



Importancia biológica y evolutiva de la estructura y de la dinámica de redes genéticas pequeñas: modelos de simulación inspirados en redes reales.

Tesis que presenta

Álvaro Chaos Cador

para obtener el grado de Doctor en Ciencias

México, D. F. 2010



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.

«Jaio zera, ilko zera»

Agradezco:

Especialmente a Elena Álvarez-Buylla su apoyo mientras las fechas para iniciar el doctorado llegaban, el haber dirigido esta tesis y haber sido mi tutora durante el tiempo en que realicé el doctorado en su laboratorio. Espero que nuestras discusiones entre los azares y los destinos produzcan otros trabajos en conjunto.

A los miembros de mi comité tutor, Pedro Miramontes y Daniel Piñero, con quienes semestre a semestre discutí mi proyecto, les agradezco sus comentarios y sugerencias durante esas sesiones, así como también el haber revisado el borrador final de este trabajo y de ser sinodales en el examen.

A Gerardo Escalera y a Pablo Padilla por las discusiones interesantes y amenas fortificadas con un cafecito y alguna que otra golosina. También les agradezco el haber revisado el borrador final de este trabajo y conformar parte de los sinodales del examen.

A Max Aldana y a Octavio Miramontes el haber accedido a ser parte del jurado. Sus críticas y observaciones hicieron que la versión final de este trabajo resultara mucho mejor.

A la Universidad Nacional Autónoma de México y a su propiedad emergente.

A CONACyT por el apoyo económico durante el tiempo en el que realicé el doctorado.

Al Doctorado en Ciencias Biomédicas por el apoyo económico para la impresión de la tesis.

Al caos.

Índice

Resumen.....	1
Introducción.....	2
Capítulo 1.....	3
Errores, adaptaciones y suertes.	
Los morfoespacios: las playas de la variación	
El reduccionismo: «modus operandi» genético.	
El ABC de las redes	
Las propuestas kaufmanianas: la vida al borde del caos	
La decisión de lambda	
La flor enredada	
Capítulo 2.....	31
De los genes a las formas florales y a la evolución: modelos dinámicos de redes genéticas reguladoras	
Capítulo 3.....	47
Morfogénesis floral: exploraciones estocásticas del paisaje epigenético de una red genética.	
Capítulo 4.....	63
Conclusiones generales y perspectivas	
Apéndices.....	67
Apéndice 1 «Vientos del pueblo»	
Apéndice 2 «ATALIA»	
Referencias generales.....	79

Resumen

Tomando como modelo a las redes genéticas binarias, este trabajo se centra en el estudio dos aspectos principales: la robustez de estas redes y la transición de un sumidero a otro sin recurrir a señales deterministas. Utiliza como modelo real la red de diferenciación floral de la arábide (*Arabidopsis thaliana* (L.) Heynh.).

Con respecto al primer aspecto, se tomó la red floral de *A. thaliana* para realizar pruebas de robustez. Primero se actualizaron, de acuerdo con la evidencia experimental reciente, las relaciones topológicas y las reglas lógicas de los genes. El resultado fue que se cambió una relación topológica y una regla lógica. El modelo original consistía de nodos binarios y de nodos con tres estados. Se tradujo esta red mixta a una red cuyos nodos fuesen exclusivamente binarios. Los resultados obtenidos sobre la dinámica del sistema (sumideros) coincidieron exactamente con los de la red original, que a su vez coinciden con los perfiles genéticos observados en las células que originarán a los verticilos florales en *A. thaliana*. Por último, se realizó un análisis que consistió en mutar la salida de cada una de las combinaciones de todas las reglas lógicas de cada uno de los genes para registrar la robustez del sistema con relación a las mutaciones. Los resultados mostraron que la red es extremadamente robusta y que la mayoría de las mutaciones son neutras (83.76%).

Dado que los modelos de redes genéticas discretas son deterministas, existe el problema de explicar cómo es que se transita entre los diferentes sumideros. Una idea para resolver dicho problema consiste en proponer la entrada de señales de tipo determinista en alguno de los genes de la red, es decir, que cierto gen reciba la señal de cambio de estado y que lo obligue a irse a otra cuenca de atracción, en la cual finalmente se llegará a otro sumidero. Sin embargo, en este trabajo se explora cómo el ruido, entendido como cualquier discrepancia de las reglas originales debido a factores intrínsecos o extrínsecos, puede generar las secuencias que se observan en la naturaleza sobre la diferenciación de los verticilos florales. Se aborda el problema con dos modelos, el modelo de redes discretas binarias y la aproximación de Glass para un modelo continuo. Se encontró que dentro de un intervalo de ruido, la secuencia que se observa se acerca mucho a la natural para los dos modelos, y que bajo un cierto nivel de ruido se recupera la secuencia natural con el segundo modelo.

Introducción

El descubrimiento del proceso evolutivo y, posteriormente, el surgimiento de una teoría que lo explique, han sido dos de los eventos más importantes en la historia de la biología por medio de los cuales esta ciencia toma un carácter maduro y de unidad. La teoría evolutiva que se utiliza normalmente y que se enseña en las universidades está conformada por muchos elementos provenientes de diferentes quehaceres biológicos, que, por cierto, en varias etapas durante la construcción de la teoría estuvieron en corrientes de pensamiento rivales y enemigas. Dicha teoría se denomina síntesis moderna o síntesis nueva.

Desde el reconocimiento de que las especies no son entidades inmutables hasta las últimas evidencias sobre las cuales se estructuró la síntesis moderna, hay una cantidad enorme de científicos y de evidencias que sería imposible hablar de todas ellas aquí. Sin embargo, es importante al menos mencionar algunos de los eventos trascendentales, que representaron un papel importante en la confección de esta teoría evolutiva, así se podrá entender sus fortalezas y sus debilidades para proponer las partes que deberán estar sujetas a revisión, ya sea para perfeccionar aquellas secciones problemáticas o en caso contrario construir un paradigma evolutivo nuevo.

En este trabajo se pretende estudiar una de las lagunas de la síntesis moderna, que es la liga entre los ámbitos genético y morfológico, y se propone que este hueco conceptual puede ser entendido mediante un enfoque integrador valiéndose de redes genéticas binarias. Para ello se presenta una semblanza brevísima sobre la edificación de la síntesis moderna en donde se destacan sus problemas, tanto teóricos como metodológicos; posteriormente se plantea de qué forma un análisis integral puede ayudar a entender y a alumbrar estas partes oscuras de la evolución para las cuales la síntesis moderna está ciega.

Capítulo 1

Errores, adaptaciones y suertes.

El año de 1492 no sólo es un parte aguas en la historia del hombre, sino también en la biológica. En los años venideros la Corona Española envió muchas expediciones científicas a las tierras recién descubiertas e ignotas, como producto de estos viajes de exploración se generaron colecciones fáunicas, florísticas, antropológicas, geográficas, culturales, entre otras, sin precedentes. Cabe resaltar aquí un par de expediciones, la de Miguel Hernández, «el preguntador», y la de Mariano Mociño y Martín Sessé. La expedición de Hernández se destaca por ser la primera expedición científica realizada. Sus tintes de planificación, de reclutamiento, de realización y de análisis inauguran la era de las expediciones científicas modernas. La segunda, la de Mociño y Sessé, además de ser una de las más grandes y mejor realizadas nunca, fue la que dio a conocer la biodiversidad americana en Europa y gracias a esto la concepción biogeográfica cambió rotundamente (Lozoya 1984). José de Acosta, quien anticipó la existencia del estrecho de Bering antes de su descubrimiento, y Buffon, por ejemplo, empezaron a sostener que las especies no eran entidades fijas, sino que por el contrario debían cambiar mientras migraban, ya que de otra forma era imposible explicar el origen de algunos grupos americanos. La idea de que las especies eran inmutables se desmoronaba.

Antes del descubrimiento de América se pensaba que la distribución geográfica de los seres dependía de sus capacidades de dispersión. Después de analizar la diversidad americana, los naturalistas empezaron a poner en duda la idea de que las especies eran colonizadoras activas. La distribución geográfica de las especies respondía a otros procesos, al parecer, a que se creaban en donde habitaban.

Con estos antecedentes ya era posible pensar en una teoría evolutiva. Lamarck, en *Filosofía zoológica (Philosophie zoologique 1809)* fue el primero en proponer un mecanismo de cambio que explicara la diversidad enorme de criaturas que apreciamos en el mundo. De manera general se basó en ciertas leyes, como la de que los organismos poseían una fuerza innata al cambio, y la del uso y desuso de los caracteres y la herencia de los caracteres adquiridos (Lamarck 1963). La primera proponía que los organismos tenían un impulso natural o tendencia hacia el cambio, es decir, que inherentemente estaban propensos a evolucionar. La segunda explicaba, según la inercia de la primera, que los organismos se transformaban. Si ocurría que cierta característica le proporcionaba una ventaja al individuo ésta se desarrollaría más en él y además la heredaría a su descendencia, la cual podía continuar el desarrollo de esa característica si las condiciones que generaron su desarrollo se mantenían. El ejemplo típico de esta transformación se personifica con la jirafa y su cuello. El lamarquismo propone que las variaciones que observamos en el mundo biológico son adaptaciones porque responden a una necesidad del organismo. Recordemos que bajo las ideas esencialistas griegas las variaciones eran errores, ya que se desviaban de la esencia de la especie en el mundo ideal. Bajo ambos enfoques la utilidad de las variaciones estaba determinada *a priori*, serán errores o adaptaciones, la variación estaba predestinada.

En 1859, Carlos Darwin publica el libro *Sobre el origen de las especies* (*On the origin of species*), en el cual amplía la idea lamarquiana de transformación introduciendo un elemento nuevo: la selección natural. Darwin continúa con la propuesta de Lamarck de que las variaciones son adaptaciones, porque está de acuerdo en la ley del uso y desuso de los caracteres y la herencia de los caracteres adquiridos; sin embargo, propone otra fuente de variación. Para él la variación podía ser azarosa. El sentido del azar no significaba que las variaciones provenían de un evento probabilista, las variaciones surgían por procesos bien definidos aunque desconocidos para el hombre; no obstante, la creación de cierta variación podía no estar ligada a una utilidad preconcebida en el mundo real, es decir, una variación podría ser beneficiosa, pero también perjudicial. El darwinismo propone, al igual que el lamarquismo, que una parte de las variaciones serán adaptaciones, pero otra parte de las variaciones estarán desligadas de las necesidades del organismo. Esta parte de las variaciones podrá ser beneficiosa, perjudicial o neutra.

Por otro lado, es fundamental recalcar que la aparición del libro de Darwin no generó un avance general en la teoría evolutiva. La propuesta de la idea de la selección natural como fuerza de evolución sin duda proporcionó un elemento sólido a la teoría evolutiva; sin embargo, propició un atraso descomunal en la biogeografía regresando a esta ciencia al modelo viejo de la dispersión, retrasando el desarrollo de ella, siendo muy benévolo, por al menos cien años. José de Acosta, y posteriormente, Buffon y Alfonso de Candolle, habían desechado la idea de que la causa de la distribución geográfica de los organismos se debía a sus capacidades de dispersión (Acosta 1962, Nelson 1978). Con ello la biogeografía tomaba un rumbo nuevo, un paradigma distinto. Ahora el centro de estudio eran las áreas de endemismo y sus relaciones históricas. Darwin, en los capítulos 11 y 12 de *Sobre el origen de las especies*, retoma la idea antiquísima sobre la importancia que tiene la dispersión en la conformación de las distribuciones geográficas de los organismos. Debido a la aceptación grande de las ideas darwinistas en materia evolutiva, se regresó a estudiar la dispersión y los centros de origen como factores principales de la distribución geográfica de las criaturas. El programa científico en biogeografía histórica consistió, durante finales del siglo XIX y mediados del XX, en la búsqueda de los centros de origen de las especies y de los caminos que habían seguido para conformar sus distribuciones geográficas. El error de Darwin consistió en creer que los organismos cosmopolitas y dominantes son la regla de la naturaleza, cuando en realidad representan la excepción. El modelo dispersionista se asimiló dentro de la síntesis nueva, sobre todo en los modos de especiación. Dentro de las clases de especiación que reconoce, se encuentra la especiación geográfica (además de la ecológica y la genética), tipo de especiación que está apoyada por una cantidad muy grande de datos que provienen de una gama amplísima de criaturas (Huxley 1963).

A inicios de la mitad segunda del siglo XX, el trabajo de León Croizat, inventor de la pambiogeografía, refutó las ideas darwinianas sobre la evolución de las especies en el espacio geográfico (Craw *et al.* 1999). Croizat se centró en los postulados fundamentales de la biogeografía de la dispersión para ponerlos a prueba. El esencial establecía que la distribución geográfica de las criaturas respondía a sus capacidades de dispersión, puesto así, se esperaba que las especies con medios más poderosos para dispersarse poseyeran distribuciones geográficas más amplias que aquellas con esta capacidad limitada. El resultado de superponer muchas distribuciones de organismos con medios y poderes de dispersión heterogéneos fue contra la intuición y contra el núcleo teórico de la biogeografía dispersionista: las áreas de distribución de los seres no dependían de sus medio para dispersarse. Este hecho, junto con el descubrimiento de formas geográficas recurrentes que

resultaban de unir a las distribuciones de los organismos emparentados, refutaron a la dispersión como la causa principal de la distribución geográfica de la biodiversidad. Así, se retomó el camino guiado por los resultados de las grandes expediciones a América, en el cual la dispersión no era el eje de la explicación biogeográfica, abandonado hacía más de un siglo.

Es intrigante que las ideas erróneas y desmentidas contundentemente sobre la evolución de los organismos en el espacio, una de las bases de la teoría de Darwin, sean todavía bases de la síntesis moderna (Colacino y Grehan 2003).

La teoría darwinista presentaba problemas innumerables, de los más importantes eran las ideas lamarquianas, la herencia y la especiación con su componente espacial. La herencia era uno de los puntos más débiles. Darwin apoyaba a la herencia mezclada (Vorzimmer 1963), lo cual es extraño porque bajo ese tipo de herencia es difícil que la selección natural pudiese funcionar, fuerza que él proponía como motor evolutivo. Pocos años después de la publicación de *Sobre el origen de las especies*, Fleeming Jeremy comprobó matemáticamente que no era posible que la selección natural funcionara bajo este tipo de herencia (Depew y Weber 1995).

El desconocimiento que tenía Darwin de la herencia, específicamente de las leyes mendelianas, evitó que, además de ser incapaz de contestar a la crítica de Jeremy, mencionada anteriormente, entendiera la base genética sobre la cual actuaba la selección natural y que no descartara los efectos «del uso y del desuso» (Dobzhansky et al. 1983). Los experimentos de A. Weissman lograron eliminar las ideas lamarquianas del darwinismo, creando lo que se denominó como neodarwinismo (Mayr 1991). El redescubrimiento de las leyes mendelianas produjo algo inimaginable, ya que en lugar de integrarse rápidamente al cuerpo de la teoría de evolución por selección natural y darle el sustento genético del cual carecía, originó una corriente nueva. No fue sino hasta 1937, con la publicación del libro *Genética y el origen de las especies* (*Genetics and the origin of species*) de T. Dobzhansky, que se logró conciliar, gracias a R. A. Fisher, J. B. J. Haldane y S. Wright, el neodarwinismo con la genética mendeliana, naciendo la síntesis moderna (Huxley 1963). Posteriormente esta teoría se ha ampliado con conocimientos provenientes de otras ramas de la biología, como por ejemplo, de la paleontología (Simpson 1944), de la sistemática (Mayr 1942) y con el descubrimiento del ADN.

Si se hace una recapitulación de los elementos más importantes que contribuyeron a la formación de la síntesis moderna se podría decir que fueron la selección natural, la genética mendeliana, el germoplasma y la genética de poblaciones. De la misma forma, como esquematiza S. J. Gould en su libro *La estructura de la teoría evolutiva* (*The structure of evolutionary theory* 2002), los tres pilares de la teoría son, en primer lugar, que la fuerza más importante de evolución es la selección natural, en segundo, que la selección natural es una fuerza creativa, y por último, que la macro-evolución es el resultado de la micro-evolución acumulada.

La teoría del equilibrio puntuado o de los equilibrios intermitentes y la teoría neutralista de evolución fueron en un principio teorías rivales de la síntesis moderna ya que parecían poner en entredicho alguno de sus pilares fundamentales, en particular, el que asegura que la selección natural es la fuerza más importante de evolución, cada teoría lo atacó desde su campo de estudio, siendo el paleontológico propio de la primera y el molecular de la segunda. Sin embargo, a la larga, ambas teorías perdieron este ímpetu revolucionario convirtiéndose en críticas de menor rango y para casos específicos.

Curiosamente los rivales más vituperados, como son las ideas de la ortogénesis de Croizat (Croizat 1977, Colacino y Grehan 2003), los saltarines evolutivos, mejor conocidos como los monstruos esperanzados de Goldschmidt, aparecidos en su libro *La base material de la evolución* (*The material basis of evolution* 1940) y la macroevolución, término acuñado por Y. Filipchenko en *Variabilidad y variación* (*Variabilität und variation* 1927), por citar algunos, se han omitido, se tocan someramente o han sido malinterpretados por la ciencia anglosajona como se puede apreciar si se revisan los libros de texto universitarios más conocidos sobre evolución: *Evolutionary biology* de D. Futuyma (1998), *Evolution* de M. Strickberger (2000) y *Evolution* de M. Ridley (2004), o si se revisa el temario de la materia Evolución I de nuestra universidad.

Para ciertos investigadores, entre los que se incluyen a R. Dawkins y al mismo Darwin, la teoría evolutiva debe de explicar preponderantemente las adaptaciones, para otros, como Gould, el fenómeno más importante que debe tratar es la extinción (Sterelny 2001). Este tipo de diferencias se originan debido a que el campo de trabajo de cada uno de ellos es muy diferente y perciben el cambio de la vida con este sesgo particular. De todas formas, de manera general, la síntesis moderna debe de proporcionar el marco teórico con el cual se debe de explicar cómo evolucionan los seres.

Los orígenes de estas disputas o críticas profundas tienen su raíz en una disyuntiva de apreciación sobre el carácter de la vida: la visión interna y la visión externa de la vida. La primera percibe a los organismos como entidades sujetas a fuerzas internas de orden. En el proceso del desarrollo de un organismo se observa que las diferentes fases del mismo se van sucediendo de acuerdo a interacciones entre diversas sustancias, tejidos y formas que surgen de acuerdo a un plan determinado. Esta secuencia de eventos que se va dando durante el desarrollo de un ser finaliza de forma general, cuando se llega a la fase adulta. Como es de esperarse, su duración dependerá del lapso de maduración de la criatura en cuestión, pero no es un proceso largo en términos históricos, por lo cual se deben de buscar sus causas en procesos no muy lejanos. El desarrollo se basa en el estudio de estas causas próximas.

La visión darwiniana¹ desliga a las criaturas de un orden divino preconcebido y estático. Por medio de la selección natural se explica la adaptación de las criaturas sin recurrir a disposiciones previas de diseño, como argumentaba la teología natural, e imprime un carácter histórico a las formas biológicas. Dicha fuerza, descubierta por Darwin y Wallace (1858), impone un planteamiento distinto. Los organismos se ven como formas acabadas constituyendo agregados que se exponen a la fuerza de la selección natural en un proceso adaptativo preponderantemente. Este ángulo es la visión externa de la vida. La selección natural causa la adaptación, que en esencia, por muy rápida que sea, se trata de un proceso de tiempos históricos mucho más largos que los tiempos del desarrollo. Sus causas deben de ser más lejanas. La evolución se basa en el estudio de estas causas lejanas.

Como sumario, cada visión persigue causas diferentes y se ajusta a un programa de investigación particular ponderando exageradamente su perspectiva sobre la otra, lo que causa la separación del cuerpo teórico de cada visión. Una de ellas se enfoca en el estudio de la organización y la forma, más precisamente de la auto-organización; la otra, en el de la transformación y la función, hoy mejor conocida como la evolución. La distinción sirve,

¹ El término «darwinismo» ha tenido connotaciones diferentes (Mayr 1991), por ejemplo, originalmente B. W. Richardson lo usó en 1856 para hacer referencia a las ideas transformistas de Erasmo Darwin, para efectos de claridad se llamará aquí «darwinismo» a la corriente de pensamiento de Carlos Darwin exclusivamente.

por un lado, para ubicar dentro de la perspectiva interna la crítica de Stuart Kauffman a la síntesis moderna, crítica sobre la cual gira gran parte de este trabajo, y por el otro, para señalar que ésta se omite en los textos universitarios y en los programas de estudio mencionados anteriormente.

La crítica kaufmaniana (Kauffman 1985) y de otros investigadores (Alberch 1989, Goodwin 1994, Solé y Goodwin 2000) esencialmente atacan al papel protagónico que se le atribuye a la selección natural al conformar los organismos y explican que existe una fuente de orden previo a ella, es decir, que la selección natural está supeditada a ella: la auto-organización.

Las visiones interna, la cual entiende a las criaturas con sus propiedades auto-organizadoras, y la externa, que las percibe como materia maleable ante las presiones del ambiente, responden a preguntas viciadas de origen. Preguntas que entienden a un ser como una unidad dual, en cuanto a su forma y su función, y que por lo mismo acotan sin querer su campo de estudio en la misma pregunta que se formulan sobre la naturaleza de los organismos. Esta dualidad permanecerá mientras se continúe haciéndose la pregunta en esencia sesgada hacia una parte de lo que son los seres. Kant postuló en *Crítica al juicio* (1790) la necesidad de entender a los organismos y sus partes como productos del por y del para, dicho en palabras evolutivas, que los organismos son uno en donde la evolución es causa y producto de la organización biológica.

Los morfoespacios: las playas de la variación

Las plantas del género *Passiflora* presentan una diversidad de formas de hojas impresionante. Se ha sugerido que son el resultado de una relación coevolutiva antagonista parcial con sus polinizadores, las mariposas del género *Heliconius* (Williams y Gilbert 1981). Al parecer esta asociación comenzó como un tipo de coevolución mutualista, los *Heliconius* obtenían néctares de las plantas y éstas conseguían un medio de polinización. En algún momento aparecieron mariposas mutantes que en lugar de poner sus huevos sobre otras especies los ponían ahora en la pasiflora. Esto complicó la relación. Las pasifloras debían de evitar ser reconocidas como hostal para huevos, pero sin ser totalmente invisibles para las mariposas porque de lo contrario perderían a su polinizador. Una de las estrategias de engaño que produjeron se basó en la capacidad de reconocimiento de los *Heliconius*. Estas mariposas identifican a las pasifloras por medio de la forma de sus hojas, entonces un cambio en la forma foliar evitaría que las mariposas ubicaran a las pasifloras. Aquí dio comienzo una carrera armamentista evolutiva en la cual la planta cambia continuamente la forma de sus hojas y las mariposas deben de ser capaces de distinguir estas versiones nuevas.

Al mirar la diversidad foliar de las pasifloras parecería que la capacidad de este género para producir hojas con morfologías diversas es inagotable. ¿Lo es? Para contestar esta pregunta se tendría que conocer de antemano cuáles son todas las hojas posibles, el universo de hojas, y así comparar este conjunto con el que son capaces de fabricar estas plantas, de esta forma se comprobaría si las pasifloras son en realidad tan creativas como aparentan. Este espacio, el de todas las formas posibles de una estructura, se denomina morfoespacio (Raup 1966). Actualmente disponemos de un conocimiento limitadísimo para construir morfoespacios.

Otra estructura natural que también parece haber sido creada por un artista prolífico es la concha. La diversidad de conchas es extraordinaria, las hay simples, las hay con formas caprichosas, churriguerescas y hasta las hay con proporciones místicas. En su trabajo clásico de 1966, Raup concluyó que para construir una concha se necesitan solamente tres parámetros, entonces se puede construir el morfoespacio de las conchas, el cual será un cubo, en el que se pueden situar a las formas de conchas que conocemos. Cada punto del cubo representa una combinación de tres medidas que origina una forma de concha. Todas las conchas que han existido, que existen y que existirán se localizarán dentro de este espacio, el morfoespacio de las conchas. Cuando se localizaron en él todas las conchas de las que se tienen noticia, sucedió que la parte ocupada por las formas reales era pequeña comparada con el volumen total del cubo (Figura 1).

Este resultado interesante provoca cuestionamientos de carácter evolutivo que, en un principio, debería contestar la síntesis moderna. Por un lado están las interrogaciones de qué es lo que sucede en el morfoespacio desocupado, es decir, ¿a qué se debe que no haya conchas en gran parte del morfoespacio? Por el otro, están las cuestiones acerca de qué pasa en el morfoespacio ocupado. La síntesis moderna contesta la primera cuestión de manera muy general, para la segunda, explica que la fuerza que origina esas formas es la selección natural principalmente. Lo que puede decir es algo muy general y vago. Veamos.

La razón de que no haya conchas en partes grandes del morfoespacio se puede atribuir a tres causas. La primera de ellas es el azar. No hay conchas en esos sitios porque no ha surgido la variación que las cree. La segunda de ellas es la selección natural. Esta fuerza es la responsable de eliminar características perjudiciales en las poblaciones de organismos, por lo que si se generase una forma de concha que fuese perjudicial para su portador, la selección natural la eliminaría como posible modelo viable. Entonces la selección natural actúa como un filtro escogiendo sólo aquellas conchas cuya eficacia tenga cierta magnitud relativa. La tercera causa son las restricciones de alguna índole, ya sean físicas, químicas o históricas, por mencionar algunas. Para hablar de restricciones es necesario tener un conocimiento muy preciso del organismo en cuestión, por ejemplo, una restricción física que tienen los insectos es su tamaño. Debido a que su modelo general de construcción posee un sustento externo, el exoesqueleto, no les es posible crecer más allá de cierto límite, si lo sobrepasan, su estructura se colapsará y se derrumbarán como edificios dinamitados. Esta es la contestación que proporciona la síntesis moderna y para la cual no se dispone de datos para saber qué conchas son imposibles y cuáles no lo son. Después de revisar estas posibilidades que proporciona la síntesis moderna se permanece en el mismo sitio, ya que no resuelve la pregunta central porque es incapaz de predecir sobre qué partes del morfoespacio podría transitar una especie y cuáles zonas le estarían vedadas. Ni siquiera puede calcular los tamaños de estas alternativas.

Con respecto a la zona ocupada del morfoespacio la síntesis moderna arguye que la selección natural es la causa principal, para algunos la única, de ese orden que se observa. La razón del por qué la parte ocupada del morfoespacio es mucho más pequeña que la desocupada sobrepasa por mucho sus capacidades de elucidación.

