional Control in Gene Networks-Methods, Recent Results, and Future Directions. *Bull. of Math. Biol.* **62**, 247-292.
- [32] The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.