La síntesis moderna explica bien, gracias a la genética de poblaciones, el cambio de las proporciones alélicas de una población a lo largo de las generaciones. Predice si será eliminado un alelo o no dependiendo de sus frecuencias, de su coeficiente de selección y de otros fenómenos que puedan alterar estos porcentajes, como es la migración. Por otro lado, esta teoría también anticipa cómo cambiarán las formas en las poblaciones de acuerdo a las eficacias de las diferentes manifestaciones de las características. Habrá unas que sean más.

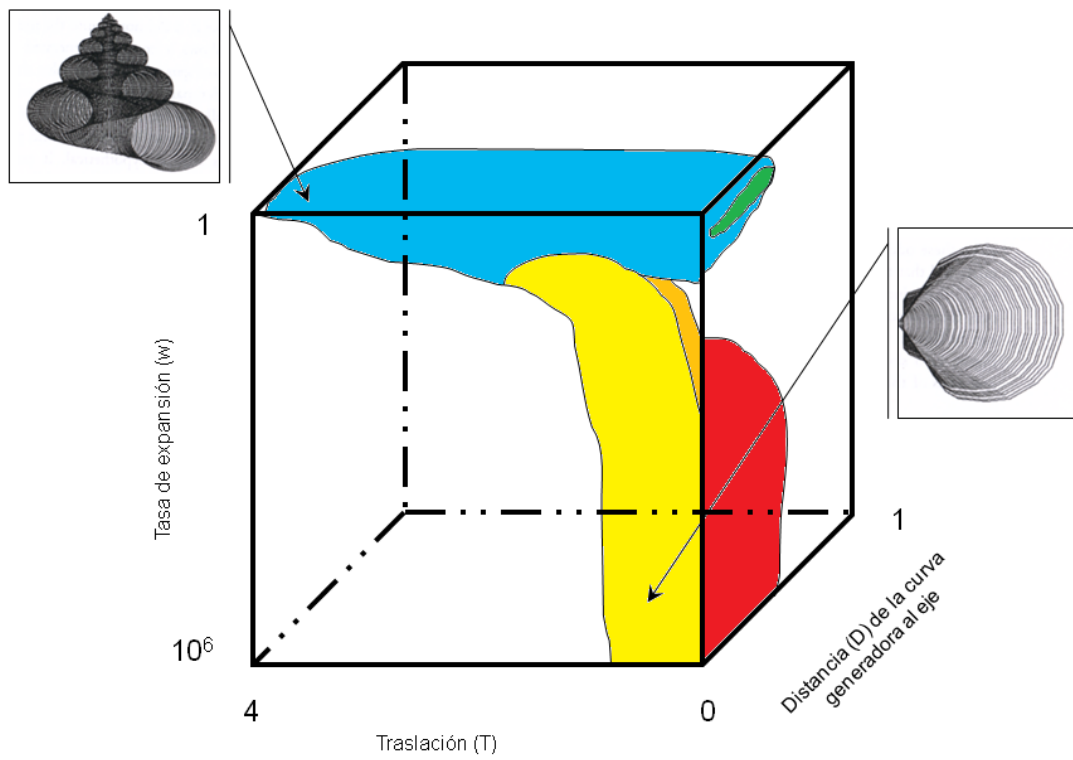


Figura 1: El morfoespacio de las conchas. Cualquier punto dentro de este cubo representa a una concha. Las conchas que se han registrado en la naturaleza se presentan como las zonas coloreadas. La zona azul representa las conchas de los gasterópodos, la verde la de los cefalópodos, la amarilla la de los bivalvos y la roja la de los braquiópodos. Se presenta la ubicación aproximada de conchas típicas de dos de estos grupos.

eficaces para desempeñar ciertas tareas ocasionando que sus portadores sobrevivan y/o se reproduzcan más provocando que a la larga, si las condiciones se mantienen, sean la forma predominante en la población. No obstante, la liga entre estos dos ámbitos, que es fundamental para el entendimiento total de la biología, está fuera de su poder explicativo

El reduccionismo: «modus operandi» genético

El reto de la biología, y por consecuencia de una teoría evolutiva, es el de explicar de qué manera están ligados el ámbito de las formas con el de los genes, es decir, cómo se mapean los efectos genéticos sobre las formas a lo largo del desarrollo y, por consiguiente, de la evolución. ¿Qué sucede en el ámbito de las formas cuando se hacen cambios en el ámbito genético? Tradicionalmente el enfoque adoptado por los científicos para responder esta pregunta ha sido el reduccionista. Este enfoque ha proveído avances impresionantes para entender la función de algunos genes, principalmente de aquellos con efectos claros y grandes en los fenotipos, y es el responsable directo de la ingeniería genética. Su manera de trabajar consiste en mutar consistentemente genes y observar los fenotipos haciendo una liga entre el cambio y el efecto de una forma unívoca. De aquí surge la idea de los «genes de algo», es decir, que si se muta cierto gen se ocasionará cierto efecto. Esta corriente ha ocasionado la búsqueda de genes con efectos deseables, normalmente para curar, prevenir o diagnosticar ciertas enfermedades o para satisfacer algunos de los deseos más primitivos del hombre y codiciados con más ahínco, como son la belleza y la inmortalidad. Algunos ejemplos de estas búsquedas de genes son el gen de longevidad, el de la homosexualidad y el de la religiosidad. Este enfoque ha producido, como ya se dijo, avances espectaculares; no obstante, muchas veces se queda en vilo ante algunos resultados inesperados. Su incapacidad no radica en falta de conocimiento, sino más bien en el modelo de investigación.

Este acercamiento genera la idea de que los genes son como herramientas, y que cuando se necesita hacer alguna función específica el organismo va a la caja de herramientas, toma el gen adecuado para la tarea, lo utiliza y lo regresa a la caja cuando ya no se necesite de él. Se sabe que hay genes cuyas mutaciones obscurecen o enmascaran el efecto de mutaciones en otros genes (epistasia), que un gen puede afectar a distintos rasgos aparentemente sin relación entre sí (pleiotropía) y que varios genes pueden estar involucrados en la formación de una característica (poligenia). Aunque se conocen estos efectos, el reduccionismo ha seguido siendo el paradigma metodológico para entender la genética, inclusive se presupone que la acción de varios genes sobre un fenotipo es aditiva y lineal normalmente. Pero varias investigaciones han puesto de manifiesto que la forma de trabajar utilizando mutantes simples fallará para revelar la organización funcional de los sistemas en los cuales la relación entre los elementos y las funciones no sea de uno a uno, entre otras cosas por las interacciones complejas y la redundancia (Deutscher *et al.* 2008). Recientemente se ha propuesto que el modelado genético debe de hacerse de una forma integradora y holística, en la cual los elementos están relacionados de formas intrincadas a manera de redes en los cuales se enmascaran muchos efectos de los elementos y los resultados no son lineales ni aditivos (Kauffman 1993, Espinosa-Soto *et al.* 2004, Wagner 2007).

Durante casi dos décadas se han estado recabando datos sobre los mecanismos genéticos que subyacen varios aspectos del desarrollo de las flores. Gracias a los mutantes

homeóticos que se han obtenido con dos sistemas modelos, *Antirrhinum majus* L. y *A. thaliana*, se han identificado los genes que controlan el destino celular durante la diferenciación de las estructuras florales: sépalos, pétalos, estambres y carpelos. Con estos genes, bajo el paradigma reduccionista, se ha construido el modelo ABC de desarrollo floral (Coen y Meyerowitz 1991). Veamos someramente cómo se ha construido dicho modelo.

Se han detectado ciertos genes, que cuando se les muta producen resultados fenotípicos anormales en la formación de las flores. Una flor perfecta se compone por cuatro anillos o verticilos concéntricos y en cada uno de ellos se forma una estructura diferente, en el primero, que es el más externo, se forman los sépalos, en el segundo los pétalos, en el tercero los estambres y en el último o central, los carpelos. Al mutar a ciertos genes se ha visto que pueden desaparecer algunas de estas estructuras, pueden surgir flores sin sépalos o sin pétalos. No sólo sucede eso, sino que además es posible que se formen algunas de estas estructuras florales en un lugar que le corresponde normalmente a otra, es decir, se puede generar una flor con pétalos en donde debería de haber estambres. Este tipo de mutantes se conoce como mutantes homeóticos.

El modelo ABC agrupa a los genes según sus efectos sobre la flor, el grupo de los genes *A*, el de los *B* y el grupo de los *C*, y propone ciertas interacciones entre estos grupos de genes (Figura 2). Si se muta un gen *A* desaparecerán los sépalos y los pétalos, y en sus lugares aparecerán carpelos y estambres. Al mutar un gen *B* desaparecerán los pétalos y los estambres, a cambio surgirán sépalos y carpelos, y por último, si se muta un gen *C* no habrá estambres ni carpelos y en esas posiciones brotarán pétalos y sépalos. De acuerdo con estos mutantes, las combinaciones de verticilos florales que generan y el lugar en donde se forman, se creó el modelo ABC que explica a los mutantes. Este modelo sí embona con los experimentos realizados; sin embargo, presenta dos problemas conceptuales y metodológicos inmensos. El primero es que narra lo que sucede en la realidad, la narrativa puede ser en algunos casos un deleite, pero en la ciencia muestra un grado bajo de desarrollo teórico y una carencia de madurez, eso se refleja en que no es posible hacer predicciones más allá de los datos con los que uno dispone. A pesar de estas limitaciones, el modelo ABC ha servido como guía que ha fomentado la investigación genética molecular del desarrollo en *A. thaliana* y en otras angiospermas debido a que la configuración floral y la función de los genes ABC se han conservado en la mayoría de las plantas con flor. El segundo problema es que el modelo es circular, los datos con los que se construye son los datos con los que se concluye, lo cual contradice la definición de modelo científico. Por lo tanto el modelo ABC, no es tal. Por otro lado, presenta el vicio de tomar en cuenta de una manera simplista y reduccionista las relaciones genéticas como lineales y aditivas en el mapeo del fenotipo. El modelo indica cuáles son los genes y el tipo de funciones que se necesitan para determinar cada uno de los órganos de la flor. Pero no establece qué componentes moleculares y cuáles interacciones son suficientes para determinar los órganos florales, ni tampoco el mecanismo que subyace el establecimiento en el espacio y en el tiempo de los propios genes ABC o de los órganos mismos.

Sin duda se necesita de un acercamiento que no tenga los defectos y limitaciones del modelo ABC.

En la determinación de los órganos florales intervienen muchos otros genes además de los genes de las funciones ABC. Los experimentos de sobreexpresión de los genes ABC en las hojas, demostraron que al menos se necesitan otros genes (los *SEPALATA*) además de las combinaciones adecuadas de los ABC para poder transformar hojas en alguno de los

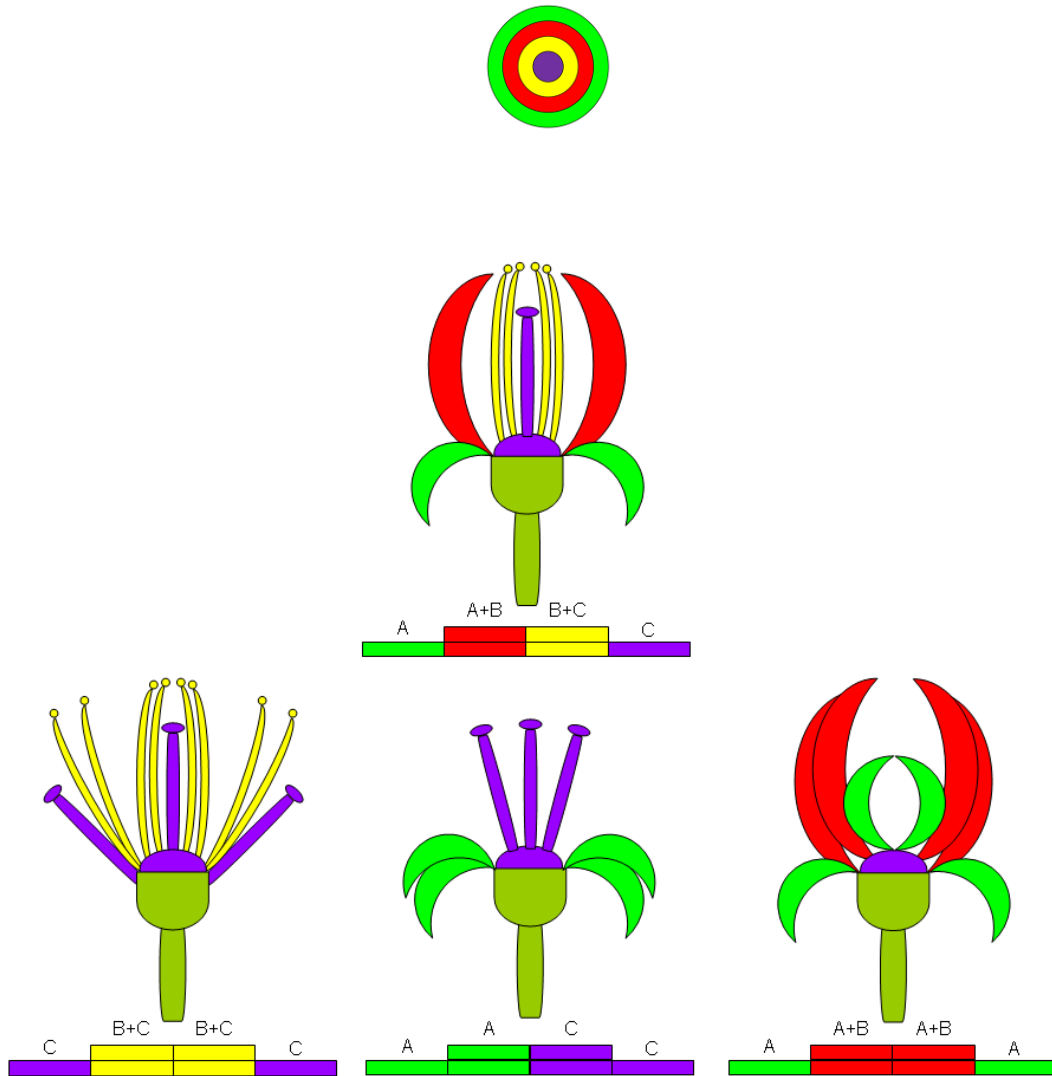


Figura 2. El modelo ABC. Hasta arriba se muestra un corte transversal del tallo de la planta en el cual se señala la disposición espacial de los verticilos florales: sépalos, pétalos, estambres y carpelos. En seguida se muestra una flor normal con sus cuatro verticilos en las posiciones correctas y la actividad de los tres grupos de genes: *A*, *B* y *C* según el modelo ABC (se muestra solamente la mitad del eje transversal del tallo). Abajo: Se enseñan los diferentes mutantes. La primera flor corresponde a mutar los genes *A*, el resultado que se produce es que se pierden los sépalos y los pétalos y en su lugar aparecen carpelos y los estambres. La segunda flor se origina cuando se mutan los genes *B* y lo que se obtiene es una flor sin pétalos ni estambres. En sus lugares surgirán sépalos y carpelos. En la última flor se han mutado los genes *C* ocasionando que desaparezcan los estambres y los carpelos y que ahora haya en esos sitios pétalos y sépalos. En todos los esquemas el verde señala a los sépalos, el rojo a los pétalos, el amarillo a los estambres y el morado a los carpelos.

órganos de la flor (Pelaz *et al.* 2001). Resulta entonces importante preguntarse cuál es la red de interacciones genéticas mínima suficiente que recupera las combinaciones de los genes ABC y de los que no son ABC que caracterizan a las células primordiales de los cuatro tipos de órganos de la flor.

Para abordar esta pregunta es necesario un enfoque teórico alternativo al modelo ABC. Este acercamiento nuevo se trata de modelar a los sistemas genéticos utilizando redes. Los primeros estudios con este método se han utilizado en *A. thaliana* (Mendoza y Álvarez-Buylla 1998 y Mendoza *et al.* 2000) y en otros organismos como en *Drosophila melanogaster* (Von Dasow *et al.* 2000 y Albert y Othmer 2003).

Antes de ingresar con detalle en el estudio de los sistemas genéticos mediante este método es necesario proporcionar un panorama del modelo general de redes, de su terminología y de sus conceptos.

El ABC de las redes

En sentido amplio² una red es un conjunto de elementos o nodos (N) unidos entre sí por conexiones (Newman 2003). Cada nodo posee ligas o uniones (K) que lo conectan con otros nodos. No todos los nodos o vértices de una red presentan el mismo número de conexiones, puede haber nodos con pocas y otros con muchas, el número promedio de conexiones (\hat{K}) de una red indica que tan conectados están sus nodos. El número de conexiones que presenta un nodo se conoce como su grado (Diestel 2000). Estas conexiones pueden tener uno o dos sentidos. Si la información fluye por ella en ambas direcciones se habla de una arista, si sólo va en una dirección, se denomina arco (Behzad *et al.* 1979). Cuando se tiene una red cuyas conexiones son arcos, las conexiones que tiene un nodo se dividen en dos, las conexiones de entrada (K_e) y las conexiones de salida (K_s). En este caso se habla de una red dirigida (Sorensen 2004).

Con este tipo de red se puede representar a diferentes sistemas de relaciones como son las redes sociales, las léxicas, como la que se presenta en la Figura 3 (para más detalles sobre ella vea el Apéndice 1), las de la Internet, las genéticas, las eléctricas, las que representan compuestos químicos, entre muchas otras (Harary 1972). Este tipo de modelo sirve para analizar la manera en la que se conectan los elementos de estas redes, es decir, se analiza su topología. En algunos de estos sistemas los nodos son inmutables ya que su estado no cambia y simplemente se estudia la forma en que se relacionan los elementos, pero en otros sistemas los elementos pueden cambiar su naturaleza, si este es el caso, será posible incorporar estos datos a la red e investigar otro ámbito de ellas como son las interacciones dinámicas. A las redes con nodos inmutables se denominarán redes estáticas y a las que tienen nodos que pueden cambiar se llamarán redes dinámicas.

Si un nodo tiene la capacidad de adquirir diferentes valores, por ejemplo, puede estar en dos situaciones diferentes, cero o uno, es un nodo binario, esta situación o condición del nodo se denominará *estado*. Al conjunto de los estados de todos los nodos de una red en cierto momento se le llamará *configuración*.

El número total de configuraciones de una red se calcula como la cantidad de estados que pueden tener los nodos elevada al número de nodos que tiene la red. Una red

² Algunas definiciones matemáticas de una red la describen como una gráfica dirigida junto con una función que asigna un número real positivo a cada línea (Harary 1972).

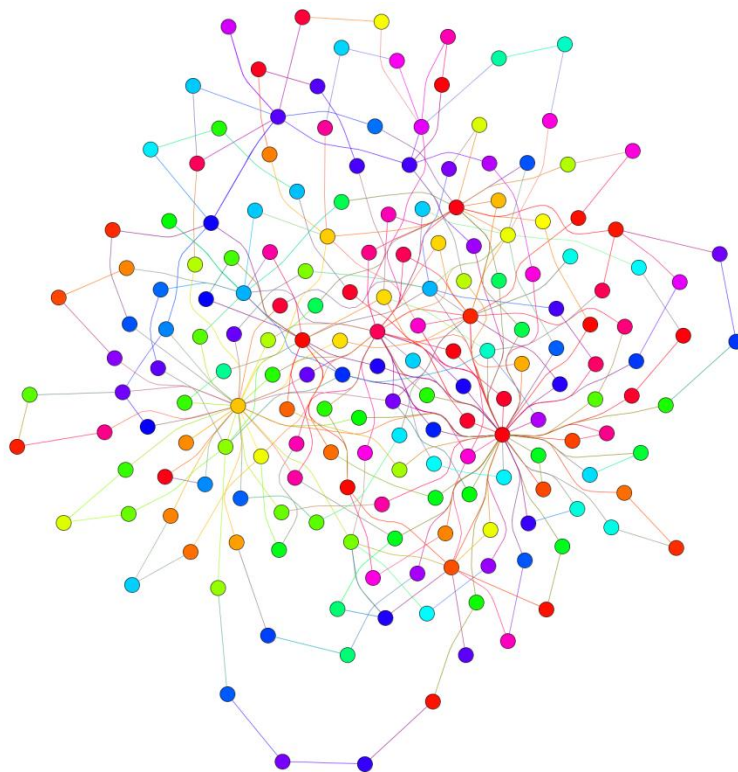


Figura 3: Un ejemplo de una red léxica, un tipo de red estática. Cada nodo representa una palabra distinta del poema «Vientos del pueblo» de Miguel Hernández.

con cuatro nodos binarios tendrá un número total de configuraciones de $2^4=16$. El conjunto de todas las configuraciones posibles de una red se le conoce como espacio omega.

Vea la red de la Figura 4, a la cual nos referiremos en adelante como Tetra. Es una red con cuatro nodos, cada uno conectado con una arista a otros dos y el espacio omega está formado por dieciséis configuraciones, por lo tanto, los valores de esta red son: $N=4$, $K=2$ (como las conexiones van en ambos sentidos la K_e de cada nodo es igual a la K_s).

¿Qué es lo que define el estado en que se encuentra un nodo? Un nodo presenta un estado de acuerdo a la combinación de los estados de los nodos que inciden sobre él. Supongamos que para que cierto nodo (el nodo objetivo) presente el estado 1 es necesario que los dos nodos que se conectan con él (los nodos de entrada) posean ambos el estado 1. Cualquier otra combinación posible de los estados de estos dos nodos, léase 0 0, 0 1 ó 1 0, provocará que el otro nodo permanezca en el estado 0. Este régimen de combinaciones de estados de los nodos de entrada a los cuales corresponde un estado del nodo objetivo se le conoce como la regla lógica del nodo (Figura 5).

En el ejemplo anterior se tiene que la regla lógica del nodo se compone de cuatro combinaciones que son solamente las configuraciones de los nodos que afectan al nodo objetivo. El número de combinaciones se calcula como el número de estados que puede presentar un nodo, elevado al número de nodos que inciden sobre el nodo objetivo, lo cual, para este caso, es $2^2=4$. Cada combinación corresponde a un estado en el nodo objetivo y se denominará *condición*. Las reglas lógicas se expresan a manera de tabla, la Tabla 1 presenta la regla lógica explicada anteriormente y que se conoce en lógica de Boole como la regla Y. Los nodos de una red pueden tener la misma regla lógica o pueden poseer diferentes reglas lógicas.

Tabla 1: Regla lógica Y del nodo C.

Estado del nodo A Entrada	Estado del nodo B Entrada	Estado del nodo C Salida
0	0	0
0	1	0
1	0	0
1	1	1

El estado de un nodo depende de los estados de los nodos que inciden sobre él y de su regla lógica, y a su vez, el estado de este nodo afectará el estado de otros nodos de la misma red sobre los cuales se conecte. Según estas reglas es posible tener configuraciones que no cambien, es decir, que después de ver qué estado le correspondería a cada nodo de la red según los estados de los demás y de las reglas lógicas, ninguno de los estados de los nodos cambie. Este tipo de configuración se denomina sumidero, específicamente este caso es un sumidero puntual. Caso contrario es una configuración que cambia y da lugar a otra configuración. Normalmente, tarde o temprano, una configuración llegará a una que no cambia, llegará a un sumidero. También es posible que se llegue a lo que se denomina sumidero cíclico, el cual consta de una serie de configuraciones que se van sucediendo de manera cíclica o repetitiva indefinidamente. Al número de configuraciones que se repiten incesantemente dentro de este ciclo se le conoce como el período del sumidero cíclico.

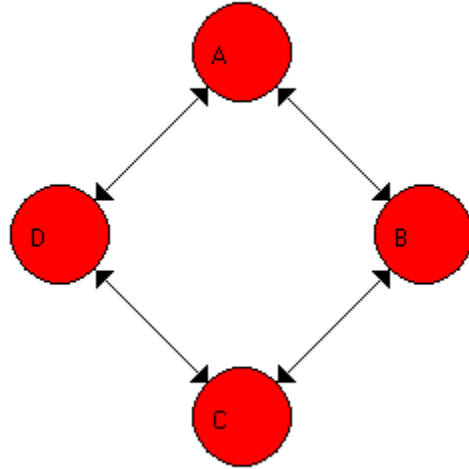


Figura 4: Una red de tipo $N=4, K=2$. A esta red se le apodará *Tetra*.

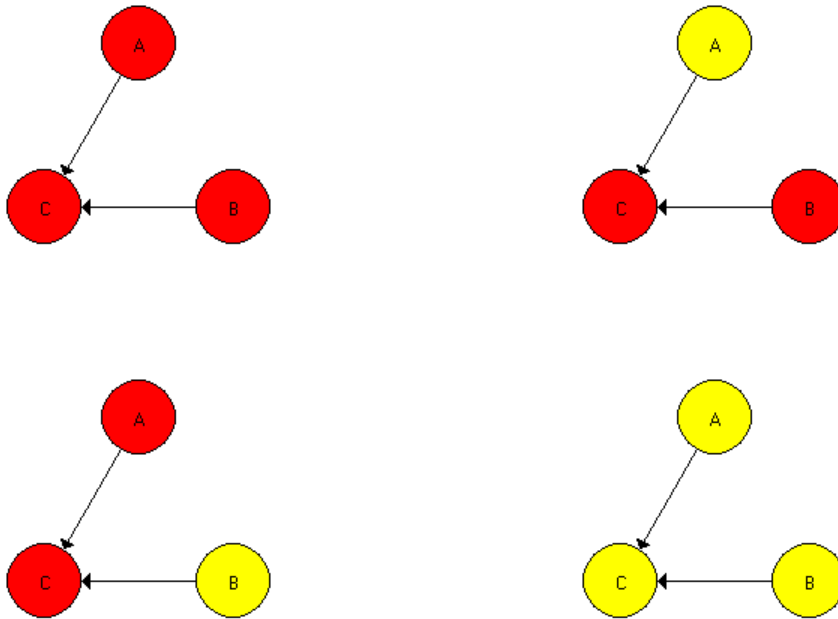


Figura 5: Cambio de estado de un nodo según los estados de los nodos que influyen sobre él de acuerdo a una regla lógica específica, en este caso es la regla *Y* que aparece en la Tabla 1. Los colores significan diferentes estados: cero en rojo, uno en amarillo. Para que el nodo C cambie de estado es necesario que los nodos A y B estén en el estado 1 (última red).

Puesto así, se puede decir que un sumidero puntual o de punto fijo es un sumidero con período igual a uno.

Para conocer el destino de todas las configuraciones es necesario rastrear a cada una de ellas hasta que lleguen a un sumidero puntual o a uno cíclico. Esto es lo que se conoce como la dinámica del sistema y se representa con una red también. La diferencia es que los nodos representan a las configuraciones y las conexiones al destino inmediato de ellas. Las conexiones siempre van en una dirección, es decir, son arcos y la red está dirigida. Otra característica importante de esta red es que cada configuración sólo puede tener un destino.

Si se desea conocer la dinámica del sistema de la red Tetra se deberá de calcular los destinos de las dieciséis configuraciones que componen el espacio omega y que se presentan en la parte izquierda de la Tabla 3. Para rastrear dichos destinos es necesario conocer las reglas lógicas de cada nodo, para este caso se tomará que todos siguen la regla lógica conocida como *O exclusivo (Ox)* y que se detalla en la Tabla 2.

Tabla 2: Regla lógica *O exclusivo (Ox)*.

Estado del nodo A Entrada	Estado del nodo B Entrada	Estado del nodo C Salida
0	0	0
0	1	1
1	0	1
1	1	0

En la Tabla 3 se presentan los destinos inmediatos de todas las configuraciones de la red de la Figura 4: Tetra.

Tabla 3: Destinos inmediatos de todas las configuraciones posibles de la red Tetra de la Figura 4.

Configuración inicial					Destino inmediato			
A	B	C	D		A	B	C	D
0	0	0	0		0	0	0	0
0	0	0	1		1	0	1	0
0	0	1	0		0	1	0	1
0	0	1	1		1	1	1	1
0	1	0	0		1	0	1	0
0	1	0	1		0	0	0	0
0	1	1	0		1	1	1	1
0	1	1	1	→	0	1	0	1
1	0	0	0		0	1	0	1
1	0	0	1		1	1	1	1
1	0	1	0		0	0	0	0
1	0	1	1		1	0	1	0
1	1	0	0		1	1	1	1
1	1	0	1		0	1	0	1
1	1	1	0		1	0	1	0
1	1	1	1		0	0	0	0

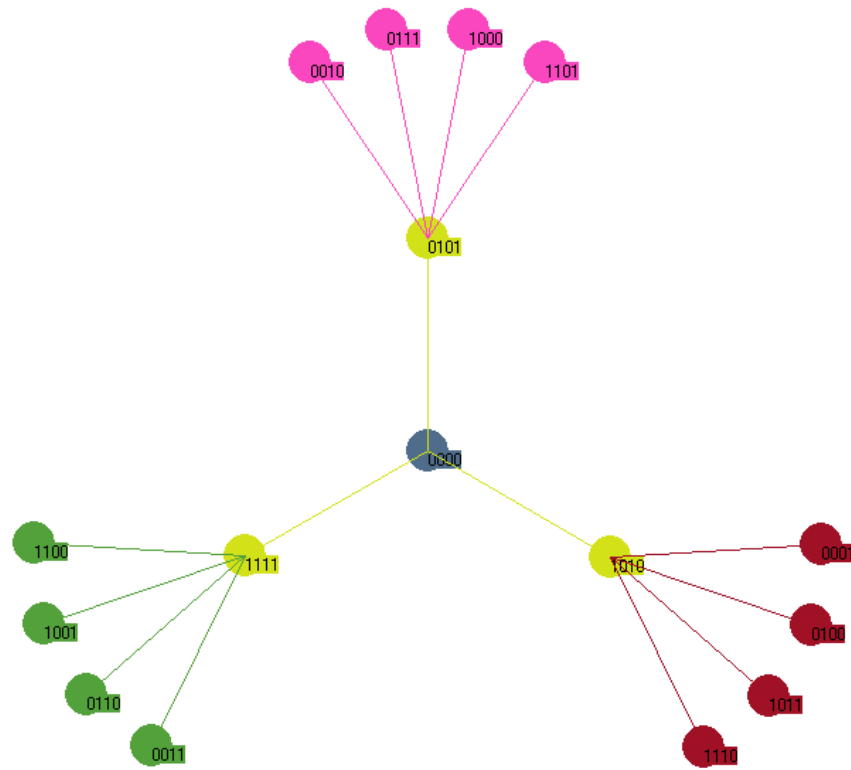


Figura 6: Dinámica de la red Tetra con la regla Ox . Se han etiquetado los nodos con las configuraciones para poder seguir el destino con detalle de cada una de las combinaciones del espacio Ω . Los colores están asignados al azar y solamente tienen valor estético.

Después de calcular los destinos inmediatos para cada configuración, que se presentan en la parte derecha de la Tabla 3, se procede a construir la red que representa a la dinámica de Tetra-Ox. Dicha red se presenta en la Figura 6.

Ya calculada la dinámica del sistema se puede ver que todas las configuraciones terminarán, tarde o temprano, en la configuración 0000 y una vez en ella se permanecerá allí para siempre, esa configuración es entonces un sumidero puntual. Al conjunto de todas las configuraciones que terminan en un mismo sumidero se le conoce como la *cuenca de atracción*. En este caso, como sólo hay un sumidero, todas las configuraciones del espacio omega pertenecen a la misma cuenca de atracción.

Es posible que una red siga teniendo el mismo sumidero aun cuando se le hayan cambiado algunas salidas de las reglas lógicas de sus nodos, esta propiedad se conoce como la robustez de la red. Si se cambiase la regla lógica *Ox* por una que tuviera en el tercer renglón un cero en lugar de un uno, llamémosle a esta regla *E* (que se detalla en la Tabla 4), la dinámica del sistema cambiaría en cuanto a los destinos inmediatos de algunas configuraciones, pero todas ellas acabarían en el mismo sumidero 0000 como se puede apreciar en la Figura 7.

En cambio, si la alteración de la regla *Ox* sucediese en el último renglón cambiando la salida de cero por la de uno, obteniendo lo que se conoce en el álgebra de Boole como la regla *O* (Tabla 5), no sólo se alteraría la dinámica del sistema, sino que además se crearían dos sumideros más, aparte del sumidero original. En la Figura 8 se pueden ver dichos sumideros, de los cuales dos son sumideros puntuales y el otro es un sumidero cíclico.

Tabla 4: Regla lógica *E*.

Estado del nodo A Entrada	Estado del nodo B Entrada	Estado del nodo C Salida
0	0	0
0	1	0
1	0	1
1	1	0

Tabla 5: Regla lógica *O*.

Estado del nodo A Entrada	Estado del nodo B Entrada	Estado del nodo C Salida
0	0	0
0	1	1
1	0	1
1	1	1

Ahora que ya se han descrito las características y los conceptos básicos de los modelos de redes, se podrá explicar cómo se utilizan con un enfoque genético.

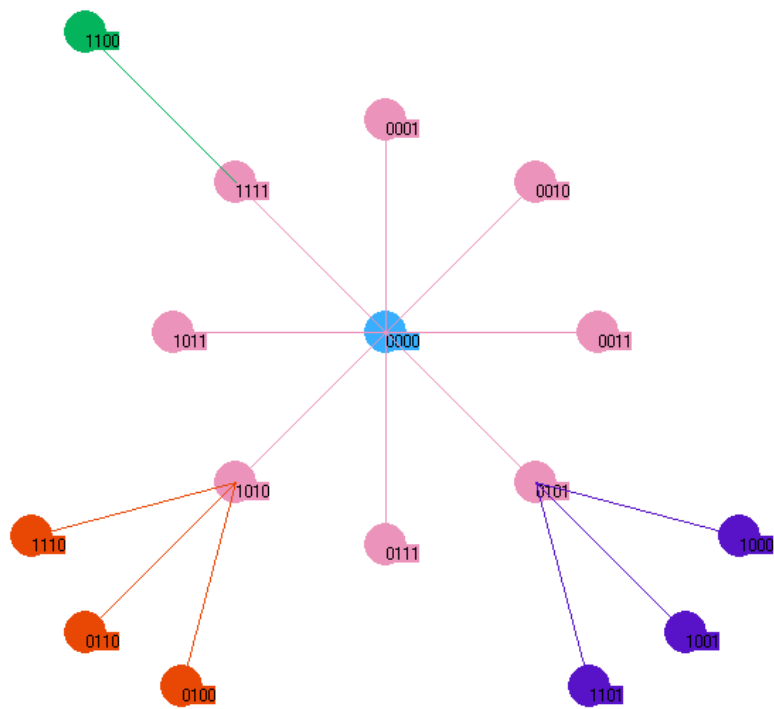


Figura 7: La dinámica del sistema Tetra-E cambia muy sutilmente con respecto a Tetra-Ox, presenta el mismo sumidero puntual (0000) que la del sistema Tetra-Ox y sólo cambian algunos de los destinos inmediatos de ciertas configuraciones. Los colores están asignados al azar y solamente tienen valor estético.

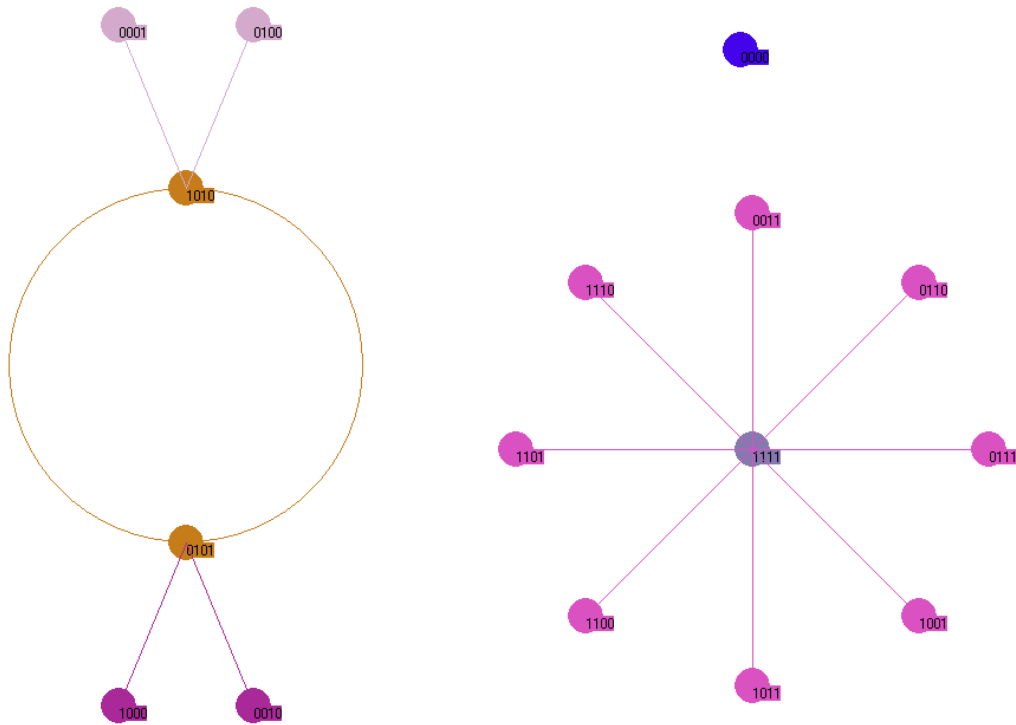


Figura 8: La dinámica del sistema Tetra-O. En este caso la dinámica cambia drásticamente debido a que se crean dos sumideros más, uno puntual (1111) y uno cíclico con período igual a 2 (1010-0101). Se sigue conservando el sumidero original (0000). El espacio omega se ha fragmentado y se forman tres cuencas de atracción con magnitudes de 9, 1 y 6. Los colores están asignados al azar y solamente tienen valor estético.

Las propuestas kaufmanianas: la vida al borde del caos

Kauffman, inspirado en los trabajos de Jacob y Monod (1961) sobre la regulación genética de los procariontes, propuso que los sistemas biológicos se podían modelar con el enfoque de las redes (Kauffman 1969). En este modelo los nodos representarían a los genes, los estados de los nodos simbolizarán la actividad del gen, *cero* si está inactivo y *uno* si está activo, y las conexiones significarán las interacciones entre los genes, es decir, si activan o reprimen a otros genes por medio de sus productos. De esta forma el estado de un gen estará determinado por la combinación de los estados de los genes de los cuales reciba entradas. Este régimen de combinaciones de entradas para las cuales el gen en cuestión responderá con un estado determinado, equivaldrá a la regla lógica, en este caso será la regla lógica propia del gen.

En su propuesta original se valió de un modelo de red binaria y con una conectividad aleatoria y uniforme, esto último significa que todos los nodos tienen la misma probabilidad de tener una conexión lo que genera que el grado de todos ellos sea casi igual, por lo tanto la distribución del grado de los nodos será aproximadamente una recta. Para este tipo de red el número aproximado de sumideros es igual a la raíz cuadrada del número de nodos (\sqrt{N}) (Kauffman 1983). Ahora bien, si los sumideros representan configuraciones estáticas que presenta cierta célula, sería posible que ellas determinasen el tipo celular en el cual se manifiesten. Para explorar esta idea de manera indirecta Kauffman hizo una relación entre el número de genes de varias especies y el número esperado de sumideros que tendrían sus genomas, entendido éste como el conjunto de redes genómicas. Por ejemplo, el resultado entre el número esperado de sumideros y de tipos celulares para el hombre coincidía bien. Calculó que el hombre poseía unos 100,000 genes, lo cual daba 370 tipos celulares, contra 254 que se calculaban en esa época (Kauffman 1991). Estos resultados generaron mucho optimismo sobre el modelaje de redes genéticas con este enfoque. Posteriormente se encontró que el número de genes que se utilizó estaba sobrevalorado, y por otro lado, el interés sobre las redes propició que se estudiaran diversos tipos de redes biológicas, entre las cuales figuraban las genéticas, y se descubrió que la conectividad de éstas no era uniforme. La conectividad que presentan es una conectividad muy sesgada y que sigue una ley de potencias (Aldana y Cluzel 2003). Una red genética posee muchos genes con pocas interacciones (K de magnitudes pequeñas) y pocos genes con muchísimas conexiones (K de magnitudes grandes) que funcionan a manera de concentradores. Esta es la razón por la cual el flujo de información es muy eficiente en una red genética o en cualquier otro tipo de red con esta topología, como son las de la Internet o las redes sociales (Barabasi y Oltvai 2004). Inclusive cuando estos descubrimientos son ciertos y los datos originales de Kauffman eran erróneos, el modelaje genético con redes binarias no se invalida como se verá más adelante.

Otra idea que se puede derivar del mismo modelo trata sobre la estabilidad de los sistemas genéticos y la evolución. Imagine una red genética que posea una robustez inusitada, a tal grado que pueda resistir cualquier cambio en sus reglas lógicas, esta red tendrá la ventaja de que no habrá mutación alguna que perjudique su desempeño. Si es una red involucrada en la formación de alguna característica, el organismo que la posea podrá estar tranquilo en que esta característica funcionará siempre de esa forma bajo las condiciones actuales. Esta situación no será favorable en un par de casos. Uno de ellos estaría relacionado con el cambio del ambiente en el que viva el organismo. Si las

condiciones en las que se desarrolla dicho organismo cambian, es probable que esa característica ya no presente alguna ventaja, por lo que su manutención en esta situación nueva quedaría en entredicho y representaría un gasto absurdo. Otro escenario más adverso estaría relacionado con la posibilidad de mejorar la estructura para incrementar la eficacia. Si la red no puede cambiar, será imposible crear variaciones de la característica en cuestión y que la selección natural entre en acción para llevar a la población a un nivel más alto sobre el pico adaptativo. Una red con una robustez extraordinaria no podrá evolucionar, estará congelada.

El caso contrario, el de una red muy flexible provocaría otro efecto. Las poblaciones de los organismos evolucionan de manera que algunas de las características que poseen se conservan y otras se modifican con el paso de las generaciones. Por muy rápida que sea la evolución, la mayoría de las características de las poblaciones se mantienen y una minoría de ellas se va modificando o perdiendo, la evolución no es un proceso que parte de la nada en el que desbarata su creación anterior para originar una nueva sin relación alguna con la original, no es una *tabula rasa*. Bajo ese esquema sería imposible la generación de organismos con cualquier tipo de complejidad. Ejemplificado con el modelo de redes, equivaldría a tener una red que construyera ciertas características y al momento evolutivo siguiente produciría unas características diferentes, la acumulación de características beneficiosas no se realizaría. Una red de este tipo no estaría congelada, sino que su régimen de cambio provocaría la desaparición caótica de los sumideros cada vez que se mutase alguna regla lógica. Sería una red revolucionada. Se puede concluir que las redes biológicas deben de tener un balance entre la capacidad de evolucionar y la robustez.

Por otro lado, es importante mencionar que Kauffman propone que las redes biológicas deben de estar en una zona que llama el «filo del caos». Si las redes tienen una conectividad de $K=2$, la dinámica del sistema se acota de forma inesperada obteniéndose muy pocos sumideros, y en caso de ser cíclicos, sus períodos son muy cortos. De cierta manera las propiedades intrínsecas de la red bajo ese esquema de conectividad crean funciones canalizadoras que restringen los caminos entre las diferentes configuraciones, lo que recuerda el paisaje epigenético propuesto por Waddington (1940). Para $K>3$, el número de sumideros se incrementa y los sumideros cíclicos aumentan muchísimo su período. En el caso extremo de $K=N$ el número de sumideros es enorme comparado con el de $K=2$. Por ejemplo, para $N=80,000$ el número de sumideros es aproximadamente de 30,000, cantidad que contrasta con los 283 sumideros cuando $K=2$ (Solé y Goodwin 2000).

Hay varios estudios recientes que apoyan las ideas del equilibrio entre la robustez y la capacidad de evolución (Espinosa-Soto 2004, Wagner 2007) así como sobre la posición de las redes biológicas en el espacio fase entre orden y caos (Balleza 2008 *et al.*).

Este es el escenario con el cual se abordará a los sistemas genéticos y, como consecuencia, a la evolución.

La decisión de lambda

Las redes kaufmanianas se prestan muy bien para modelar a los sistemas de regulación genética. De esta forma, los nodos representan a los genes y las flechas (arcos o aristas) simbolizan las interacciones que se dan entre ellos y que ocurren de manera indirecta debido a que interaccionan sus productos. Por ejemplo, si un gen produce un factor de transcripción que está involucrado necesariamente en la transcripción de otro gen,

la manera de representar esa relación es uniendo dos nodos, que representan al gen que produce el factor de transcripción y al gen objetivo, mediante una línea que va del primero al segundo. La línea que toca al gen objetivo se remata con una flecha, lo cual indica que la relación se da en ese sentido y que además es una relación de activación, es excitadora. Por otro lado, es posible que la proteína que fabrica un gen bloquee al operador de otro gen. En este caso la representación gráfica será la misma, con la excepción de que el remate de la línea no será una flecha, sino un círculo o una línea pequeña perpendicular indicando que la relación tiene el mismo sentido que la anterior, es decir, del gen que genera el producto hacia el gen regulado, pero este tipo de extremo muestra que la relación es inhibitoria. Cabe hacer una aclaración sobre la representación gráfica de la red. En las redes kaufmanianas simplemente se muestran las uniones entre los nodos; sin embargo, es posible y enriquecedor representar las uniones distinguiendo si son lazos que activan o inactivan a los genes. Esta representación no entra en conflicto con alguno de los supuestos de este tipo de redes, a cambio provee detalles de regulación útiles desde un punto de vista genético. Es por esta razón que las conexiones en el esquema de la Figura 9 tienen dos tipos de terminaciones: si son puntiagudas, se trata de una relación activadora, si son esféricas, la relación es inhibitoria.

Sírvanos la decisión que toma un virus sobre la fase que adoptará cuando infecta una bacteria para ejemplificar lo anterior con un caso real del cual se obtendrá la red de regulación según las actividades de los genes, de los factores de transcripción y de sus genes objetivos.

El bacteriófago lambda es un virus compuesto por una hélice doble lineal de ADN de 48,502 pares de bases con extremos cohesivos. Está formado por una cápside con forma de icosaedro, en donde guarda su material genético, y por una cola con la cual se ancla a la víctima para insertar su ADN posteriormente. Es un parásito de la bacteria *Escherichia coli* (Brock y Madison 1991).

Lambda puede existir de dos maneras, comportándose como fago o como profago. La primera forma, que se presenta durante la fase de lisis, ocasiona un comportamiento típico de un virus que acaba por producir copias múltiples de sí mismo y por romper la célula en la cual se introdujo matándola en el acto. La segunda se presenta durante la fase de lisogenia, conserva al virus dentro de la célula invadida en un estado de inactividad total, excepto por la producción de una proteína que mantiene inactivos a todos los demás genes del virus y quien causa precisamente ese estado de latencia. A los virus que poseen esta característica se les denomina atemperados (Lewin 1987).

Este bacteriófago posee uno de los circuitos de duplicación viral más intrincados. La cascada que origina la lisis está compuesta por dos genes reguladores que controlan las fases de la cascada de expresión génica, pero el circuito que controla la fase lisogénica se encuentra insertado dentro del circuito lítico.

Durante la infección, se expresa parte del conjunto de genes tempranos que producirán la lisis, lo que determinará cuál de las dos fases tendrá el virus dependerá de la acción competitiva de los productos de los genes tempranos y de la influencia de los factores de la bacteria. Es decir, la decisión de permanecer como profago se toma después de invadir a la bacteria. A grandes rasgos, una infección protagonizada por el virus lambda sigue los pasos siguientes.

Una vez que se ha anclado a la superficie de la célula bacteriana se dan lugar varios procesos cuyo fin es introducir la hélice doble de ADN a la bacteria. La infección comienza primero en transformar la hebra doble de ADN en una circular valiéndose de sus extremos

cohesivos. Posteriormente se activan los promotores P_L y P_R cuya transcripción expresa los genes N y Cro produciendo las proteínas correspondientes N y cro (Figura 9). Después cro se pega al operador del gen cI (Or_3) impidiendo su transcripción. N se pega a dos sitios Nut , uno en el gen N , en el marco de lectura izquierda, y el otro en el gen Cro , en el marco de lectura derecha. Debido a que la proteína N es un anti-terminador, los marcos de lectura se alargan a ambos lados y se transcriben los genes $cIII$, xis e int , del marco izquierdo de lectura, y cII , OP y Q del derecho. El producto de $cIII$ ($cIII$) protege de la digestión de ciertas proteasas que se encuentran en la bacteria al producto de cII (cII) y lo estabiliza también. Este momento es crucial, si la proteína cII se estabiliza, el virus entrará en la fase de lisogenia, si se degrada, entrará a la lítica. Hay ciertos factores conocidos que inclinan la decisión a favor de la lisis, algunos de ellos son la temperatura baja, las condiciones de hambruna y las infecciones múltiples.

Si las condiciones existentes favorecen la lisis, es decir, que cII se degrade, entonces la proteína cro bloquea la síntesis de cI pegándose a su promotor y la cascada de actividad genética desemboca en la lisis. Entre ellas, se encuentra la acción de la proteína Q , quien se encarga de dirigir la producción de proteínas líticas y estructurales que conforman las partes del virus.

Si por el contrario, cII se estabiliza, entonces promoverá la transcripción de los promotores P_{RE} , P_I y P_{antiq} . P_{RE} produce ARNm en contrasentido del gen Cro bajando la producción de la proteína cro , quien mantenía detenida la expresión de cI , y también de las proteínas necesarias para la lisis y las estructuras de los virus. De la misma forma P_{antiq} detiene la transcripción de Q . Al detener su actividad, no hay transcripción de los genes más allá de Q que producen la lisis y las partes del virus. El promotor P_I expresa la proteína int , quien se encargará de integrar el genoma del virus al de la bacteria.

Al bajar la concentración de cro , deja de obstruir el promotor de cI , entonces comienza la síntesis de este gen, la proteína cI se anclará en los sitios operadores de N y de Cro , desactivando su síntesis. Como la cascada de eventos comienza con N y Cro activos, al reprimirlos, todos los genes cascada abajo acabarán por apagarse. De esta forma el virus quedará en la fase de lisogenia. Una vez que se establece la fase lisogénica, la bacteria quedará inmunizada contra otros virus iguales debido a que la proteína inhibidora de lambda (cI) evitará que cualquiera de estos intrusos se pueda duplicar dentro de la célula en caso de que infecten al mismo anfitrión (Brock y Madigan 1991).

Cuando lambda se encuentra en lisogenia permanece inactivo, produciendo solamente la proteína cI ; sin embargo, el virus tiene la capacidad de salir de esa fase y entrar en lisis bajo ciertas condiciones. Por ejemplo, cuando ocurre un daño severo en el ADN de la bacteria, por causa de rayos X o rayos UV, el sistema SOS de emergencia de la bacteria se activa. Una de las proteínas de este sistema, cuya función radica en digerir ciertas proteínas que reprimen la acción de los genes de reparación de emergencia (operones inducibles), es la $RecA$. Esta proteasa tiene la capacidad de digerir también a la proteína represora de lambda (cI), lo cual provoca que el gen Cro pueda expresarse y que su proteína se pegue en la zona del promotor de la proteína represora de lambda, inhabilitando la producción de esta última. La inducción es una consecuencia indirecta de que se active el sistema SOS.

Después de revisar el modo de infección del virus lambda es posible establecer las relaciones entre sus genes y a partir de ellas construir la topología de la red genética que toma esta decisión. La red que se obtiene se presenta en el centro de la Figura 9. Las reglas.

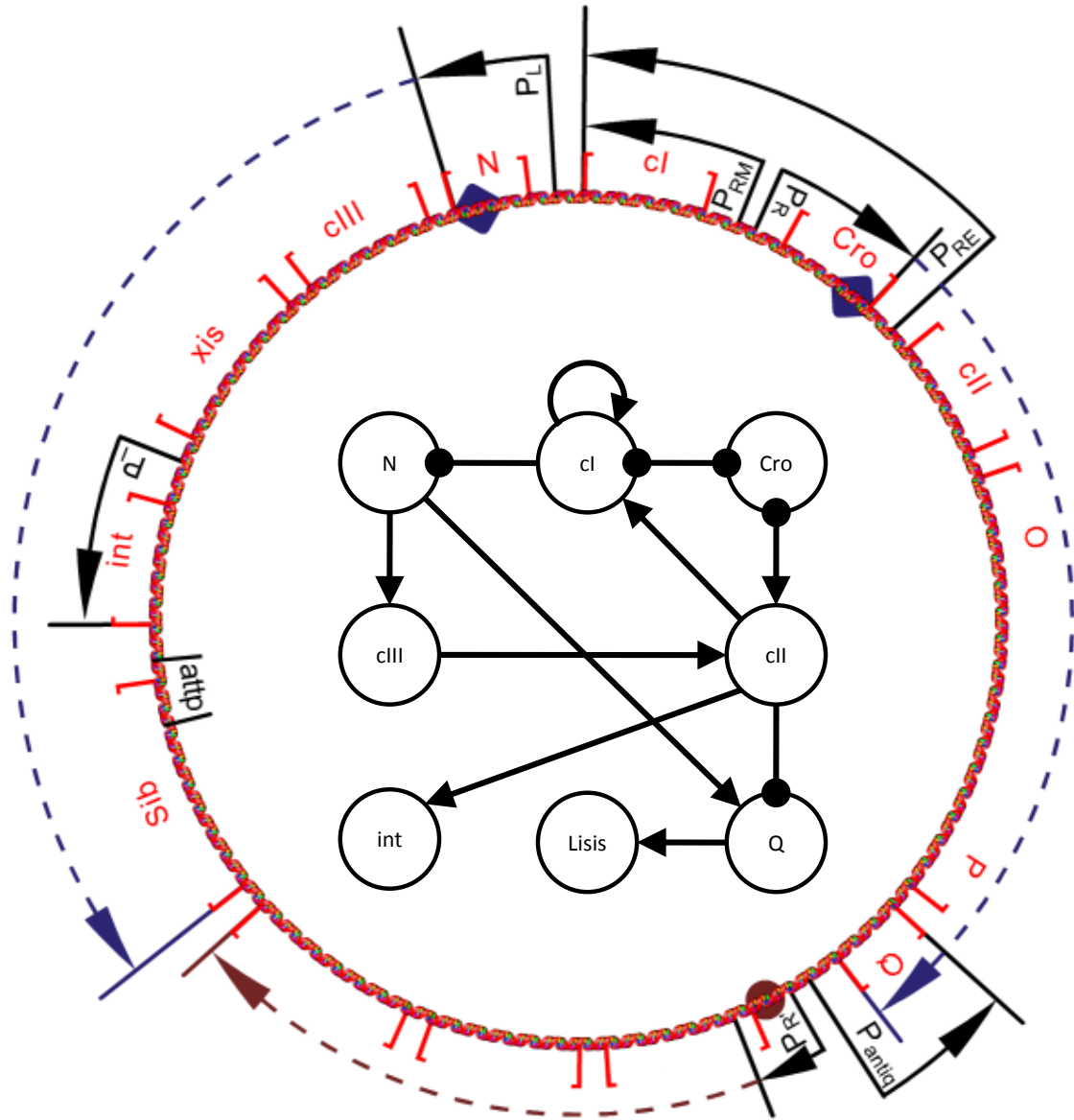


Figura 9. Mapa del genoma del virus lambda en donde se marcan los sitios donde se encuentran los genes que deciden durante la infección la fase que tomará el virus, lítica o lisogénica. En el centro de este mapa se muestra la red que describe el mismo proceso. Las puntas de las líneas indican el tipo de relación entre los genes: flecha si es activación y círculo si es inhibición. (Figura del mapa genético del bacteriófago lambda de Richard Wheeler)

se deducirán de evidencias experimentales sobre la actividad génica normal y de los estudios con mutantes

Para redondear el ejemplo, se ilustrará el caso del gen Q. El gen N activa a Q, mientras que el gen cII lo inhibe. Para cualquiera de los dos casos en los que sólo está activo un gen regulador el resultado es claro: si sólo N está activo, Q también lo estará; si cII es el único encendido, Q se apagará. Con relación a los casos en donde ambos genes reguladores se encuentran activos, situación que tarde o temprano terminará en la lisogenia, se deduce del ciclo infeccioso de lambda que el gen Q acabará inactivo. Para el caso último, donde ambos genes reguladores están inactivos, condición ya de lisogenia establecida, el estado de Q será de inactividad también. De este modo la tabla de verdad o la regla lógica del gen objetivo Q quedará de acuerdo a las entradas que recibe de los genes reguladores como se ve en la Tabla 6.

Tabla 6. Regla lógica del gen Q del bacteriófago lambda.

Estado del gen excitador: N	Estado del gen inhibidor: cII	Estado del gen objetivo: Q
0	0	0
0	1	0
1	0	1
1	1	0

Por lo anterior, el modelo de la red genética binaria que se utiliza en los trabajos que se presentan aquí consta de un conjunto de variables discretas y binarias donde cada una de ellas representa a un gen de la red. La configuración de la red en cierto momento queda expresada mediante el vector: $(g_1, g_2, g_3, \dots, g_N)$, donde N simboliza el número de genes que componen dicha red y g_n es el estado de expresión del gen n -ésimo. Si el gen se expresa $g_n=1$, si no lo hace $g_n=0$. El estado de cada uno de los genes cambia con el tiempo según la ecuación siguiente:

$$g_n(t + 1) = f_n(g_{n1}(t), g_{n2}(t), \dots, g_{nk}(t)) ,$$

donde $\{g_{n1}, g_{n2}, \dots, g_{nk}\}$ son los estados de los k genes que regulan la actividad del gen n y f_n es la regla lógica asociada al estado del mismo.

En realidad el modelo no prohíbe el caso de que las variables tengan más de dos estados. Por ejemplo, la red floral con la que se comenzó a trabajar era una red compuesta por una combinación de nodos binarios y de nodos que podían adoptar uno de tres estados. En el Capítulo 2 se verá cómo se tradujo esta red a una red binaria y sus implicaciones ontológicas.

Es importante hacer notar que la dinámica de este sistema es determinista, lo cual significa que dada una configuración inicial, el sistema siempre llegará a la misma configuración final. En el Capítulo 3 se analiza el papel que podrían representar fuentes de naturaleza aleatoria que modifican momentáneamente el comportamiento de las reglas lógicas en la evolución biológica.

La flor enredada

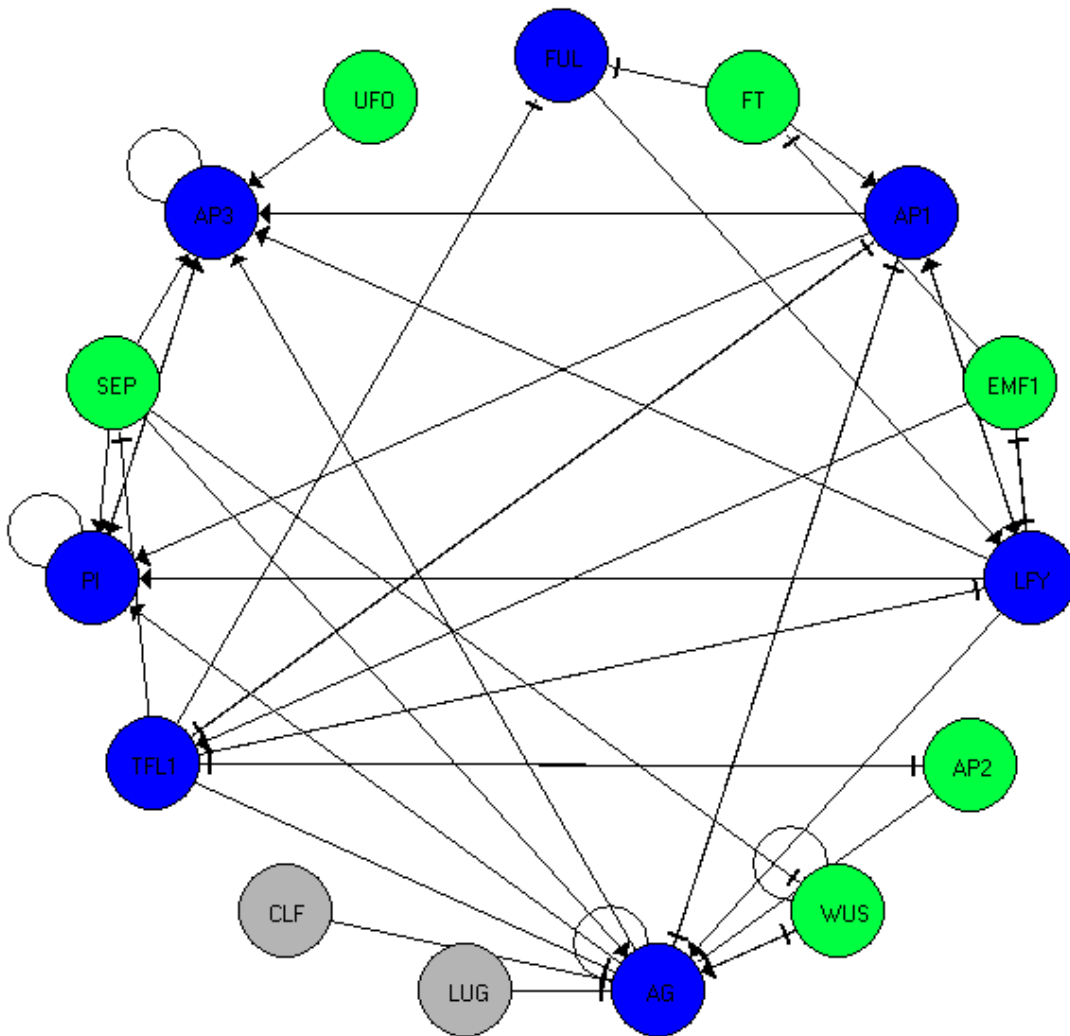
Para construir una red genética se necesita de muchísima información sobre los genes responsables de cierta característica y de un modelo real adecuado. Actualmente hay muy pocas redes genéticas construidas a partir de datos experimentales, entre las primeras que se han obtenido se encuentran la red floral de *A. thaliana* (Mendoza y Álvarez-Buylla 1998, Mendoza *et al.* 2000, Espinosa-Soto *et al.* 2004) y la red que determina la polaridad de los segmentos en el embrión de la mosca de la fruta (*D. melanogaster*) (Albert y Ohtmer 2003). Este trabajo tomó como punto de partida el estudio de Espinosa-Soto *et al.* sobre la determinación floral.

Bajo la perspectiva del modelo ABC se han identificado los genes de los grupos A, B y C que están involucrados en la formación de la flor. Para empezar a hacer la red genética de la diferenciación de la flor se tomaron los mismos genes como punto de partida. La diferencia con el modelo ABC radica en que además se documentaron las interacciones que tienen estos genes entre sí y con otros que no pertenecen necesariamente al conjunto de genes ABC.

En un primer intento (Mendoza y Álvarez-Buylla 1998) se construyó una red a partir de la información experimental que se disponía entonces. Dicha red constaba de once genes; sin embargo, no se dedujeron las reglas lógicas de los genes, en lugar de eso utilizaron evidencia experimental para encontrar las magnitudes relativas de las interacciones. La manera en que un gen se activa en este trabajo depende de la fuerza sumada de cada una de las interacciones individuales que recibe. Para encontrar las magnitudes de las interacciones y los umbrales de respuesta de cada uno de los genes utilizaron algoritmos genéticos. El marco de comparación fue el mismo modelo ABC, lo cual generó una circularidad en el estudio y que no se pudiese discernir entre los efectos de la red en sí o del modelo de optimización. Posteriormente realizaron otro acercamiento basado en las asas de retroalimentación positivas en el cual obtuvieron los mismos seis sumideros que en el primer trabajo (Mendoza *et al.* 2000). Justificaron las incongruencias de los sumideros encontrados con los perfiles naturales a que a la red le faltaban genes que no se habían identificado.

En un trabajo subsiguiente Espinosa-Soto *et al.* (2004) recopilaron una gran cantidad de datos provenientes de estudios actuales y con estos datos lograron tener un panorama mucho más completo de la red. Identificaron cuatro genes más, encontraron relaciones nuevas y, lo más importante, formalizaron las interacciones reguladoras de los genes a partir de datos experimentales en funciones lógicas para cada uno de los genes de la red. Dicha red consta de quince genes (Figura 10), de los cuales dos de ellos siempre permanecen en el mismo estado, seis de ellos son binarios y siete tienen tres estados.

El espacio omega de esta red consiste de 139,968 configuraciones. Cuando se calcula la dinámica del sistema se encuentran 10 sumideros puntuales cuyas configuraciones coinciden con los perfiles de expresión genéticos que están documentados experimentalmente en las células de los órganos y meristemos florales (Tabla 7). La red también reproduce las expresiones genéticas de los mutantes.



Atalia

Figura 10: Red de *A. thaliana* de Espinosa-Soto *et al.* 2004 que se compone de 15 genes. Los genes en color azul son genes que presentan uno de tres estados (0, 1,2), los genes en color verde son binarios (0,1) y los genes en color gris sólo presentan un estado (1). Las flechas indican que la interacción es una activación y las rayas que es una inhibición. Todas las asas de retroalimentación son activaciones.

Tabla 7: Las configuraciones de los 10 sumideros puntuales obtenidos por Espinosa-Soto *et al* (2004) junto con el tipo de célula al que están asociados.

FT	EMF1	TFL1	LFY	FUL	AP1	AP3	PI	AG	UFO	WUS	AP2	SEP	LUG	CLF	Célula
0	1	2	0	0	0	0	0	0	0	0	0	0	1	1	Inflorescencia
0	1	2	0	0	0	0	0	0	1	0	0	0	1	1	Inflorescencia
0	1	2	0	0	0	0	0	0	1	1	0	0	1	1	Inflorescencia
0	1	2	0	0	0	0	0	0	0	1	0	0	1	1	Inflorescencia
1	0	0	2	0	2	0	0	0	0	0	1	1	1	1	Sépalos
1	0	0	2	0	2	2	2	0	1	0	1	1	1	1	Pétalos
1	0	0	2	0	2	2	2	0	0	0	1	1	1	1	Pétalos
1	0	0	2	2	0	2	2	2	1	0	1	1	1	1	Estambres
1	0	0	2	2	0	2	2	2	0	0	1	1	1	1	Estambres
1	0	0	2	2	0	0	1	2	0	0	1	1	1	1	Carpelos

Ahora sí se dispone de un modelo sustentado en evidencia experimental clara en el cual no se obtienen datos de las relaciones o de los pesos de forma circular. Este modelo apoya fuertemente la idea de que el modelado genético mediante redes discretas es posible y fructífero, además de que sustenta la hipótesis kaufmaniana de asociar los sumideros con tipos celulares. Es precisamente de esta red con la cual se partirá para realizar los trabajos que a continuación se muestran. Además, en este trabajo se hicieron simulaciones que muestran que los sumideros que corresponden a las configuraciones de activación genética de las células primordiales de sépalos, pétalos, estambres y carpelos, son robustos a alteraciones puntuales de las reglas lógicas.

Capítulo 2

De los genes a las formas florales y a la evolución: modelos dinámicos de redes genéticas reguladoras

El enfoque de redes genéticas ha probado ser un método útil para describir y entender las configuraciones genéticas de los diferentes tipos celulares en el desarrollo de la flor (Espinosa-Soto *et al.* 2006). Lo cual también apoya que la dirección en la cual se deben de abordar los temas y estudios evolutivos es la de modelar los sistemas vivos como ensambles de redes, bajo una perspectiva holística e integradora, en lugar del esquema tradicional y reduccionista de concebirlos como simples cajas de herramientas.

En el estudio que se presenta en este capítulo se actualizó la red de *A. thaliana* con referencias nuevas para tener la red de diferenciación floral con mayor precisión de acuerdo a los datos que han surgido recientemente. Con base en la idea de que la lógica genética funciona de forma binaria, ya que un gen puede estar encendido o apagado, situación descrita en el modelo original de Kauffman y basado en el modelo del operón, se planteó la hipótesis de que la red con algunos nodos de tres elementos podría ser reducida a una red binaria con la misma arquitectura general sin alteración de los sumideros. Entonces, en este trabajo se simplificó la red en cuanto al número de estados que consideraba. La red original tomó en cuenta tres estados, bajo nuestro concepto de cómo se deben de modelar estos sistemas creemos que los sistemas biológicos deben de ser modelados con dos estados, siendo uno de ellos cuando el gen se activa y el otro cuando se desactiva. Al hacer esta reducción o traducción, los resultados que se obtienen son exactamente iguales a los previos con el modelo de tres estados, lo cual indica que se justifica el modelaje con dos estados de estos sistemas genéticos. Ese mismo análisis también puede verse como una prueba adicional de la robustez de la red.

Posteriormente se realizó una prueba de robustez más directa que consistió en mutar sistemáticamente la salida de todas las combinaciones de cada una de las reglas lógicas de todos los genes. Los resultados demostraron que la red floral de *A. thaliana* es muy robusta, lo que concuerda con lo que se esperaría bajo la teoría, es decir, que las redes biológicas deben de ser muy robustas con respecto a las mutaciones, pero sin llegar a ser totalmente resistentes a ellas. Lo cual apoya la idea original de Kauffman, en la cual postula que la vida se encuentra entre la fase congelada y la caótica del comportamiento de las redes. Después de ciertas exploraciones con redes de cinco organismos que provienen de cuatro reinos se encontró evidencia que sustenta esta idea (Balleza *et al.* 2008).

From Genes to Flower Patterns and Evolution: Dynamic Models of Gene Regulatory Networks

Álvaro Chaos,^{1*} Max Aldana,^{2*} Carlos Espinosa-Soto,¹
Berenice García Ponce de León,¹ Adriana Garay Arroyo,¹ and
Elena R. Alvarez-Buylla^{1*}

¹Laboratorio de Genética Molecular, Desarrollo y Evolución de Plantas, Dpto. de Ecología Funcional, Instituto de Ecología, UNAM, C. P. 04510, México DF, México; ²Centro de Ciencias Físicas, UNAM, C. P. 62251, Apartado Postal 48-3, Cuernavaca Morelos, México

ABSTRACT

Genes and proteins form complex dynamical systems or gene regulatory networks (GRN) that can reach several steady states (attractors). These may be associated with distinct cell types. In plants, the ABC combinatorial model establishes the necessary gene combinations for floral organ cell specification. We have developed dynamic gene regulatory network (GRN) models to understand how the combinatorial selection of gene activity is established during floral organ primordia specification as a result of the concerted action of ABC and non-ABC genes. Our analyses have shown that the floral organ specification GRN reaches six attractors with gene configurations observed in primordial cell types during early stages of flower development and four that correspond to regions of the inflorescence meristem. This suggests that it is the overall GRN dynamics rather than precise signals that underlie the ABC model. Furthermore, our analyses suggest that the steady states of the GRN are robust to random alterations of the logical functions that define the gene interactions. Here we

have updated the GRN model and have systematically altered the outputs of all the logical functions and addressed in which cases the original attractors are recovered. We then reduced the original three-state GRN to a two-state (Boolean) GRN and performed the same systematic perturbation analysis. Interestingly, the Boolean GRN reaches the same number and type of attractors as reached by the three-state GRN, and it responds to perturbations in a qualitatively identical manner as the original GRN. These results suggest that a Boolean model is sufficient to capture the dynamical features of the floral network and provide additional support for the robustness of the floral GRN. These findings further support that the GRN model provides a dynamical explanation for the ABC model and that the floral GRN robustness could be behind the widespread conservation of the floral plan among eudicotyledoneous plants. Other aspects of evolution of flower organ arrangement and ABC gene expression patterns are discussed in the context of the approach proposed here.

Álvaro Chaos, Max Aldana and Elena Alvarez-Buylla contributed equally to this work.

Received: 29 June 2006; accepted: 30 June 2006; Online publication: 26 November 2006

*Corresponding author; e-mail: elena.alvarezbuylla@gmail.com; ealvarez@miranda.ecologia.unam.mx

Key words: Gene regulatory networks; Flowers; Pattern formation; Dynamic models; Discrete models; Evolution; Development; Boolean networks; Self-organization; ABC model

INTRODUCTION

Gene interactions are prevalent and key during development, and thus development may bias, restrict, or mediate the variation that evolutionary forces may act on. Therefore development should be explicitly considered in evolutionary models that aim at understanding the origin and evolutionary dynamics of phenotypical traits (Lewontin 1974; Salazar-Ciudad and others 2000).

Complex networks of interacting components at different levels of biological organization are being uncovered at the molecular level. Genes and proteins form complex dynamical systems or gene regulatory networks (GRN) that can reach several steady states (attractors; Figure 1). These have been associated with distinct cell types (Kauffman 1969), and this proposition has been recently supported experimentally (Acar and others 2005; Huang and others 2005). Our aim is to contribute to a general framework for understanding the interplay between the structure, function, and evolution of GRN and their role in pattern formation that is central for understanding development.

The approach we have followed has been to construct and validate qualitative models of specific biological processes that have been functionally characterized. The long-term aim is to infer novel and general regulatory principles of cellular networks by analyzing the structural and dynamical features of the specific gene regulatory networks studied and modeled. We think that this approach will complement the more theoretical ones that aim at uncovering general principles for network assemblage and dynamics using mathematical and simulation models or intend to infer global GRN architectures relying on functional genomics data. All approaches should provide feedback and complement one another.

Structural and dynamical analyses of well-characterized subnetworks grounded on solid experimental data are also useful to integrate experimental data and to find holes and inconsistencies in such data. Gene regulatory network models grounded on experimental data are also important to validate inferences of network connectivity from functional genomic data. Indeed, the network that we proposed for cell fate determina-

tion during floral organ specification (Mendoza and Alvarez-Buylla 1998) has been used by several researchers to validate methods of GRN architecture inference (Perkins and others 2004; Aracena and Demongeot 2004).

We have focused on plant systems to pursue our aims. In plants, multicellularity evolved independently from animals. A comparative approach of multicellular development in its broadest sense should therefore consider plant as well as animal systems (Meyerowitz 2002). Moreover, plants are relatively simpler than animals in their cellular organization, and therefore it will be more feasible to propose models that scale from network behavior to cell and morphogenetic patterns. Within plants, comparative analyses of networks controlling cell-fate determination or pattern formation may help uncover subnetworks that have been fairly fixed during evolution, modules that are co-opted for several cell-fate determination functions within or among plant species, and critical network motifs or subnetworks underlying important functions and evolutionary innovations.

The ABC's of Flower Development

Flowers are the most complex and showy multicellular structures of angiosperms, and they have attracted the attention of many researchers. Upon induction to flowering, the shoot apical meristem that only produces vegetative structures from its flanks turns into an inflorescence meristem that also produces flowers. Adult floral morphology originates from floral meristems that arise from the flanks of the inflorescence meristem in a spiral arrangement. One day after flower induction a flower buttress arises, and 1.25 days afterwards the flower primordium forms in the flanks of each floral meristem. The sepal primordia arise first in the periphery of the floral meristem at stage three (3 days) of flower development; then the petal and stamen primordia appear in the second and third whorls, respectively, almost simultaneously. Last, the carpel primordium is formed from the remaining central portion of the floral meristem (Bowman 1994).

Thus, a few days after flower induction, the floral meristem is partitioned into four regions, each

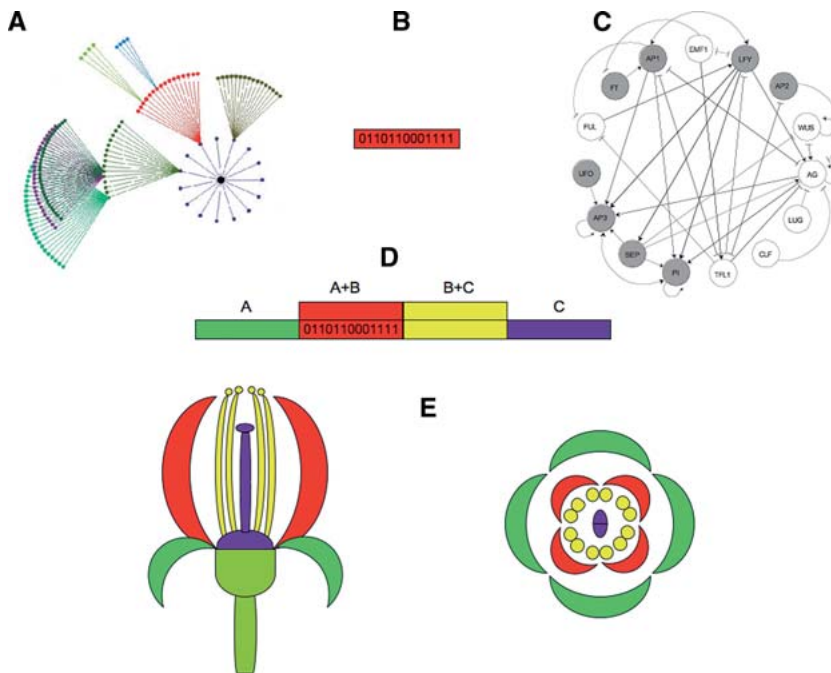


Figure 1. Floral structure of *Arabidopsis thaliana*, ABC model and GRN. (A) Basin of attraction of the petal-2 point attractor. (B) Pattern of expression of the genes, where 0 means “off” and 1 means “on.” (C) The same pattern depicted on the network; the dark circles indicate that the corresponding node is “on” and the rest are “off.” (D) The pattern of petal-2 corresponds to the overlapping activity of A + B genes. (E) The structure emerging from this pattern of activity: petals are depicted in red in the longitudinal and radial floral diagrams.

composed of primordial sepal, petal, stamen, and carpel cells that are sequentially differentiated from the outer whorl to the center of the meristem. Floral organ primordia continue to grow and differentiate into the distinct cell types that constitute each of the mature floral organs. The models that we have developed refer to the GRN that underlie primordial cell fate determination and hence underlie the combinatorial gene activities that characterize the four regions into which the flower meristem is partitioned during the early stages of flower development.

A wealth of data on the molecular genetic mechanisms underlying different aspects of flower development have accumulated during the last 15 years (Coen and Meyerowitz 1991). The genes controlling cell fate determination during floral organ specification have yielded the ABC model of flower development that was derived from genetic analyses of floral organ homeotic mutants in two plant study systems: *Antirrhinum majus* L. and *Arabidopsis thaliana* (L.) Heynh. (Coen and Meyerowitz 1991). These species have the stereotypical floral arrangement of almost all eudicotyledoneous species (see Vergara-Silva and others 2003; Ambrose and others 2006 for an exception), with sepals in the outermost whorl, then petals, stamens, and carpels in the flower center. The ABC model states that the identities of the floral organ types are established by combinations of genes grouped in three main classes, A, B, and C. A genes alone determine sepal identity; A plus B, petal identity;

B plus C, stamen identity; and C alone, carpel identity (Figure 1). However, the ABC model does not provide an explanation for how such combinatorial selection of gene activity is established during floral organ primordia specification, and how the spatiotemporal pattern of ABC and non-ABC gene expression is established. Furthermore, the conserved pattern of floral organ determination and also the overall conservation of the ABC gene patterns of expression among eudicotyledoneous species suggest a robust mechanism underlying such combinatorial selection of gene activities. The ABC model by itself does not provide an explanation for such robustness either. We have been developing qualitative GRN models for tackling these questions

Discrete Gene Regulatory Network Models: The Formalism

Regulatory network models are abstractions of the multiple regulatory interactions among genes during cell-fate determination and pattern formation. Each gene undergoes a network of mRNA transcription, translation, protein modifications, and protein transport until its final functional destiny. Most data available for plant development is at the transcriptional regulatory level, although recent findings suggest that miRNA regulation is important during plant development (Chen 2004).

Gene regulatory network models may incorporate continuous or discrete functions for the gene-

activation kinetics. Because the temporal scales of the molecular processes of the pathway that goes from DNA transcription to protein function are relatively short in comparison to those of the processes of pattern formation, it is reasonable to focus only on the qualitative regulatory interactions. Continuous implementations can incorporate more detail and can yield quantitative predictions. However, experimental data that provide estimates of such models are lacking. Furthermore, formal analyses of equivalent continuous and discrete models have analytically shown that both yield equivalent results (Thomas and others 1995). Furthermore, in systems of many components with many non-linearities, the behavior of the system depends mostly on the qualitative aspects of the GRN topology rather than on the kinetic details for each interaction and component. Therefore, we postulate models that assume discrete functions for the kinetics of gene activation. In the simplest case, Boolean functions that may only have two values: 0 ("off") or 1 ("on") are assumed. The activity of each gene depends on updating rules that are formalized as logical functions grounded on experimental data. Boolean models have also been successfully used in animal systems. For example, Albert and Othmer (2003) have shown that a Boolean GRN recovers the same patterns for the *Drosophila melanogaster* segment polarity genes as those recovered by a continuous model (Von Dassow and others 2000) analysis that had already been shown to be robust to parameter perturbations.

Discrete models are further justified because recent experimental evidence suggests that gene expression is digital and stochastic at the individual cell level rather than continuous. In cell aggregates, gene expression may appear to change continuously as an inductive signal increases. However, recent experimental data at the individual cell level suggests that it is the probability of a particular template to be active within a certain time window, rather than the rate of transcription from this template, that seems to be subject to regulation. Hence, genes within individual cells have a distinct probability of responding to a given concentration of stimulus of transcription, and the gene is either "on" or "off" in a certain time window (Hume 2000; Ozbudak and others 2002; Elowitz and others 2002; Blake and others 2003; Paulsson 2004; Walters and others 1995; Fiering and others 2000; Ho and others 1996). Rossi and collaborators (2000) have proposed a mechanism that is necessary and sufficient to account for the binary response in inducible gene expression that proposes that for each site in a promoter region there are transcrip-

tion factors with opposing functions (for example, activation and repression) for the same target promoter motif.

Since their proposal (Kauffman 1969), random Boolean networks (RBN) have successfully described in a qualitative way several important aspects of gene regulation and cell differentiation processes (Kauffman 1993, 1995). Discrete GRN grounded on experimental data have been inspired in the RBN, but in experimentally grounded GRN genes are not connected by chance as in a RBN. Recent data are suggesting that the connectivity of real GRN obeys a power law (Aldana and Cluzel 2003). The type of discrete GRN models that we are developing consist of a set of discrete variables, $\epsilon_1, \epsilon_2, \dots, \epsilon_N$, each acquiring a discrete value corresponding to the two states of gene expression. The state of each gene is regulated by a set of K other genes, $\epsilon_{n1}, \epsilon_{n2}, \dots, \epsilon_{nK}$, according to logical rules that are grounded on experimental data (see example in Espinosa-Soto and others 2004). The network dynamics are then given by:

$$\epsilon_n(t+1) = f_n(\epsilon_{n1}(t), \dots, \epsilon_{nK}(t)),$$

where f_n is a logical rule associated with the n th element of the network. These rules are derived from experimental data.

A particular GRN has a large number of possible gene activity configurations or states, and each one always leads to another state or to itself, in which case that state is a point attractor (steady state). Eventually all states, the Ω space, will reach an attractor. A system may have more than one attractor. The set of all states that lead to a specific attractor is the basin of that attractor. The basins of attraction and attractors of a GRN depend on the number of elements, the number of possible states of each element, the topology of the network, and on the logical rules of each gene (Figure 2). Consequently, the dynamics of the system are deterministic, a condition in which the fates of all states are known.

Gene Regulatory Network Models for *Arabidopsis thaliana* (L.) Heynh. Cell Type Determination during Organ Type Specification

We are interested in understanding the GRN structure and dynamics underlying cell-fate determination during floral organ specification. We have mainly focused on *Arabidopsis thaliana*, for which most data are available. On a first approach, Men-

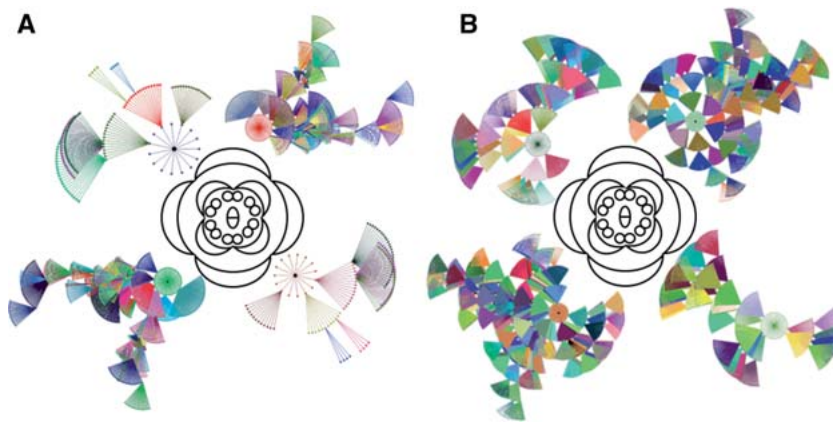


Figure 2. Basins of attraction and point attractors of the four floral primordial cell types (sepal, petal, stamen, and carpel) around the flower diagram for *A. thaliana*. **(A)** Boolean network model. **(B)** Three-state network model. The basins of attraction end in the point attractors that correspond to gene expression profiles that characterize sepal (lower right), petal (upper left), stamen (lower left), and carpel (upper right) primordial cells in both **A** and **B**. Note that the state space is much larger in the second case.

doza and Alvarez-Buylla (1998) proposed a GRN dynamic model with 11 nodes that included ABC and non-ABC genes that attained six attractors that corresponded to the genetic activities proposed in the ABC model. However, in this first model there was circularity in the interaction weights estimates, and some attractors did not correspond to observed gene activity combinations. Mendoza and collaborators (1999) then analyzed the dynamic behavior of the same model in terms of positive and negative feedback circuits. Interestingly, they recovered the same attractors as the ones found in their first approach, and they were able to identify the functional positive feedback loops underlying such attractors. However, once again interaction parameters were not grounded on experimental data. Nonetheless, this latter approach shows that positive regulatory circuits play a central role in cell differentiation and that modularity that may be described in terms of functional loops is a central characteristic of GRN (Thieffry and Sánchez 2004). It is also useful to detect possible missing nodes and/or connections that could form additional functional positive loops with nodes within the GRN or in other GRN that could explain observed gene configurations or cell behaviors (Thomas 1991).

Recently, in Espinosa-Soto and collaborators (2004), we integrated all available data on ABC and non-ABC genes in a GRN model of 15 nodes that significantly improves previous studies on GRN modeling for flower organ specification in our laboratory. The GRN was modeled as a three-state discrete model to capture all the experimental data at hand, and the problem of circularity was avoided by explicitly establishing the logical rules based on the experimental data available. The updating was assumed synchronous, making it possible to avoid subjective interpretations concerning the temporal order with which genes are actualized. Of the total number of initial combinations (139,968), the net-

work attained only 10 states, which coincide with the gene-expression profiles documented experimentally in cells of *A. thaliana* inflorescence meristems and floral organ primordia, and it also reproduces mutants' gene-expression patterns.

These data suggest that the gene regulatory model incorporates the key components of a developmental module or subnetwork that underlies the combinatorial gene activities predicted in the ABC model. Thus it provides a dynamical explanation for the ABC model of floral organ determination, showing that precise signaling pathways are not required to restrain organ primordia cell types during *A. thaliana* flower development. Our model suggests that these are rather determined by overall gene network topology and dynamics. This was confirmed by robustness analyses of random perturbations of gene-interaction parameters (Espinosa-Soto and others 2004).

One value of the approach that we pursue is that it provides an opportunity to integrate all previously published experimental evidence on particular genetic interactions, and this might aid in detecting contradictions and holes in the experimental evidence, as well as guiding functional genomic approaches. Although our GRN studies are not historical accounts of all work that has been done, but rather a formal qualitative model of the behavior of the gene network that regulates floral organ cell differentiation, the work should nonetheless incorporate the critical evidence for each interaction. Since we published the Espinosa-Soto and others (2004) article we have detected a few instances in which some key research papers that first documented a particular interaction were not cited, and we update these cases here. We have also detected a few minor errors in the article of Espinosa-Soto and others (2004). We list the additions and corrections in Appendix 1 and show that these do not affect the model or the study results in any way.

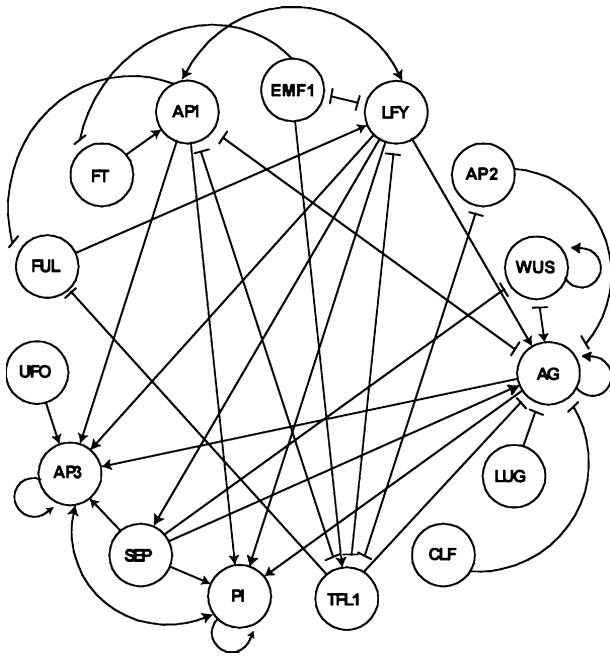


Figure 3. Updated floral network of *A. thaliana*. Nodes correspond to genes or proteins, and edges correspond to regulatory interactions (arrows are positive and blunt-end are negative). See also Espinosa-Soto and others (2004).

Updated Floral Network Structure and Dynamics: Predictions Met and New Interactions

We present here an updated GRN for cell-fate determination during floral organ specification (Figure 3). Two recent experimental findings confirmed two predictions made by our analyses (Espinosa-Soto and others 2004). A publication that came out during the review process of our work showed that the *SEP1, 2, 3* genes are indeed activated by *LFY* (Schmid and others 2003). We had predicted that a floral meristem identity gene should activate the *SEP*, or an inflorescence identity gene should repress them. We had postulated that *TFL1* repressed them. Hence, we have now deleted this interaction and instead postulated that *LFY* activates *SEP* (Figure 3; detailed tables of logical rules are available on request). Second, we had predicted that *AG* should have a positive direct or indirect feedback loop, and this has been confirmed in a recent article (Gómez-Mena and others 2005).

New experimental evidence suggests that the activity of *AP2* is postranscriptionally repressed by a specific miRNA (Chen 2004), whose abundance depends on the activity of the DICER-like protein *HUA ENHANCER 1* (*HEN1*, Park and others 2002). However, even though *HEN1* is expressed

throughout the plant (Chen and others 2002), miRNA-dependent *AP2* repression seems to occur only in the inner floral whorls (Chen 2004). This suggests that a cofactor, expressed in the inner floral whorls, is required to give specificity to *HEN1*-dependent repression of *AP2*. One possibility is that this cofactor is downstream of *AG* activity, supporting the mutual repression by A and C genes, proposed in the original ABC model (Coen and Meyerowitz 1991). We tested this possibility in our model. Because *HEN1* is expressed throughout the plant, we can consider it to be permanently “on,” and hence it is only required to add, in the *AP2* logical rules, that the inactivity of *AG* is needed for *AP2* function. The only change we observed in this simulation was that in the carpel and stamen attractors *AP2* is “off.”

Because both systems—the one that considers *AP2* repression by *AG* and the one that does not—are so similar, we decided to work with the latter to avoid what may be unrealistic assumptions.

It has also been found that *FT* expression favors *SEP3* and *FUL* activity (Teper-Bamnolker and Samach 2005). However, these interactions could be mediated by genes such as *LFY* and *TFL1*, and because this explanation is more parsimonious, we decided to avoid these additional interactions.

Interestingly, the updated GRN behavior is almost identical to that of the previously published (Espinosa-Soto and others 2004) network. We obtained the same number of point attractors as in the previous network. All the basins of attraction are of the same size and contain the same set of initial conditions as in the previous GRN, except for the sepal and the petal-1 basins, in which eight states of the petal-1 basin now belong to the sepal basin.

Reduction of the Three-state to a Boolean GRN

The original network had three states for some genes: inactive (0), medium active (1), and active (2). We reduced this three-state GRN to a Boolean network. We performed this analysis as an additional robustness analysis and also to test if it is useful to incorporate intermediate levels of gene activity (see Appendix 2 for details in GRN reduction). Although it is obvious that the basins of attraction are larger and more complicated in the three-state than in the two-state GRN, interestingly, the same attractors are reached in both cases (Figure 2). We then addressed if both GRN respond in the same way to systematic perturbations of the logical rules.

Table 1. Results of Systematically Mutating each Output of All Logical Rules in the Three and Two State Networks

3 states		Number of point attractors recovered							Total number of perturbations
Gene	4	5	6	7	8	9	10		
<i>FUL</i>	0	0	2	4	0	0	12	18	
<i>FT</i>	1	0	1	0	0	0	0	2	
<i>API</i>	0	0	2	4	0	0	102	108	
<i>EMF1</i>	1	0	1	0	0	0	1	3	
<i>LFY</i>	0	0	2	3	0	1	102	108	
<i>AP2</i>	1	0	1	0	0	0	1	3	
<i>WUS</i>	0	0	0	1	2	0	9	12	
<i>AG</i>	0	0	0	4	0	0	5180	5184	
<i>TFL1</i>	0	0	2	3	1	0	66	72	
<i>PI</i>	0	0	0	0	0	2	970	972	
<i>SEP</i>	1	0	1	0	0	0	1	3	
<i>AP3</i>	0	0	0	0	0	4	1940	1944	
TOTAL	4	0	12	19	3	7	8384	8429	
2 states		Number of point attractors recovered							Total number of perturbations
Gene	4	5	6	7	8	9	10		
<i>FUL</i>	0	0	1	2	0	0	1	4	
<i>FT</i>	1	0	1	0	0	0	0	2	
<i>API</i>	0	0	1	2	0	0	13	16	
<i>EMF1</i>	1	0	1	0	0	0	0	2	
<i>LFY</i>	0	0	1	1	0	1	13	16	
<i>AP2</i>	1	0	1	0	0	0	0	2	
<i>WUS</i>	0	0	0	1	2	0	5	8	
<i>AG</i>	0	0	0	2	0	0	126	128	
<i>TFL1</i>	0	0	1	1	0	0	14	16	
<i>PI</i>	0	0	0	0	0	2	62	64	
<i>SEP</i>	1	0	1	0	0	0	0	2	
<i>AP3</i>	0	0	0	0	0	3	125	128	
Total	4	0	8	9	2	6	359	388	

The columns list the number of point attractors recovered of the total (last column) when a rule is mutated. No rule recovered 0, 1, 2, or 3 point attractors when mutated.

We mutated each of the outputs of all logical rules and followed the dynamics, counting the number of the original point attractors that were conserved after the perturbation (Table 1). Although the Ω space is quite different for both GRN, it is interesting to note that the qualitative behavior of the two networks is almost identical with respect to the number of point attractors recovered, and both GRN are quite robust. As expected, genes with more outputs are the most stable ones, whereas the genes with fewer outputs are the most sensitive ones. Although it makes functional sense that attractors are buffered against perturbations to most connected genes, these could also be older and hence have had more time to accumulate connections. Alternatively, robustness of the GRN to perturbations in these gene logical functions may have been selected for. As discussed in Espinosa-Soto and others (2004), it seems that the new

attractors are not very different from the original ones.

Themes and Variations in Eudicotyledoneous Flower Evolution

Our previous studies had suggested that the GRN that underlies cell-fate determination during floral organ specification constitutes a robust developmental module or subnetwork that may underlie the widely conserved basic floral plan of eudicotyledoneous flowering plants. Furthermore, comparative analyses of flower patterning and floral organ determination gene network components and interactions suggest that this subnetwork has persisted over long periods irrespective of the evolution of the rest and largely uncovered molecular components that interact with this subnetwork (Ferrario

and others 2004). These other components should underlie other, relatively more variable aspects of flower morphology (symmetry, organ number, organ color, etc). If this hypothesis is true, then the floral GRN should be conserved among most eudicotyledoneous species, and this could be tested experimentally.

The GRN model may also be a basis for explorations of evolutionary differences in network architecture between *A. thaliana* and other species. *Petunia hybrida* Vilm. is another experimental plant system in which the overall network of cell fate determination during flower organ development seems to be conserved with respect to *A. thaliana* (Ferrario and others 2004). Nevertheless, *P. hybrida* B-gene mutant phenotypes are different from those described in *A. thaliana* because these genes are duplicated in the former species. Interestingly, the *A. thaliana* GRN with duplicated B genes recovered the observed expression profiles observed in *P. hybrida* floral organ primordial cells under wild type and also when the node corresponding to the *P. hybrida* AP3 ortholog was simulated as mutated (Espinosa-Soto and others 2004). This result further validates the model and suggests that the GRN underlying floral organ cell-fate determination is robust also to gene duplications.

Models that couple the floral GRN dynamics in explicit spatial domains will be useful in exploring possible mechanisms underlying the divergent floral patterns and the role of ABC genes among basal angiosperms. Preliminary analyses of a single-cell stochastic version of the GRN model considered here suggest that the temporal sequence with which floral organ primordia fate is determined may result, at least in part, from the dynamics of the organ specification regulatory network studied in the face of stochastic noise (Alvarez-Buylla, Chaos, Cortés, Espinosa-Soto, and Padilla-Longoria, unpublished work).

An extension of this model will be useful to postulate hypotheses underlying the variations of floral organs and ABC gene expression patterns observed among angiosperms. Probably the most striking floral organ arrangement is that observed in the endemic heterotrophic species of the monocotyledons, *Lacandonia schismatica* E. Martínez & Ramos (Vergara-Silva and others 2003; Ambrose and others 2006), which has a homeotic inversion of stamens and carpels with the latter in the third rather than the fourth central whorl. Spatiotemporal models that explicitly consider the type of GRN model proposed here will be needed to address which mutations may underlie this unique floral phenotype.

Although 75% of angiosperms are eudicotyledoneous, ABC gene-expression patterns and floral organ identities and arrangements are more variable among the basal flowering plants. Recently, Kim and collaborators (2005) reported that (1) the expression pattern of *AG* (a C class gene) of basal angiosperms is quite similar to that observed in core eudicotyledoneous species, (2) *AP3* and *PI* (B class genes) homologs seem to have a wider spatial expression pattern in basal angiosperms than that observed in eudicotyledoneous species, and (3) the expression pattern of *API* (an A class gene) differs greatly from that observed in eudicotyledoneous plants. Coincidentally, the floral organs that are specified by the A and B genes also show the most divergent patterns with respect to those characteristic of the core eudicotyledoneous species. For example, in the basal angiosperm *Nuphar advena* (Aiton) W. T. Aiton, the outer perianth organs are green and sepal-like whereas the inner perianth organs are more like petals because they are brightly colored. However, another basal angiosperm, *Asimina longifolia* Kral., has a perianth that is well differentiated into sepals and petals (Kim and others 2005).

This leads to the question of whether GRN models represent useful frameworks for unraveling the molecular evolution of flower patterns. Interestingly, the basins of attraction of the reproductive organs that seem to be the most stable among angiosperms are also the largest and, in consequence, the most stable in our simulations. This could have a historical explanation, and it could imply that the genes and/or interactions involved in the development of stamens and carpels evolved first, whereas those underlying sepal and petal specification arose later during evolution. The variation observed among basal angiosperms can reveal the unstable footsteps of the assemblage of the floral organ specification GRN that occurred during the evolution of the early lineages represented by extant basal angiosperms. The use of formal GRN models may help explore these hypotheses.

Our formal and dynamic GRN approaches suggest that the emergence of structures could have been the result of relatively slight modifications in GRN. New floral organs may have arisen as additional genes were added to a primitive network that originally yielded only reproductive structures. Data from gymnosperms (Sundstrom and Engstrom 2002) does indeed suggest that part of the angiosperm floral organ GRN could have been present before the origin of flowering plants. Adding a gene to the GRN may multiply the attractors and enlarge the potential adaptive landscape. The origin of

sepals and petals could have been the result of such a process, and this could have expanded the morphospace on which natural selection could act, yielding new evolutionary possibilities. If our rationale holds, perturbation analyses of an explicitly spatiotemporal model of coupled GRN could reveal which are the most stable steady states of the network and hence the probable primitive core within it that has remained unchanged during long evolutionary times. Such an approach not only allows for much more flexibility but also seems more realistic in evolutionary terms than postulating that the emergence of a new structure has relied on addition of genes from A to Z endlessly. It also allows for more specific predictions that could be first tested *in silico* and then *in vivo* by pursuing comparative analyses of key genes within the GRN in species selected according to their phylogenetic position in the tree of angiosperm evolution.

Cell identity attainment can be conceptualized as the product of a series of developmental decisions in which alternative gene activity patterns are acquired. The ABC model succeeded as a mechanistic proposal of floral organ specification because it showed that floral organ cell identity could be interpreted as the outcome of two developmental decisions: A or C gene activity, and B gene activity or inactivity (hereafter “ABC decisions”). Despite new discoveries on genes involved in floral morphogenesis and their interactions, the logic behind the ABC model remains unaltered. We now know that, in wild-type plants, the ABC decisions depend on an earlier decision: the floral or non-floral meristem decision (F-NF decision), in which genes such as *LFY* and *TFL1* are involved. A consequence of the F-NF decision is the upregulation of the *SEP* genes, because *LFY* promotes their expression (Schmid and others 2003). In later events, on the other hand, ovule identity genes, such as *SEEDSTICK* (Pinyopich and others 2002) seem to require C class gene activity as a necessary condition, because C mutants lack ovules. Hence, a hierarchical series of decisions in which some can only be taken if there is a certain outcome in a previous one seems to account for floral cell identity attainment. Therefore, as data accumulate, it is becoming clear that the developmental decisions during flower development considered in the ABC model hold, despite the fact that additional earlier and later developmental decisions have been discovered.

On the other hand, the inclusion of factors that provide positional information and that bias the ABC decisions, such as *UFO* (Parcy and others 1998) and *WUS* (Lenhard and others 2001; Lohman and others 2001), suggest that these genes may be key

inputs for determining the specific time and site where the combinations of gene activities considered in the ABC model are established. Explicitly spatial models that consider this will be useful to understand the mechanisms underlying the observed spatiotemporal patterns of ABC gene expression. In this context, the ABC model promises to continue to be a valid and useful core framework for understanding floral organ pattern, development, and evolution

ACKNOWLEDGMENTS

We acknowledge R. Vicencio and A. Navarrete for support in various tasks while preparing this manuscript. Financial support was from PAPIIT (Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica), UNAM (Universidad Nacional Autónoma de México; IN 230002, IN212995 and IN221406), and CONACYT (Consejo Nacional de Ciencia y Tecnología; 41848-Q and 0435/B-1) grants to E.A.B; E.A.B. also was supported by a SFI International fellowship while completing this article, and M.A. received CONACYT grant P47836-F. A.C. and C.E.S. received PhD fellowships from Consejo Nacional de Ciencia y Tecnología and Universidad Nacional Autónoma de México.

Appendix I. Corrections of GRN in Espinosa-Soto, Padilla-Longoria and Alvarez-Buylla 2004

We list some errors detected in Espinosa-Soto and others (2004). The truth table for *FT* on page 2925 has two outcomes, 0 or 1, which defines *FT* as a binary gene; nevertheless, on the same page in the truth table of *API*, *FT* has three values which is a contradiction. A closer look at the later truth table reveals that for the first rule of *API* the state of *FT* has no effect on *API* because it can be any state. For the second and the third rules, the state of *FT* is the same, which is active. In addition, for the other rules (3, 4 and 5) the state of *FT* can be non-active or medium active. Therefore, to avoid that contradiction between the truth table of *FT* and its states defining the truth table of *API* we made some changes. The first rule remains the same. In the second and the third rules of *API*, we substituted the value of 2 by 1, which means it is active. As the other rules of *API* do not make a distinction between the state of no activity and the state with medium activity, we changed them by 0 which means no activity. This contradiction does not affect the results obtained in Espinosa-Soto and others (2004).

The truth table of *WUS* (page 2926) has three inputs, which are *WUS*, *AG*, and *SEP*. The gene network architecture on the same page depicts that the inputs of *WUS* are just *SEP* and *AG*, so it needs a connection between *WUS* and itself to be coherent with the truth table. The printed truth table of *PI* in Espinosa-Soto and others (2004) had a few mistakes and here we present the correct truth table of *PI* (Table A1). In the two latter cases the correct truth tables shown here were used in the simulations supporting the results of Espinosa-Soto and others (2004).

In the logical rules for *LFY* the evidence for the fact that in *ap1* and *ap1cauliflower (cal)* double mutants, *LFY* expression is reduced, Bowman and others (1993) besides Piñeiro and Coupland (1998) should be cited. In the logical rules for *API* the evidence for the direct activation of *API* by *LFY* that binds to the *API* promoter was first documented by Parcy and collaborators (1998) and then by Wagner and collaborators (1999). On *APETALA3* logical rules, the evidence for the fact that in *lfy ap1* double mutants, *AP3* mRNA is not detected, but in *ap1* mutants, *AP3* is expressed as in wild-type plants Weigel and Meyerowitz (1993) as well as Hill and collaborators (1998) should be cited. On *PI* logical rules, the evidence that the level of *PI* mRNA and its domain of expression are severely reduced in *lfy-6* null mutants is found in Weigel and Meyerowitz (1993) as well as in Honma and Goto (2000). On *AGAMOUS* logical rules, the evidence that in strong *ap2* mutant alleles, *AG* RNA accumulates in the four floral whorls, and in *ap2-2* and *ap2-9* alleles, 1st whorl organs are converted into carpels or carpeloid sepals, and 2nd whorl organs are absent was first presented in Bowman and collaborators (1989, 1991) and later by Deyholos and Sieburth (2000). Also, support for the fact that *AG*, *PI*, *AP3*, *WUS*, and others are transcription factors is also found in Parcy and collaborators (1998) and Lohmann and collaborators (2001) as well as in Coen and Meyerowitz (1991) and Mayer and collaborators (1998).

Appendix II. Reduction of the Three-state to a Boolean GRN

To perform this reduction we had to decide to which state, active or inactive, the medium active state corresponded. To translate each rule for all truth tables we proceeded as follows: if the state was 2 it was changed to 1, if it was 0 it remained in that state. For some rules a combination of two states, gave the same output. Those cases were combina-

Table A1. Correct Truth Table of *PI*

<i>API</i>	<i>LFY</i>	<i>AG</i>	<i>PI</i>	<i>SEP</i>	<i>AP3</i>	<i>PI</i>
0	0	0	x	x	x	0
0	0	1	0	x	x	0
0	0	1	1	0	x	0
0	0	1	1	1	0	0
0	0	1	1	1	1	1
0	1	0	x	x	0	0
0	1	0	x	x	1	1
0	1	1	x	x	x	1
1	0	0	0	x	x	0
1	0	0	1	0	x	0
1	0	0	1	1	0	0
1	0	0	1	1	1	1
1	0	1	0	x	x	0
1	0	1	1	0	x	0
1	0	1	1	1	0	0
1	0	1	1	1	1	1
1	1	0	x	x	0	0
1	1	0	x	x	1	1
1	1	1	x	x	x	1

An "x" means any state (that is, 0 or 1).

tions of 0 and 1, and 1 and 2. In such cases, the first one was translated to 0 because there was no distinction between inactivation and intermediate activation, and in the second one it was translated to 1 because the rule gave the same output when the input was either 1 or 2. If the state was just 1 we assigned an uncertainty indicator because we could not assign a state of activity (1) or inactivity (0). For all possible inputs of each gene (for example, for a gene with three input genes there are 2^3 different inputs), we assigned an output according to the translated rules (that is, the rule that corresponds to that specific combination of inputs). Because of the reduction of the states it is possible that a rule governs an input that is already governed by other rules. In these cases there are two possibilities. Overlapping rules can be either redundant, if all rules give the same outputs, or contradictory, if some of all the overlapping rules yield different outputs. If the outputs of all the rules for a particular combination of inputs are unassigned, we call it a total ambiguity. If the output of one rule has a defined state, but one rule governing the same combination of inputs has an undefined state, we took the output of the defined rule and we call this case a partial ambiguity.

The total number of combinations of inputs of all genes from the original network with three states was 5500, whereas the Boolean network has only

772 combinations of inputs for all genes. For example, *FUL* was regulated by two genes with three entries for each in the three-state case, that yielded 9 different combinations of inputs. In contrast, in the Boolean translation, *FUL* is still regulated by the same two genes, but now these two have only two states each, yielding four combinations.

After assigning all states of all genes as mentioned, we found 96 total ambiguities and only 1 contradiction. We tested several combinations for these 97 problematic cases to try to recover the same behavior of the network with 3 states, starting with all of them with the state of inactivity (0), continuing with different combinations of 0 and 1, and finally assigning all of them a state of activity (1). When we assigned to all the problematic states the value of activity, we obtained exactly the same dynamics as for the original network: 10 point attractors with the same pattern of gene activity.

This can be understood in two ways. The observation that some products of certain genes differ in concentrations was the cause of the abstraction of the 3 states. If a particular threshold to activate a gene is surpassed, then the gene will enter in a state of activation. As we do not know the activation threshold of the genes, it could happen that regardless there is a marked difference in the concentration of the products of the genes; in both cases, medium activation and activation, the concentration surpasses the specific gene activation threshold. Consequently, we will obtain the same dynamics with the three-state network as that from the two-state network. On the other side, it is possible that the observed difference in gene product quantities is important and meaningful in the context of the system, but the robustness of the architecture network is sufficient to tackle this alteration of the parameters in gene product concentrations and to perform a normal behavior. This approach is a robustness analysis because the network is robust to this magnitude of change.

REFERENCES

- Acar M, Becskei A, Oudenaarden A Van . 2005. Enhancement of cellular memory by reducing stochastic transitions. *Nature* 435:228–232.
- Aracena J, Demongeot J. 2004. Mathematical methods for inferring regulatory networks interactions: application to genetic regulation. *Acta Biotheor* 52:391–400.
- Albert R, Othmer HG. 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.
- Aldana M, Cluzel P. 2003. A natural class of robust networks. *Proc Natl Acad Sci USA* 100:8710–8714.
- Ambrose BA, Espinosa-Matías S, Vázquez-Santana S, Vergara-Silva F, Martínez E, others . 2006. Comparative developmental series of the mexican triurids support a euanthial interpretation for the unusual reproductive axes of *Lacandonia schismatica* (Pandanales, Liliopsida). *Am J Bot* 93:15–35.
- Blake W, Kaern M, Cantor C, Collins J. 2003. Noise in eukaryotic gene expression. *Nature* 422:633–637.
- Bowman JL, Smyth DR, Meyerowitz EM. 1989. Genes directing flower development in *Arabidopsis*. *Plant Cell* 1:37–52.
- Bowman J, Smyth D, Meyerowitz E. 1991. Genetic interactions among floral homeotic genes of *Arabidopsis*. *Development* 112:1–20.
- Bowman J. 1994. *Arabidopsis—an Atlas of Morphology and Development*. Berlin Heidelberg New York, NY, USA: Springer-Verlag p 480.
- Bowman JL, Alvarez J, Weigel D, Meyerowitz EM, Smyth D. 1993. Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. *Development* 119:721–743.
- Chen X, Liu J, Cheng Y, Jia D. 2002. HEN1 functions pleiotropically in *Arabidopsis* development and acts in C function in the flower. *Development* 129:1085–1094.
- Chen X. 2004. A microRNA as a translational repressor of *APETALA2* in *Arabidopsis* flower development. *Science* 303:2022–2025.
- Coen ES, Meyerowitz E. 1991. The war of the whorls: genetic interactions controlling flower development. *Nature* 353:31–37.
- Deyholos MK, Sieburth LE. 2000. Separable whorl-specific expression and negative regulation by enhancer elements within the *AGAMOUS* second intron. *Plant Cell* 12:1799–1810.
- Elowitz M, Levine A, Siggia E, Swain P. 2002. Stochastic gene expression in a single cell. *Science* 297:1183–1186.
- Espinosa-Soto C, Padilla-Longoria P, Alvarez-Buylla E. 2004. A gene regulatory network model for cell-fate differentiation during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles. *Plant Cell* 16:2923–2939.
- Ferrario S, Immink RG, Angenent GC. 2004. Conservation and diversity in flower land. *Curr Opin Plant Biol* 7:84–91.
- Fiering S, Whitelaw E, Martin D. 2000. To be or not to be active: the stochastic nature of enhancer action. *BioEssays* 22:381–387.
- Gómez-Mena C, de Folter S, Costa MMR, Angenent GC, Sablowski R. 2005. Transcriptional program controlled by the floral homeotic gene *AGAMOUS* during early organogenesis. *Development* 132:429–438.
- Hill TA, Day CD, Zondlo SC, Thackeray AG, Irish VF. 1998. Discrete spatial and temporal cis-acting elements regulate transcription of the *Arabidopsis* floral homeotic gene *APETALA3*. *Development* 125:1711–1721.
- Ho S, Biggar S, Spencer D, Schreiber S, Crabtree G. 1996. Dimeric ligands define a role for transcriptional activation domains in reinitiation. *Nature* 382:822–826.
- Honma T, Goto K. 2000. The *Arabidopsis* floral homeotic gene *PISTILLATA* is regulated by discrete cis-elements responsive to induction and maintenance signals. *Development* 127:2021–2030.
- Huang S, Eichler G, Bar-Yam Y, Ingber DE. 2005. Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Physical Rev Lett* 94:128701–128704.
- Hume D. 2000. Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. *Blood* 96:2323–2328.
- Kauffman SA. 1969. Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* 22:437–467.

- Kauffman SA. 1993. *The Origins of Order: Self-organization and Selection in Evolution* Oxford, England, UK: Oxford University Press.
- Kauffman SA. 1995. *At Home in the Universe*. New York, NY, USA, Oxford University: Press.
- Kim S, Koh J, Yoo M, Kong H, Hu Y, others . 2005. Expression of floral MADS-box genes in basal angiosperms: implications for the evolution of floral regulators. *Plant J* 43:722–744.
- Lenhard M, Bonhart A, Jurgens G, Laux T. 2001. Termination of stem cell maintenance in *Arabidopsis* floral meristems by interactions between *WUSCHEL* and *AGAMOUS*. *Cell* 105:805–808.
- Lewontin R. 1974. *The Genetic Basis of Evolutionary Change*. New York, NY, USA: Columbia University Press, p 352.
- Lohmann JU, Hong RL, Hobe M, Busch MA, Parcy F, others. 2001. A molecular link between stem cell regulation and floral patterning in *Arabidopsis*. *Cell* 105:793–803.
- Mayer KFX, Schoof H, Haecker A, Lenhard M, Jürgens G, others. 1998. Role of *WUSCHEL* in regulating stem cell fate in the *Arabidopsis* shoot meristem. *Cell* 95:805–815.
- Mendoza L, Alvarez-Buylla ER. 1998. Dynamics of the genetic regulatory network of *Arabidopsis thaliana* flower morphogenesis. *J Theor Biol* 193:307–319.
- Mendoza L, Thieffry D, Alvarez-Buylla ER. 1999. Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis. *Bioinformatics* 15:593–606.
- Meyerowitz E. 2002. Plants compared to animals: the broadest comparative study of development. *Science* 295:1482–1485.
- Ozbudak EM, Thattai M, Kurtser I, Grossman AD, Oudenaarden A Van . 2002. Regulation of noise in the expression of a single gene. *Nat Genet* 31:69–73.
- Parcy F, Nilsson O, Busch MA, Lee I, Weigel D. 1998. A genetic framework for floral patterning. *Nature* 395:561–566.
- Paulsson J. 2004. Summing up the noise in gene networks. *Nature* 427:415–418.
- Park W, Li J, Song R, Messing J, Chen X. 2002. *CARPEL FACTORY*, a Dicer homolog, and *HEN1*, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol* 12:1484–1495.
- Perkins T, Hallet M, Glass L. 2004. Inferring models of gene expression dynamics. *J Theor Biol* 230:289–299.
- Piñeiro M, Coupland G. 1998. The control of flowering time and floral identity in *Arabidopsis*. *Plant Physiol* 117:1–8.
- Pinyopich A, Ditta GS, Savidge B, Liljgren SJ, Baumann E, and others. 2002. Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* 424:85–88.
- Rossi FMV, Kringstein AM, Spicher A, Guicherit OM, Blau HM. 2000. Rheostat mechanism of transcriptional regulation converted to ON/OFF switch. *Mol Cell* 6:723–728.
- Salazar-Ciudad I, García-Fernández J, Solé RV. 2000. Gene networks capable of pattern formation: from induction to reaction-diffusion. *J Theor Biol* 205:587–603.
- Schmid M, Uhlenhaut NH, Godard F, Denmar M, Bressan R, and others. 2003. Dissection of floral induction pathways using global expression analysis. *Development* 130:6001–6012.
- Sundstrom J, Engstrom P. 2002. Conifer reproductive development involves B-type MADS-box genes with distinct and different activities in male organ primordial. *Plant J* 31:161–169.
- Teper-Bamnolker P, Samach A. 2005. The flowering integrator FT regulates *SEPALLATA3* and *FRUITFULL* accumulation in *Arabidopsis* leaves. *Plant Cell* 17:2661–2675.
- Thieffry D, Sánchez L. 2004. Qualitative analysis of gene networks: toward the delineation of cross-regulatory modules, pp. 222–243 In: Schlosser G, Wagner G, editors. *Modularity in Development and Evolution*. Chicago, IL, USA, The University of Chicago Press, 600 p.
- Thomas R. 1991. Regulatory networks seen as asynchronous automata: a logical description. *J Theor Biol* 153:1–23.
- Thomas R, Thieffry D, Kaufman M. 1995. Dynamical behaviour of biological regulatory networks—I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. *Bull Math Biol* 57:247–276.
- Vergara-Silva F, Espinosa-Matías S, Ambrose BA, Vázquez-Santana S, Martínez-Mena A, Márquezand others. 2003. Inside-out flowers characteristic of *Lacandonia schismatica* evolved at least before its divergence from a closely related taxon, *Triuris brevistylis*. *Int J Plant Sci* 164:345–357.
- von Dassow G, Meir E, Munro EM, Odell GM. 2000. The segment polarity network is a robust developmental module. *Nature* 406:188–193.
- Wagner D, Sablowski RWM, Meyerowitz EM. 1999. Transcriptional activation of *APETALA1* by *LEAFY*. *Science* 285:582–584.
- Walters MC, Fiering S, Eidemiller J, Magis W, Groudine M, and others. 1995. Enhancers increase the probability but not the level of gene expression. *Proc Natl Acad Sci USA* 92:7125–7129.
- Weigel D, Meyerowitz EM. 1993. Activation of floral homeotic genes in *Arabidopsis*. *Science* 261:1723–1726.

Adenda

Como apoyo gráfico al artículo se presentan aquí unos análisis que se hicieron posteriormente gracias a que se terminó de escribir la parte gráfica del programa *Atalia*, con el cual se hicieron esos análisis, que ahora incorpora rutinas y gráficas que simplifican el estudio y la interpretación de los resultados.

En el análisis original, el trabajo sólo registró el número de sumideros originales que se recuperaban cuando se mutaban las salidas de las combinaciones de las reglas lógicas, estos resultados aparecen en la Tabla 1 del artículo citado. Gracias a las rutinas nuevas del programa *Atalia* es posible representar eso de manera gráfica y de profundizar el análisis, no sólo registrando el número de sumideros conservados, sino que además es posible saber si se producen más sumideros. Entonces se pueden obtener cuatro casos cuando se hacen las mutaciones. El primero de ellos se da cuando al mutar la salida se obtienen los mismos 10 sumideros, ni más ni menos, a este régimen lo denominaremos neutro. El segundo caso, que llamaremos oportunista, será cuando después de mutar la salida se produzcan los 10 sumideros originales y además sumideros nuevos. Es posible que a continuación de hacer la mutación se pierdan sumideros originales y que no se creen nuevos, este régimen es el reaccionario. Si en cambio, además de perder sumideros originales, se ganan sumideros nuevos, será el régimen intrépido. Bajo este análisis con cuatro regímenes o fases en los que se clasifican los efectos de las mutaciones se detalla mucho mejor y con más profundidad las consecuencias de hacer cierto cambio. Los resultados de este análisis se presentan en la Figura 11.

Es posible representar esta información de otra manera mucho más detallada. En la Figura 12 se presentan los porcentajes de sumideros recuperados cuando se muta la salida de cierta combinación. Como se puede apreciar, de las 388 salidas que se mutaron, en la mayoría de ellas se recuperan los 10 sumideros originales.

Como se mencionó anteriormente, que no se recuperen los 10 sumideros originales no implica que el sistema tenga una cantidad menor de sumideros. Para apreciar eso se presenta la Figura 13, en ella se puede ver que hay casos en los que se pueden producir hasta 25 sumideros (la salida número 33 general que corresponde a la salida 9 del gen *LFY*) en los cuales no están todos los originales, por lo tanto, esa es una mutación intrépida.

Los análisis presentados aquí toman en cuenta que la red se actualiza de forma sincrónica, es decir, que todos los genes cambian al mismo tiempo; sin embargo, también es posible hacer las simulaciones de forma asíncrona. Tal vez sea más real pensar que los genes no se actualizan todos al mismo tiempo, el problema para hacer una simulación asíncrona consiste en que se deben de tener datos sobre este orden y desgraciadamente no se dispone de ellos todavía. La opción de probar todas las combinaciones es irreal debido al número astronómico de ensayos que se tendrían que hacer. Para sondear como afectaría esta escalada consecutiva de activaciones sobre los resultados obtenidos aquí, se realizaron algunas simulaciones tomando en cuenta lo poco que se sabe sobre el orden de cambio de los genes. Se efectuaron un par de pruebas para sondear los efectos de la asincronía de acuerdo con datos muy incompletos.

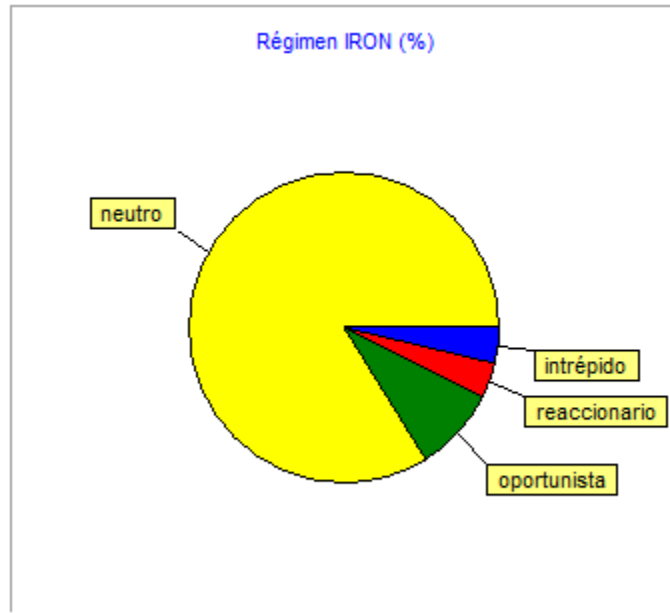


Figura 11: Las fases o regímenes posibles (régimen IRON) que resultan de hacer mutaciones en las salidas de las combinaciones de las reglas lógicas. Esta gráfica corresponde a la red de *A. thaliana*. De las 388 mutaciones posibles, 325 son neutras (83.76%), 34 son oportunistas (8.76%), 15 son reaccionarias (3.87%) y 14 son intrépidas (3.60%).

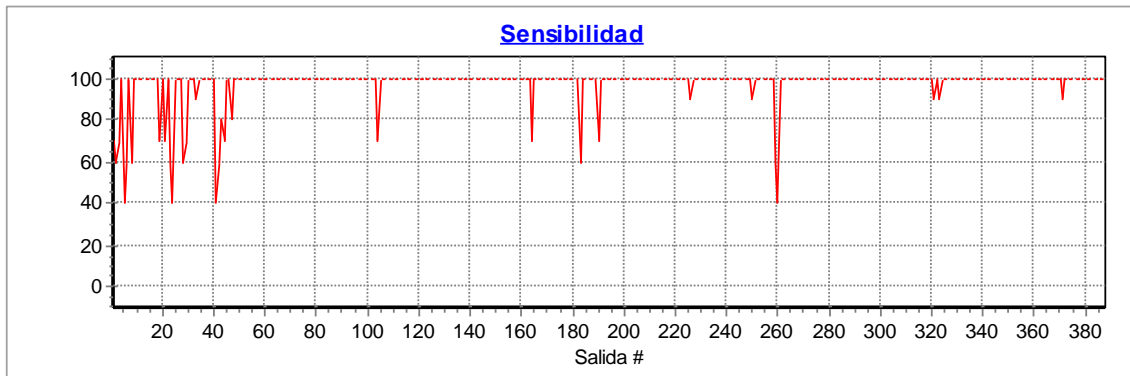


Figura 1: Los porcentajes de sumideros originales que se mantienen después de hacer un cambio en cada una de las 388 combinaciones o salidas de las reglas lógicas de la red de *A. thaliana*.

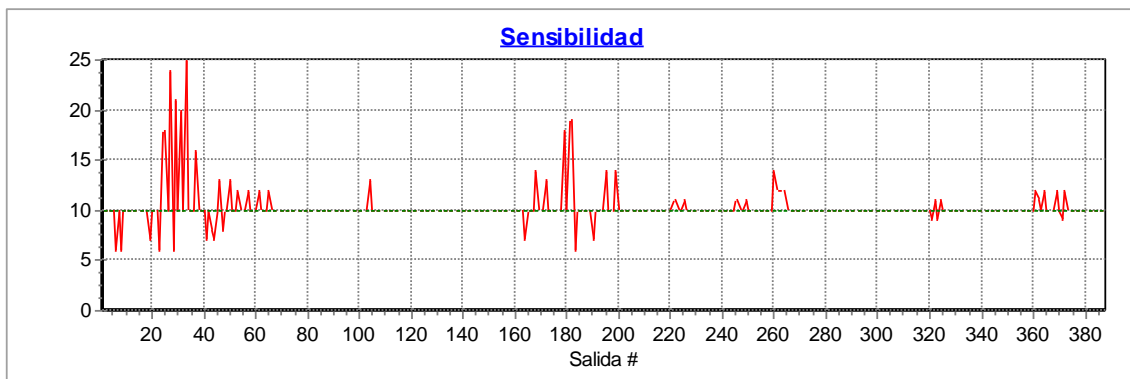


Figura 13: Los números totales de sumideros que se originan después de hacer un cambio en cada una de las 388 combinaciones o salidas de las reglas lógicas de la red de *A. thaliana*.

En la primera de ellas se establecieron tres grupos de actualización, el primero formado por tres genes (*FT*, *EMF1* y *TFL1*), el segundo por seis genes (*FUL*, *API*, *LFY*, *AP2*, *WUS* y *AG*) y el último por cuatro genes (*PI*, *SEP*, *AP3* y *UFO*). Para la segunda prueba solamente se cambió al gen *AG* (el único gen del grupo C), se puso en un cuarto grupo de actualización. En ambas pruebas se obtuvieron exactamente los mismos diez sumideros. Donde sí hubo cambios fue en el tamaño de las cuencas de atracción y en sus estructuras, de todas formas, las diferencias fueron mínimas. En la Tabla 8 se presentan estos resultados. Sin lugar a dudas, también son una prueba de robustez de la red.

Tabla 1: Comparación de los tamaños de las cuencas de atracción con sincronía y dos casos de asincronía.

Configuración	Célula	Tamaño de las cuencas de atracción		
		Síncrona	Asíncrona 1	Asíncrona 2
0001000010000	Inflorescencia	128	160	128
0001000010001	Inflorescencia	128	160	96
0001001010000	Inflorescencia	64	96	64
0001001010001	Inflorescencia	64	96	96
0110110000100	Sépalos	152	152	398
0110110001110	Pétalos	8	8	16
0110110001111	Pétalos	160	160	160
1100110101110	Estambres	136	136	176
1100110101111	Estambres	3744	3680	3743
1100110101100	Carpelos	3608	3544	3315

Capítulo 3

Morfogénesis floral: exploraciones estocásticas del paisaje epigenético de una red genética.

El modelo de redes genéticas binarias explicado en capítulos previos ha demostrado ser un buen método para entender cómo la acción concertada de un grupo de genes conectados entre sí regula la diferenciación celular, en este caso de la flor, y de esta forma es posible abstraer el funcionamiento de los seres vivos. Es un modelo determinista, lo que significa que una vez establecida una configuración génica el camino a seguir será siempre el mismo. De esta manera es imposible transitar entre los diferentes sumideros, es decir, si una célula comienza en una configuración que corresponde a la cuenca de atracción de los sépalos, irremediablemente acabará en el sumidero que corresponde a los sépalos y dicha célula tendrá la identidad de un sépalo. Esto genera dos explicaciones con respecto a la diferenciación de los órganos de la flor.

La primera de ellas es que las células no cambian de identidad, por lo que no transitan entre las configuraciones de las diferentes cuencas de atracción. Lo que hace que se formen los diferentes verticilos florales es que a partir de un cúmulo de células indiferenciadas una parte de ellas acabará teniendo la identidad de sépalos, posteriormente otra fracción se diferenciará en pétalos y así sucesivamente en estambres y en carpelos hasta que se halla agotado el meristemo floral. Lo que faltaría saber es cuál es la señal y en dónde se origina la directriz de esta cascada de eventos de diferenciación celular.

La segunda consiste en que las células que transitan por ser sépalos reciben algún estímulo interno o externo lo cual hace que cambien de cuenca de atracción generando la formación, después de los sépalos, de los demás verticilos florales: pétalos, estambres y carpelos. De esta manera, el cúmulo inicial de células indiferenciadas podría ir adquiriendo diferentes identidades transitorias mientras cada una de las células alcanza su destino final. En ambos casos se requiere de cierta señal que haga el cambio de cuenca. Existen dos maneras de generar la señal de cambio. La primera de ellas es una señal que venga de un gen externo a la red que provoque el cambio de configuración, esta señal estaría dictada por el desarrollo del organismo, y sería determinista también. La segunda explicación se relaciona con cualquier cambio inducido al sistema por estímulos externos o internos ajenos al mismo e inesperados y que dependen de la influencia del medio externo o del interno que provocan cambios aleatorios en la dinámica del sistema, bajo esta perspectiva, la señal que provocaría el cambio de cuenca de sumidero no sería una señal determinista y precisa, si no una aleatoria o ruidosa. La naturaleza de estos ruidos puede provenir de varias fuentes. El caso en el que se origina algún cambio en las reglas lógicas debido a cierta mutación en el ADN que haga que la célula responda de manera diferente a la previa, donde la alteración es permanente porque la célula cambia su comportamiento, no es un tipo de explicación alterna. Aunque en este caso el sistema cambia, volvería a comportarse de forma determinista, sólo que con una regla lógica cambiada, lo cual dejaría las cosas en el mismo caso que el anterior: un sistema determinista que necesita de una señal para poder transitar entre sus sumideros.

Es posible que los ruidos provengan del exterior también, pero que no ocasionen cambios perdurables sino momentáneos. Las fluctuaciones térmicas son un ejemplo de ello. Si el sistema es sensible a este tipo de variaciones se puede generar un error momentáneo en la actualización de las reglas lógicas, lo que ocasionaría un cambio en la dinámica normal y determinista del sistema y que se pudiese alcanzar un sumidero diferente al predestinado bajo condiciones normales. Otra fuente de ruido podría originarse dentro del mismo sistema, es decir, si el sistema muestra cierta imprecisión o ineficacia en su funcionamiento, las rutas que se seguirían normalmente se verían alteradas por eventos intrínsecos del sistema. Imagine que se necesita cierta cantidad de moléculas para el desempeño correcto del sistema, si por alguna incapacidad momentánea no se le pueden suministrar, éste alterará el destino normal de las configuraciones del espacio omega. En estos casos el cambio de sumideros se deberá al ruido ambiental, es decir, a mutaciones transitorias que perturban al sistema o a una imprecisión innata y momentánea del mismo.

En el estudio de este capítulo se aborda el caso en el cual se perturba al sistema de manera continua, por medio de cambios momentáneos, y aleatoria con diferentes niveles de alteraciones para luego determinar el efecto de éstos sobre el cambio de sumideros. Con base en este enfoque se podría averiguar si la red floral de *A. thaliana* sujeta a ruido da como resultado la secuencia temporal de transición por los sumideros que se observa en la naturaleza sin necesidad de suponer la presencia de una señal determinista. Aunque ésta podría existir y estar reforzada por el mismo sistema dinámico no lineal implicado en la red.

Las simulaciones se realizaron con el modelo binario de redes y con una implementación como un sistema continuo utilizando la aproximación de Glass que consiste en un sistema de ecuaciones diferenciales. Ambos modelos casi recuperan la secuencia observada en la naturaleza para ciertos intervalos de ruido (1-10%). La discrepancia radica en los dos verticilos últimos, ya que bajo las simulaciones del modelo discreto lo que se observa es que primero surgen los carpelos y después los estambres, cuando debería de ser al revés; sin embargo, el modelo continuo sí recupera esa secuencia solamente para un nivel de ruido (3%).

Según estos resultados es posible pensar que el ruido externo o el intrínseco del sistema que producen cambios momentáneos, pero constantes, son los causantes en realidad de la secuencia que se observa y ella no se debe a la selección natural. Como se mencionó arriba, es posible también que existan señales deterministas que coincidan con las aleatorias y que potencien los efectos, lo que ocasionaría que entren en concordancia con las primeras. De esta manera la señal determinista y la señal ruidosa podrían sincronizarse espontáneamente provocando un efecto que suma ambas señales y dando como resultado una secuencia de diferenciación más robusta. Otro aspecto interesante que se desprende de estos resultados es que de cierta manera se podría tener una idea de los niveles de ruido a los que está expuesta la red de diferenciación floral de *A. thaliana*. Por poner un ejemplo, si el ruido se debe exclusivamente a fenómenos intrínsecos relacionados con la fiabilidad del sistema, el origen de la flor estaría originado simplemente por la imprecisión del sistema, la flor perfecta no sería más que un producto colateral de un sistema genético que se desempeña imperfectamente.

Cabe aclarar aquí que las simulaciones no se hicieron con todos los sumideros. Sólo se utilizaron los de las estructuras florales y se dejaron afuera los cuatro sumideros de las inflorescencias. Esto se debió a que si se hacen las simulaciones con todos, el sistema brinca de los sumideros de las inflorescencias a los sumideros de las estructuras

reproductoras de la flor. Es por esto que no es posible descartar la existencia de señales deterministas en algún momento de la diferenciación de la estructura.

Otro punto importante en la discusión es el que se refiere a la permanencia de las células en un sumidero dado. Si las células estuviesen reaccionando al ruido constantemente no tendrían un destino final y estarían brincando de un sumidero a otro. Las transiciones entre los sumideros se deben de congelar en algún momento determinando el futuro final de las células en los diferentes tipos celulares. La señal que puede congelar o detener esta cascada de diferenciación podría depender de un número de células que funcione como un umbral y que una vez alcanzado, gracias a que la probabilidad de llegar al sumidero sea máxima, fungirá como el núcleo de células que tomará la identidad del verticilo en cuestión. Este núcleo o masa crítica de células variará en número según el sumidero del que se trate. Por ejemplo, para un sumidero de pétalos, cuya cuenca de atracción es de 160, la masa crítica de células consistirá de un número mucho menor que la masa crítica para el sumideros de los carpelos que tiene una cuenca de atracción de 3,608 configuraciones.

Estos resultados refuerzan los de robustez obtenidos en el Capítulo 2 así como los de otros trabajos (Von Dassow *et al.* 2000, Aldana y Cluzel 2003, Espinosa-Soto *et al.* 2004) ya que muestran que la forma típica en la que está construida una flor en términos de los cuatro tipos celulares básicos y de su secuencia temporal con la cual aparecen durante el desarrollo son robustas ante simulaciones estocásticas.

Los resultados obtenidos en este artículo y en el anterior son sin duda un principio para entender el desarrollo de una estructura como es la flor típica y serán un punto de partida para entender a otras estructuras de las que en un futuro podamos disponer de sus redes y de esta forma llegar a generalizaciones sobre la evolución de las criaturas. Pero no sólo se quedan allí, si no que también proporcionarán las bases para comprender a los famosos monstruos esperanzados, que en el caso de las flores, está representado por la *Lacandonia schismatica* Martínez y Ramos, único representante entre aproximadamente 257,000 plantas con flores que posee invertido el orden de los últimos verticilos, es decir, presenta como tercer verticilo a los carpelos y como último verticilo a los estambres que se ubican en el centro de la flor.

Floral Morphogenesis: Stochastic Explorations of a Gene Network Epigenetic Landscape

Elena R. Álvarez-Buylla^{1,2,3*}, Álvaro Chaos^{1,2,3}, Maximino Aldana^{2,3}, Mariana Benítez^{1,2}, Yuriria Cortes-Poza^{2,4}, Carlos Espinosa-Soto^{1,2}, Diego A. Hartasánchez^{2,3}, R. Beau Lotto⁵, David Malkin⁵, Gerardo J. Escalera Santos^{1,2}, Pablo Padilla-Longoria^{2,4*}

1 Instituto de Ecología, Universidad Nacional Autónoma de México, Cd. Universitaria, México, D. F., México, **2** C3, Centro de Ciencias de la Complejidad, Cd. Universitaria, UNAM, México, D. F., México, **3** Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México, **4** Instituto de Investigación en Matemáticas Aplicadas y Sistemas, Universidad Nacional Autónoma de México, Cd. Universitaria, México, D. F., México, **5** Iottolab, University College, London, United Kingdom

Abstract

In contrast to the classical view of development as a preprogrammed and deterministic process, recent studies have demonstrated that stochastic perturbations of highly non-linear systems may underlie the emergence and stability of biological patterns. Herein, we address the question of whether noise contributes to the generation of the stereotypical temporal pattern in gene expression during flower development. We modeled the regulatory network of organ identity genes in the *Arabidopsis thaliana* flower as a stochastic system. This network has previously been shown to converge to ten fixed-point attractors, each with gene expression arrays that characterize inflorescence cells and primordial cells of sepals, petals, stamens, and carpels. The network used is binary, and the logical rules that govern its dynamics are grounded in experimental evidence. We introduced different levels of uncertainty in the updating rules of the network. Interestingly, for a level of noise of around 0.5–10%, the system exhibited a sequence of transitions among attractors that mimics the sequence of gene activation configurations observed in real flowers. We also implemented the gene regulatory network as a continuous system using the Glass model of differential equations, that can be considered as a first approximation of kinetic-reaction equations, but which are not necessarily equivalent to the Boolean model. Interestingly, the Glass dynamics recover a temporal sequence of attractors, that is qualitatively similar, although not identical, to that obtained using the Boolean model. Thus, time ordering in the emergence of cell-fate patterns is not an artifact of synchronous updating in the Boolean model. Therefore, our model provides a novel explanation for the emergence and robustness of the ubiquitous temporal pattern of floral organ specification. It also constitutes a new approach to understanding morphogenesis, providing predictions on the population dynamics of cells with different genetic configurations during development.

Citation: Álvarez-Buylla ER, Chaos Á, Aldana M, Benítez M, Cortes-Poza Y, et al. (2008) Floral Morphogenesis: Stochastic Explorations of a Gene Network Epigenetic Landscape. PLoS ONE 3(11): e3626. doi:10.1371/journal.pone.0003626

Editor: Sui Huang, University of Calgary, Canada

Received: October 31, 2007; **Accepted:** October 5, 2008; **Published:** November 3, 2008

Copyright: © 2008 Álvarez-Buylla et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Financial support was provided by Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, Universidad Nacional Autónoma de México (UNAM) IN230002 and IX207104, University of California-MEXUS ECO IE 271 and Consejo Nacional de Ciencia y Tecnología (CONACYT) CO1.41848/ A-1, CO1.0538/A-1 and CO1.0435.B-1. Grants to E.A.B. PAPIIT-UNAM grant IN112407-3 y CONACYT grant and P47836-F grants to M.A. Ph.D. and postdoctoral scholarships from CONACYT and UNAM (CEP-UNAM) to C.E.-S, M.B., G.J. E. and A.C. These students were in the Doctorado en Ciencias Biológicas graduate program at the Instituto de Ecología from UNAM.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: elena.alvarezbuylla@gmail.com (ERA-B); pablo@mym.iimas.unam.mx (PP-L)

† These authors contributed equally to this work.

Introduction

“All [the] epistemological value of the theory of probability is based on this: That large scale random phenomena in their collective action create strict, non random regularity”. (From: B.V. Gnedenko and A.N. Kolmogorov, Limit Distributions for Sums of Independent Random Variables, Reading, Ma: Addison-Wesley, 1954).

The development of multicellular organisms consists of cell differentiation and spatiotemporal patterning. Since these processes arise from complex interactions among genetic and non-genetic elements, mathematical and computational models are useful to study the concerted action of these elements. Gene regulatory network (GRN) models, which are grounded in experimental data, have been able to recover fixed profiles of gene activation, that mimic those characterizing different cell types

in both plants and animals (e.g., [1–3]). Such profiles correspond to the attractors of these networks, and have been interpreted as cell fates [4–7].

Some studies have explored cell-fate decisions by modeling transitions among attractors with stochastic gene regulatory networks (e.g. [8,9]); however, models grounded in experimental data that are able to recover patterns of cell-fate attainment for a particular living system are only now starting to appear. Herein, we attempted to construct an integrative model driven by noise that explores the patterns of temporal cell-fate attainment in the experimental plant, *Arabidopsis thaliana* (L.) Heynh.

In plants, morphogenesis takes place during the entire life cycle from groups of undifferentiated cells called meristems. Within meristems, cell fate is mostly determined by position rather than by cell lineage [10]. Flower meristems are formed from the flanks of

the inflorescence meristem, which is found at the apex of an *Arabidopsis thaliana* plant once it has reached a reproductive stage (Figures 1A and B). Early in flower development, a floral meristem is sequentially partitioned into four regions, from which the floral organ primordia are formed and eventually give rise to sepals in the outermost whorl, then to petals in the second whorl, stamens in the third, and carpels in the fourth whorl in the central part of the flower (Figures 1B and C). This spatio-temporal sequence is widely

conserved among the quarter of a million flowering plant species [11]; however, the dynamic mechanisms underlying this robust pattern are not yet understood.

In this study, we used a previously characterized Boolean GRN, which converges to ten attractors (Figure 1), to explore the dynamics of cell-fate decisions during the early stages of flower development. The ten attractors correspond to the main cell types observed during early flower development, namely, meristematic

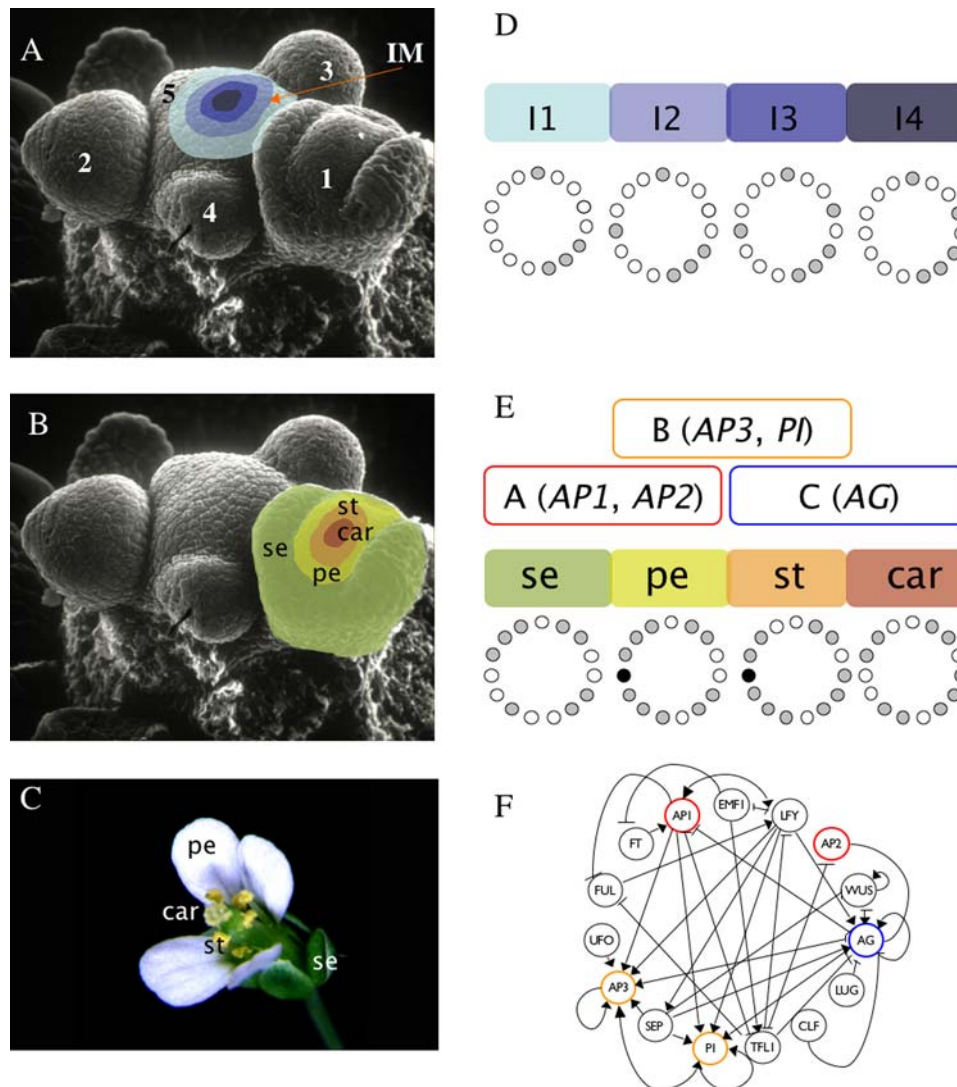


Figure 1. Flower development and gene network underlying primordial floral organ cell-fate determination in *Arabidopsis thaliana*.

(A) The inflorescence meristem (IM in the Scanning Electron Micrography) is found at the apex of a reproductively mature plant. Within the IM, four regions can be distinguished. Interestingly, the experimentally observed gene activation configurations of each one of these regions are mimicked by the I1, I2, I3, and I4 attractors of the 15-gene GRN. Flower meristems arise in a helical pattern from the flanks of the IM. The order in which floral meristems appear is indicated with numbers (1, oldest; 5, youngest). (B) Young flower meristems can be subdivided into four regions, each one containing the primordial cells that will eventually develop into the flower organs. In each floral meristem, the outermost region, which is first determined, will give rise to the sepal (se) primordium, the next to petals (pe) and finally, the primordial corresponding to stamens (st) and carpels (car) are determined in the center third and fourth whorls of the flower bud, respectively. (C) The mature flower of *Arabidopsis thaliana*. (D) I1, I2, I3, and I4 regions of the IM correspond to four of the attractors of the 15-gene GRN model. The expressed genes for each attractor are represented as gray circles, while the non-expressed genes correspond to white circles. (E) The other six attractors of the GRN model match gene expression profiles characteristic of sepal, petal (p1 and p2), stamen (st1 and st2), and carpel primordial cells. Black circles represent a gene (*UFO*) that can be either expressed or not expressed in the petal and stamen attractors, thus yielding two attractors for petal and stamen primordial cell-type. The gene activation profiles of the attractors recovered for the 15-gene GRN are congruent with the combinatorial activities of A, B, and C-type genes predicted by the ABC model of floral organ determination. See the Results section and [3,12] for details. (F) Gene regulatory network model underlying cell fate determination in the IM and the flower meristem. A-genes (red), B-genes (yellow), and C-genes (blue) from the ABC model are indicated in the network.

doi:10.1371/journal.pone.0003626.g001

cells of the inflorescence, which is itself partitioned into four regions (I1, I2, I3, and I4; Figures 1A and D), and sepal, petal (P1 and P2), stamen (S1 and S2), and carpel primordial cells within flower meristems (Figures 1B and E) [3,12]. This network was grounded in experimental data for 15 genes, wherein their interactions were formalized as logical functions. Among the 15 genes, five are grouped into three classes (A-type, B-type, and C-type), whose combinations are necessary for floral organ cell specification [13]. A-type genes (*API* and *AP2*) characterize sepal identity, A-type together with B-type (*AP3* and *P1*) petal identity, B-type and C-type (*AGAMOUS*) stamen identity, and the C-type gene (*AG*) alone for carpel primordia cell identity. The so-called ABC model describes such combinatorial activities during floral organ determination (Figures 1E and F) [13].

Different sets of initial conditions (basins of attraction) of the 15-gene regulatory network converge to the ABC-gene combinations necessary for floral organ determination [3,12] (Figures 1E and F); however, this deterministic GRN does not enable studies of the transitions among the attractors. In this study, we investigated the temporal sequence with which attractors are visited in this GRN when noise or random perturbations to the output of the updating rules drive the system from one attractor to any other.

The obtained results demonstrate that noise alone is able to drive transitions among attractors with temporal patterns that mimic the sequence with which ABC-genes are activated (first A genes, then B genes, and finally the C gene) during early flower development [13]. These results are in line with the finding that the GRN in question is a robust developmental module that is widely conserved among flowering plant species [3]. Furthermore, the temporal cell-fate pattern during early stages of flower development seems to emerge from such a robust network in the presence of noisy perturbations. The results presented herein support the idea that random fluctuations in a system may be important for physiological adaptation, plasticity, and cell differentiation (examples in: [14–24]).

Results

A stochastic Boolean model of the GRN enables the study of transitions among network attractors

We first present the results obtained from the Boolean model of the GRN, and in the next section, we present the equivalent results obtained from a continuous model. The Boolean approach focuses on the state of genes' expression rather than on the concentration of their products. Thus, each gene in the network is represented by a Boolean variable x that takes the value $x = 1$, if the corresponding gene is expressed, and the value $x = 0$, if it is not expressed. The state of expression of the genes in the entire network (herein, configurations of the GRN, which correspond to “dynamic state of the network” used by some authors), is then represented by a vector with the set of Boolean variables $\{x_1, x_2, \dots, x_N\}$, where x_n is the state of expression of the n^{th} gene and N is the total number of genes in the network. The state of expression of each gene changes in time according to the dynamic equation:

$$x_n(t + \tau) = F_n(x_{n_1}(t), x_{n_2}(t), \dots, x_{n_k}(t)). \quad (1)$$

In the above equation, $\{x_{n_1}(t), x_{n_2}(t), \dots, x_{n_k}(t)\}$ are the regulators of the gene x_n , and F_n is a Boolean function, also called a logical rule, which is constructed according to the combinatorial action of the regulators of x_n . The additional parameter τ is a measure of the *relaxation time*, namely, of the time that it takes for a gene to change its state of expression under a change in the

expression of its regulators. In the Boolean model, it is common to take $\tau = 1$. Each gene in the network has its own associated Boolean function. This particular GRN includes 15 genes (Figure 1) whose logical functions are grounded in experimental biological data, as explained in [3]. The updated truth tables used here are available in [12].

Note that the dynamics given by Eq. (1) is deterministic: For a given set of Boolean functions, the configuration of the network at time t completely determines the configuration of the network at the next time step $t + \tau$. Also note that since the number of dynamic states or configurations of the network is finite ($\Omega = 2^N$), under the dynamics given in Eq. (1), the network will eventually come back to a previously visited configuration, after which the network enters into a periodic pattern of expression. Such a periodic pattern is called an *attractor*, and all the initial configurations that eventually fall into that attractor constitute its *basin of attraction*. The deterministic version of the Boolean GRN modeled here recovered 10 fixed point attractors, each with a period equal to one, implying that the GRN remains in one of the 10 fixed 15-gene configurations after it reaches one of them.

Therefore, in the deterministic model defined in Eq. (1), once the system reaches an attractor, it remains there for all subsequent iterations; however, if noise is introduced into the logical rules, there is a finite probability for the system to “jump” from one basin of attraction to another. Our central aim herein was to address whether noisy perturbations of the logical rules in *A. thaliana* GRN are sufficient to recover the observed sequences of transitions among attractors (i.e., gene activity configurations characteristic of the primordial cell types within the floral meristem) during the development of this particular biological system.

The ten attractors of the 15-node GRN used here are as follows (Figure 1): Four corresponding to the four regions of the inflorescence meristem (I1, I2, I3, and I4), and six to the four floral organ primordial cells within the flower meristem (S, P1, P2, S1, S2, and C). The two attractors corresponding to petals (P1 and P2) are identical except for the state of activation of the *UFO* gene, and the same holds for the two stamen attractors (S1 and S2).

In the simulations of the stochastic versions of the GRN presented in this work, we did not consider the inflorescence attractors (I1–I4) because they are substantially separated from the floral primordia attractors. The distance between the two sets of attractors (inflorescence and floral) is clearly depicted by the way they are grouped in a phenogram (Figure 2). This is a branching diagram that groups entities according to their similarity (see Methods). The inflorescence meristem and floral organ primordia attractors cluster into two clearly distinct groups (Figure 2). Indeed, in simulations that considered all of the attractors, we found that, for a wide range of noise levels, the system never leaped out of the inflorescence attractors. On the other hand, when large noise magnitudes were considered, the system went from the inflorescence attractors to the carpel or stamen attractors, without visiting the sepal and petal attractors. Dismissing the I1–I4 attractors in the simulations allows for a better exploration of the temporal pattern in which the attractors corresponding to each of the four floral organ primordial cells are attained.

We used the GRN depicted in Figure 1 to examine which of the attractors (S, P1, P2, S1, S2, and C) the system is most likely to reach when it is initialized at a particular attractor and then is driven by noise to a different one. In order to obtain the transition probabilities among the different attractors (i.e., the entries of the so-called Markov matrix, see the detailed description below), the possible initial configurations of the system were exhaustively explored. Given any possible configuration (defined by an array of 15 entries with zeros and ones representing the activation states of

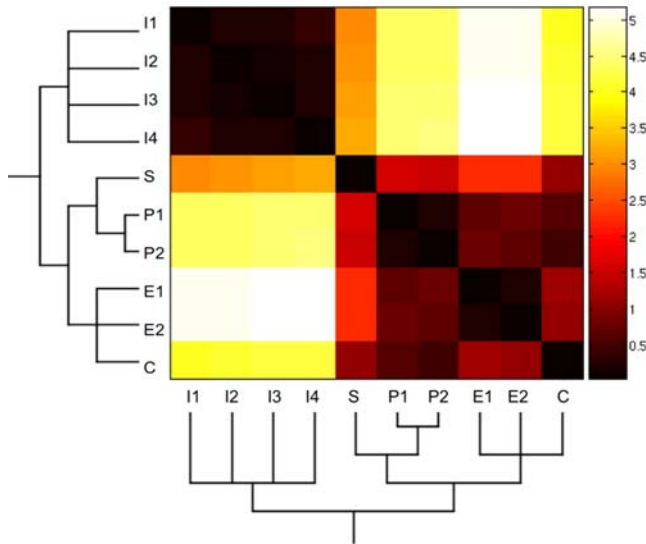


Figure 2. Heat map of the similarity matrix among the ten attractors of the GRN. A strict consensus phenogram was obtained for the GRN attractors (vectors of zeros and ones) by using the Manhattan distance similarity index (see Methods). This phenogram is shown below the attractors that are ordered along the X and Y axes of the heat map. Attractors that group together had the highest similarity indexes between them (i.e. the lowest Manhattan distance). Color scale: darker colors indicate more similar, while lighter ones indicate more different attractors in the pairs compared.
doi:10.1371/journal.pone.0003626.g002

the genes), the system was updated every iteration step according to the deterministic logical rules [12] with an error probability η . In other words, at each time step, each gene “disobeys” its Boolean function with a probability η , such that the dynamic rule in the presence of noise can be given by:

$$x_n(t + \tau) = \begin{cases} F_n(t) & \text{with prob. } 1 - \eta \\ 1 - F_n(t) & \text{with prob. } \eta \end{cases} \quad (2)$$

Note that the above equation reduces to Eq. (1) for $\eta = 0$. [In order to simplify the notation, we have written just $F_n(t)$ instead of $F_n(x_{n_1}(t), x_{n_2}(t), \dots, x_{n_k}(t))$.] These perturbations are applied independently and individually to each gene at each iteration.

If, after applying noise in one time step, the system remains in the same attractor or the same basin of attraction that it was before the noise was applied, one count is added to the main diagonal in the entry of the Markov matrix corresponding to that basin of attraction. If the configuration ended up in a different basin, a count is added to the row corresponding to the recipient basin in the Markov matrix (Table 1). This was repeated 10000 times for each of the $\Omega = 2^N$ possible initial conditions. The number of realizations was fixed to a considerably larger number than that at which the matrix entries become stable (data not shown). The transition probabilities $P(n|m)$ of the Markov matrix (Table 1) give the probability that a network in attractor m jumps to attractor n in the presence of noise, and are calculated by dividing the number of counts in each matrix entry by the total number of configurations that started in the corresponding matrix row.

Since we wanted to find the most probable sequence of transitions among the attractors representing the various cell types, we followed the changes in the probability of reaching a certain attractor throughout time given that the system was initialized in a

Table 1. Markov matrix.

	sep	pe1	pe2	st1	st2	car
sep	0.939395	0.001943	0.009571	0.000083	0.00049	0.048517
pe1	0.036925	0.904162	0.00925	0.0339	0.000488	0.015275
pe2	0.009067	0.000464	0.941609	0.000024	0.048374	0.000461
st1	0.000084	0.001893	0.00002	0.936514	0.00996	0.05153
st2	0.00002	0.000001	0.002074	0.000356	0.987953	0.009597
car	0.002045	0.000034	0.00002	0.001951	0.01002	0.98593

Matrix of transition probabilities among all possible pairs of attractors. The entries of each column in this matrix correspond to the probabilities $P(n|m)$ of reaching attractor n , given that the system is at attractor m at time $t = 0$ (see Results and Methods, noise magnitude used for this case is 1%).
doi:10.1371/journal.pone.0003626.t001

particular attractor at time $t = 0$ (see Figure 3). In order to achieve this, note that the Markov matrix (herein denoted as \mathbf{M}) in Table 1 contains the conditional probabilities $P(n|m)$ of reaching attractor n at time $t + \tau$, given that the system is at attractor m at time t . In order to obtain the temporal sequence in which attractors are most likely reached, it is necessary to repeatedly multiply the Markov matrix \mathbf{M} by the vector $\vec{v}(t)$, whose entries contain the fraction of cells at each attractor in a given population at time t . In other words, $\vec{v}(t) = (v_1(t), v_2(t), \dots, v_m(t))$, where $v_1(t)$ is the fraction of cells in the population whose configurations at time t are in the basin of attraction of the first attractor, $v_2(t)$ is the fraction of cells at time t in the basin of attraction of the second attractor, and so on. Starting out from a population with a given distribution $\vec{v}(0)$ of cells among the attractors, the distribution of cells at time t is given by: $\vec{v}(t) = \vec{v}(0)[\mathbf{M}]^t$.

Since we did not consider the four inflorescence attractors, only six attractors are involved in the dynamics. Therefore, \mathbf{M} is a 6×6 matrix and \vec{v} is a 6-dimensional vector. We also assumed that the total number of cells in the population always remains constant; hence, the sum of the six components of \vec{v} must sum to 100 (there are no “probability leaks” because transitions to the inflorescence attractors are extremely rare for the error levels used).

It is worth noting that the different attractors have basins of vastly different sizes. For instance, the basins of attraction of sepals and petals are very small in comparison to those of stamens and carpels. Therefore, the absolute probabilities for the attractors of sepals and petals are inevitably smaller than those of stamens and carpels; hence, in order to clearly observe the time at which each attractor attains its maximum probability, we divided each absolute probability value by the maximum of each attractor’s curve, and plotted the relative probabilities for each attractor probability distribution. Note that since each curve was normalized in relation to its own maxima, the probabilities in these graphs no longer add up to 1 at every moment.

It is important to notice that although the Markov matrix \mathbf{M} provides information about the probability of going from any attractor m at time t to any attractor n at time $t + \tau$, this matrix alone is not sufficient to derive the most probable sequence of transitions among attractors. The latter is only evident when the matrix \mathbf{M} is recursively multiplied by the vector \vec{v} containing the fraction of cells per attractor, ideally until the system reaches a steady probability distribution.

Since sepal cells are the first to attain their fate in flower development, we used an initial vector $\vec{v}(0)$ with $v_1(0) = 100$ and $v_m(0) = 0$ in all of the other entries (the first entry corresponds to the sepal configuration). Thus, initially, all of the population of cells

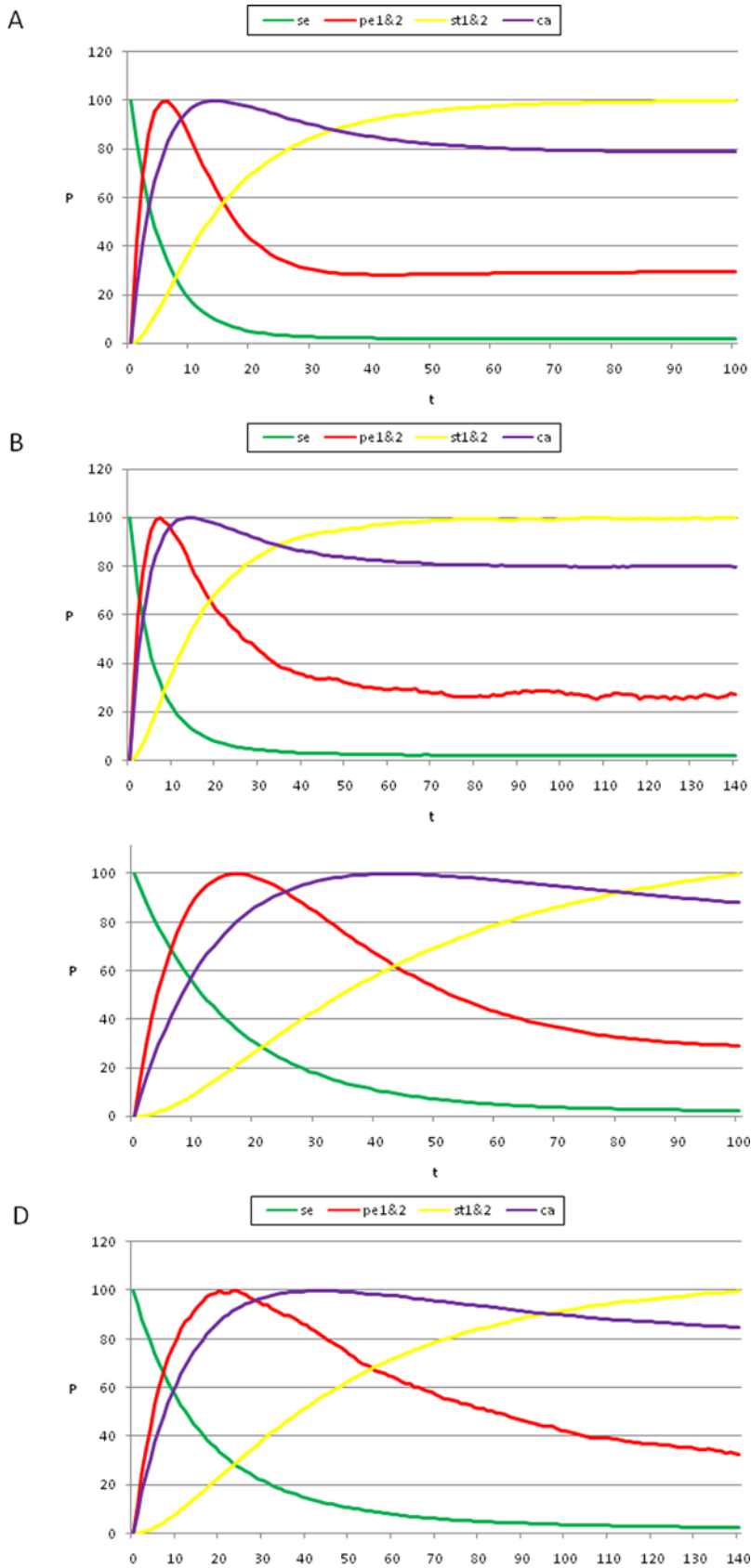


Figure 3. Temporal sequence of cell-fate attainment patterns under the Boolean dynamics with noise. Maximum relative probability ("Y" axis) of attaining each attractor, as a function of iteration number or time ("X" axis). (A) Probability of attaining each attractor (i.e., cell type) obtained by multiplying the Markov matrix \mathbf{M} by a population vector \vec{v} initialized at the sepal attractor. The error probability in computing this graph

was $\eta = 0.03$. The most probable sequence of cell attainment is: Sepals, petals, carpels, and stamens. (B) Probability of attaining each attractor (i.e., cell type) at each iteration when 80000 randomly chosen "sepal" configurations were selected and followed for 140 steps. Noise was introduced in the updating of each gene independently, with a $\eta = 0.03$ probability at each iteration. The probabilities for the petal (p) and stamen (st) attractors correspond to the sum of $p1+p2$ and $st1+st2$, respectively. All maxima correspond to 100 because each absolute probability value was divided by the maximum of each attractor's curve (see Results and Methods). Equivalent graphs to those in (A) and (B) for $\eta = 0.01$ are shown in (C) and (D), respectively.

doi:10.1371/journal.pone.0003626.g003

within a floral primordia is in the sepal attractor. We then followed the changes in the probability of reaching each one of the other attractors over time, given that the entire system started in the sepal configuration (see Figure 3A). Every attractor has a maximum or peak in the probability of being reached at particular times. This maximum corresponds to the moment at which the corresponding primordial cell fate is most likely.

The use of the probability peaks to determine the time at which each cell multigenic configuration is most probable follows the standard reasoning in deriving maximum likelihood estimators in statistics [25]. The time at which the probability peak appears corresponds to the maximum of the associated transition probability for that particular attractor. The order of appearance of the peaks shown in Figure 3 matches the order of formation of the maxima of the transition probabilities. Recall that when using the maximum likelihood methodology [25], the main assumption is that the set of real data is precisely observed because they are more likely to happen than other possible data sets. In other words, they maximize the probability of being observed among all possible samples of the same size. Conversely, if we want to know when a specific event is more likely to happen, the most natural assumption is that it will be at a maximum of the corresponding probability distribution. This is precisely what we claim based on the graphs of the frequencies of visits to each attractor. Also notice that the locations of the maxima are not affected by normalization.

This interpretation hence implies that, given that a particular attractor will be reached (i.e. that a specific event will occur), it is natural to assume that the most likely time for it to occur is when the probability of reaching that particular attractor is maximal. Therefore, we propose that the temporal sequence in which attractors are attained will correspond to the sequence in which their maximum probabilities are reached.

A related important issue has to do with the interpretation of the transition probabilities. There are at least two possibilities that are consistent with the traditional approaches in statistical studies of collective behavior [26]. First, it is possible to consider that each agent (in this case, a single cell) will spend some time at each equilibrium configuration and then will jump to another with a certain probability. This would imply that each cell transits through different configurations. In our case, for example, a particular cell might attain a sepal primordia identity, then transit to a petal primordial cell, then to a stamen primordial cell, and finally to a carpel primordial cell. An alternative interpretation is that, from a given initial population of cells, the number of individual cells at a certain attractor at any given time, is proportional to the transition probability of reaching that particular attractor.

These two interpretations are equivalent or are assumed to be so (ergodic hypothesis) in many applications of statistical physics. This is often summarized by saying that averaging quantities in time is the same as averaging them in space [26]; however, in the case we have considered here, the second interpretation seems more appropriate. Future experimental studies that actually follow gene configurations over time at the individual cellular level will directly test these two alternative interpretations. For now, if we accept the overall population of undifferentiated cells in the floral meristem as our system, it is consistent to assume that the

proportion of them reaching a particular configuration will be in accordance with the transition probabilities.

Therefore, we present a stochastic GRN that can be interpreted as a model of cell population dynamics. This model describes the dynamics of cells within the flower meristem, in which different fractions of cells sequentially attain distinct configurations. Therefore, it does not imply that individual cells transit through different identities or configurations, but rather that once in a floral meristem, one set of cells attains a certain identity first (sepal primordia) and then, from the remaining cells, another fraction attains a second cell fate (petal primordia), and so on, until all the cells in the floral primordium have reached an identity corresponding to each of the four floral organ primordia. Later in development, primordia will grow and differentiate to form the four floral mature organs: Sepals, petals, stamens, and carpels. The latter events are regulated by other GRNs. We explored whether the observed dynamics of cell-fate attainment can be recovered by the stochastic Boolean GRN model presented here.

Simulated temporal transitions among attractors (cell types) mimicked the sequence in which A, B, and C genes are expressed in real flower meristems

By following the procedure presented above, we found that, by starting from the gene configuration associated with sepal primordial cells ($t=0$ in Figures 3 A and C), the next maximum probability was observed in the petal curves, P1 plus P2 ($t=18$ in Figure 3A). Afterwards, the peaks for the probability of attaining first the carpel and then the stamen (S1 plus S2) identity appeared ($t \approx 45$, $t \approx 100$ in Figures 3A and C). Interestingly, the same sequence was observed when applying a range of noise magnitudes from 0.5 to 10%; however, the peaks corresponding to the stamen and carpel cell fates became closer, almost simultaneous, as the noise magnitudes increased (compare Figures 3A–C). Nonetheless, it is noteworthy that the probability peak of the carpel configuration appeared before the peak of the stamen configuration.

The sequence resulting from the aforementioned model mimics the observed temporal pattern for A, B, and C gene expression: A-genes are expressed first, followed by B-genes, and finally by the C-gene [27,28]. Furthermore, our model predicts that the gene configuration characteristic of carpels most probably appears before that corresponding to stamens during early flower development. This would, in fact, be the case if the C gene was first expressed in the flower center and then its expression expanded to the peripheral whorls. This should be tested experimentally by gathering data on the population dynamics of cells with different genetic configurations during early stages of flower development.

It is noteworthy that, among all of the tested noise levels, the only non-trivial temporal sequence of A, B, and C gene combinations recovered was: A, then AB, then C and finally BC. Although the latter two appeared almost simultaneously as error magnitudes used increased. This sequence is congruent with the ABC temporal pattern in *Arabidopsis thaliana* (Figures 3A and C) in which the A genes are turned on first, then the B and finally the C genes; hence BC and C combinations are defined at the same time. The trivial behaviors are: i) remaining in the initial configuration forever, and ii) transitions depending only on the size of the basins of attraction

(i.e., the system behaves according to only noise). If the magnitude of the noise is increased, for example to 50%, the system goes from sepal to stamen1 or carpel configurations directly. This is because the basins of attraction corresponding to petals are very small in comparison to those of stamens and carpels.

In addition to the Markov matrix approach, we also performed simulations by directly following trajectories starting in randomly chosen configurations from the basin of attraction corresponding to the “sepal” configuration. We followed each of 80000 such configurations for 140 iterations in order to compute the probabilities of directly attaining each attractor at each iteration (see Methods). This latter simulation is directly comparable to that performed for the Glass system discussed in the following section. It is noteworthy that the sequence of probability peaks we found for each attractor over time is the same as the one that we had obtained using the Markov Matrix approach: Sepal, petal, carpel, and stamen (Figures 3B and D).

Continuous GRN model with noise

In order to develop a continuous model based on the differential equations of the flower development GRN considered here, one would need to know all of the kinetic reaction constants, promoter affinities, degradation rates, and many other parameters involved in the dynamics. To the best of our knowledge, these have not yet been identified; however, a first step towards a continuous description of this GRN is to implement the Glass dynamics in the network [29]. This can be accomplished by considering the parameter τ in Eq. (1) as a small quantity, and expanding the left-hand side of that equation to the first order in powers of τ , which gives:

$$\frac{dx_n(t)}{dt} = \alpha[F_n(x_{n_1}(t), x_{n_2}(t), \dots, x_{n_k}(t)) - x_n(t)], \quad (3)$$

where $\alpha = 1/\tau$ is a measure of the “relaxation” time in the gene expression profile. Although the above equation is formally correct, it has the problem that the Boolean function F_n on the right-hand side has to be evaluated using discrete variables, whereas the derivative on the left-hand side treats the x_n 's as continuous variables. Therefore, each continuous variable x_n has to be transformed into a discrete variable in order to evaluate the Boolean function. This is accomplished by introducing the discrete variables \hat{x}_n defined as:

$$\hat{x}_n = H(x_n - \theta_n), \quad (4)$$

where θ_n is a threshold, and $H(x)$ is the Heaviside function. ($H(x) = 1$ if $x \geq 0$ and $H(x) = 0$ if $x < 0$). Thus, each continuous variable x_n , representing the *level of expression* of a given gene, has an associated discrete variable \hat{x}_n that represents the *state of expression* of that gene: “ON” if x_n is above the threshold θ_n , and “OFF” if x_n is below θ_n . In principle, each gene x_n could have its own threshold θ_n . Our simulations show that the results are qualitatively the same if we randomly assign the thresholds in the interval $\theta_n \in [0.35, 0.65]$. Thus, in what follows, we fixed $\theta_n = 0.5$ for all of the genes.

The continuous piece-wise linear Glass dynamics of the network can thus be given by:

$$\frac{dx_n(t)}{dt} = \alpha[F_n(\hat{x}_{n_1}(t), \hat{x}_{n_2}(t), \dots, \hat{x}_{n_k}(t)) - x_n(t)] \quad (5)$$

We will refer to the set of continuous values $\{x_1(t), x_2(t), \dots, x_N(t)\}$ as the *microscopic configuration* of the network, and to the set of

corresponding discrete values $\{\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_N(t)\}$ as the *Boolean configuration* of the network. Note that there are infinitely many microscopic configurations compatible with the same Boolean configuration. Finally, we will refer to the dynamics generated by Eq. (5) as *Glass dynamics*.

It has been pointed out that the discrete model given in Eq. (1) and the corresponding continuous piece-wise linear model defined in Eq. (5) are not necessarily equivalent, since the attractors of the two models can be different, even when the Boolean functions F_n are the same in both cases. Nonetheless, our numerical simulations show that for the *A. thaliana* network, the Glass dynamics generate exactly the same ten point attractors obtained in the Boolean model, and only those ten attractors. Therefore, from now on, we will make no distinction between the attractors of the Boolean model and the attractors of the continuous model, referring to them simply as the attractors of the floral GRN.

Even when the Boolean dynamics and the Glass dynamics produce the same ten attractors, their basins of attraction do change from one model to the other. This is so because two different initial microscopic configurations that correspond to the same Boolean configuration may end up in two different attractors under the Glass dynamics. In order to show that this is indeed the case, for each of the $\Omega = 2^N$ Boolean configurations of the network, we probed 10,000 compatible microscopic configurations. We evolved these 10,000 microscopic configurations in time until an attractor was reached, and determined the configuration in which the network fell. Figure 4 depicts in a color map the probability $P_C(n|m)$ that the network ends up in attractor n under the Glass dynamics, given that it started in a microscopic configuration whose corresponding Boolean configuration was in the basin of attraction of attractor m . As can be seen, the highest probabilities lie along the diagonal; however, the non-vanishing off-diagonal elements indicate that two different microscopic configurations corresponding to the same Boolean configuration may end up in two different attractors.

On the other hand, Table 2 shows the fractional sizes of the basins of attraction in both the Boolean and the continuous

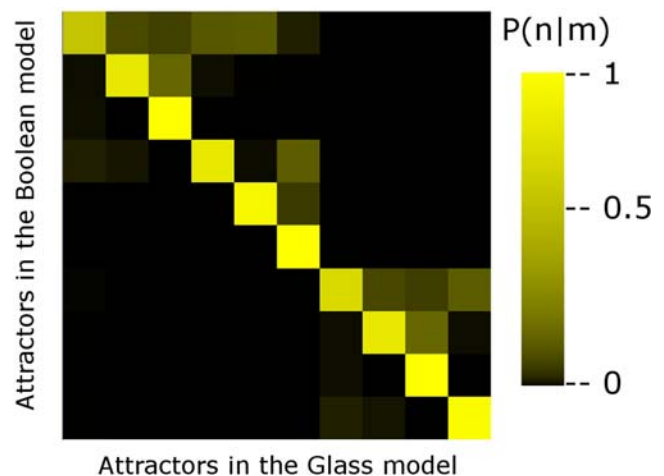


Figure 4. Changes in the basins of attraction of the continuous model with respect to the Boolean model. Color map of the probability $P(n|m)$ that a microscopic configuration whose associated Boolean configuration belongs to the basin of attraction of attractor m , ends up in attractor n using Glass dynamics. Note that the main transitions occur along the diagonal where attractors are reached by both dynamics (Boolean and Glass); however, the non-diagonal elements indicate that two microscopic configurations that correspond to the same Boolean configuration may end up in different attractors. doi:10.1371/journal.pone.0003626.g004

Table 2. Basins of attraction.

Attractor	Boolean Dynamics	Glass Dynamics
Inflorescence 1	0.0156	0.0500
Inflorescence 2	0.0156	0.0500
Inflorescence 3	0.0078	0.0380
Inflorescence 4	0.0078	0.0381
Carpel	0.4404	0.2622
Sepal	0.0185	0.0670
Stamen 1	0.4570	0.3331
Stamen 2	0.0166	0.0710
Petal 1	0.0195	0.0786
Petal 2	0.000976	0.0116

This table shows the fractional sizes of the basins of attraction in the Boolean and Glass models. The data for the Glass dynamics were obtained by sampling 10,000 microscopic configurations for each of the $\Omega = 2^N$ Boolean configurations, and by counting the frequency with which these microscopic configurations end up in each of the ten attractors. doi:10.1371/journal.pone.0003626.t002

models. It is apparent from this table that, when passing from the Boolean to the continuous description, the largest basins of attraction (carpel and stamen1) lose about 30 to 40 percent of their configurations, which are redistributed among the smaller basins of attraction. Thus, even when the predicted cell types (attractors) are the same in the two models, the basins of attraction are not.

The stochastic continuous model of the GRN yields a cell-fate attainment sequence similar to the Boolean stochastic model

In order to implement noise in the continuous model, we followed a procedure similar to the one indicated in Eq. (2); namely, with a probability η , each gene will disobey its Boolean function F_n , replacing it by $1 - F_n$; however, since the system in this case is governed by differential equations, this “perturbation” will occur during a *finite time interval* Δt_p , rather than being instantaneous. In other words, if at time t one particular gene x_n is perturbed and chosen

to disobey its Boolean function, then from time t to time $t + \Delta t_p$ its state will not be determined by Eq. (5), but rather by the equation:

$$\frac{dx_n(t)}{dt} = \alpha[1 - F_n(\hat{x}_{n_1}(t), \hat{x}_{n_2}(t), \dots, \hat{x}_{n_k}(t)) - x_n(t)] \quad (6)$$

After the time interval Δt_p , the state of x_n will be determined again by Eq. (5), and a new set of “disobeying genes” will be chosen. We will call these disobeying genes the *perturbed genes*.

We have to choose the value of Δt_p in such a way that the gene has enough time to relax to its new state after the perturbation has been produced. In other words, Δt_p has to be larger (or at least of the same order of magnitude) than the relaxation time $\tau = \alpha^{-1}$ appearing in Eq. (5). Figure 5 shows two typical noisy realizations of the temporal evolution of a particular $x_n(t)$ as a function of time, for two different choices of τ and Δt_p : One for $\Delta t_p = 2.5$ and $\tau = 1$ (black curve), and the other for $\Delta t_p = 2.5$ and $\tau = 1/20$ (red curve). The two realizations started out from the same initial conditions, and underwent the same set of perturbations. The only difference was the value of τ . As can be seen from this figure, the trajectories are qualitatively the same as long as $\Delta t_p > \tau$. In what follows, we selected $\Delta t_p = 2.5$ or 1 (Figures 6A and B, respectively), and $\tau = 1$ to simulate Glass dynamics with noise (see methods for further details).

In order to determine the cell-fate attainment patterns in the *A. thaliana* network under Glass dynamics with noise, we analyzed the transitions between attractors over time in a population of 80 000 cells subject to the perturbations described above. At time $t = 0$, all of the cells were initialized in different random microscopic configurations corresponding to the sepal basin of attraction. In every cell, each gene was independently chosen to be perturbed with a probability $\eta = 0.03$. The non-perturbed genes then evolved in time according to Eq. (5), whereas the perturbed genes evolved following Eq. (6). After a time interval $\Delta t_p = 2.5$ or 1 (for Figures 6A and B, respectively), a new set of perturbed genes in the entire population was chosen again, and so on.

At each unit of time, we looked at the microscopic configuration of each cell and determined to which attractor this microscopic configuration would have evolved in the absence of perturbations. This allowed us to associate a given attractor at each unit of time to each cell configuration. The results of this simulation are

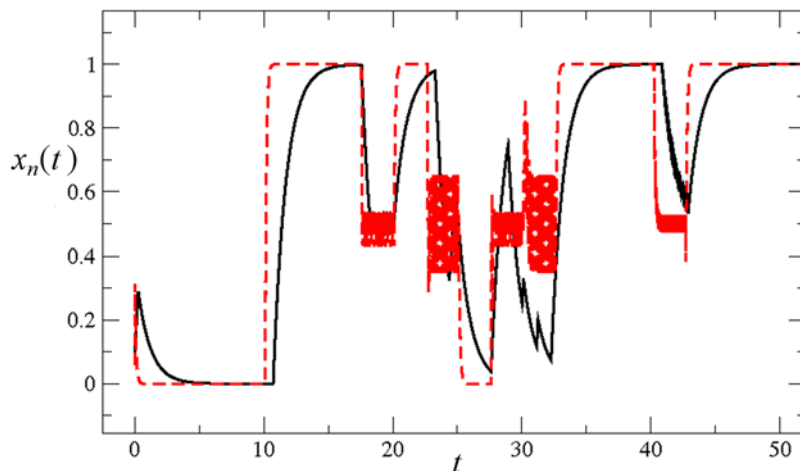


Figure 5. Effects of the choice of the relaxation time on Glass dynamics with noise. Two typical realizations of Glass dynamics for a given gene x_n showing that the choices of the relaxation time τ and the perturbation time Δt_p do not affect the qualitative dynamics, so long as $\Delta t_p > \tau$. Both trajectories started from the same initial conditions, and were followed through the same set of perturbations. The black trajectory corresponds to $\Delta t_p = 2.5$ and $\tau = 1$, whereas the red trajectory corresponds to $\Delta t_p = 2.5$ and $\tau = 1/20$. doi:10.1371/journal.pone.0003626.g005

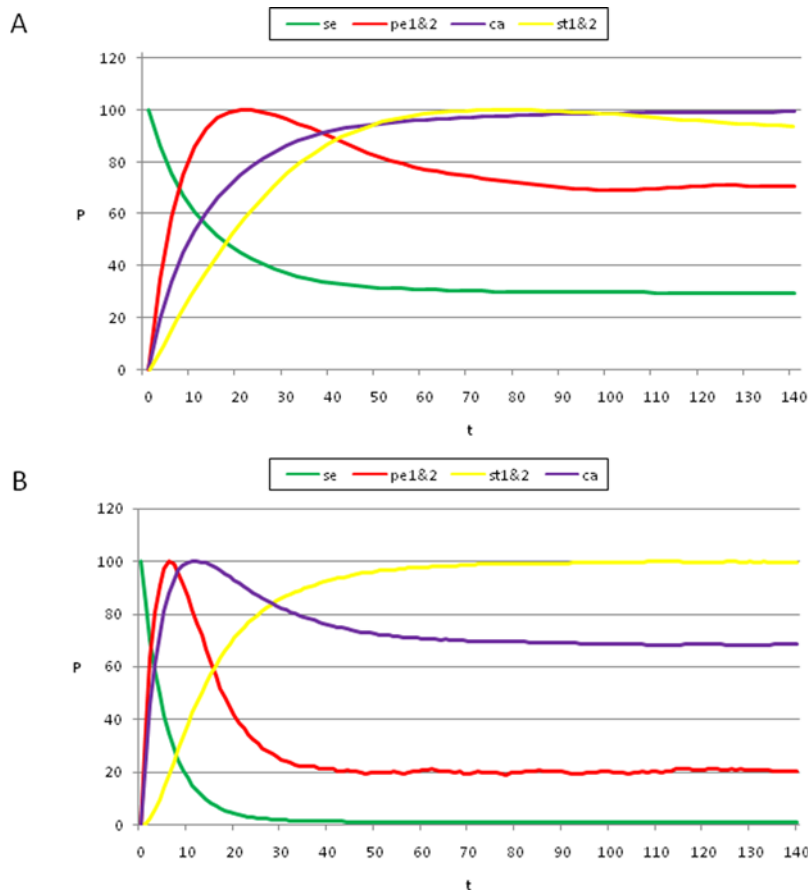


Figure 6. Temporal sequence of cell-fate attainment patterns under the Glass dynamics with noise. Maximum relative probability (“Y” axis) of attaining each attractor as a function of iteration number or time (“X” axis). (A) The maxima of the cell-fate curves are attained in a particular sequence in time, which in this case is sepal, petal, stamen, and carpel. Parameters used: $dt = 0.01$, $\tau = 1$, and $\Delta t_p = 2.5$. (B) When the simulations mimic the Boolean case ($dt = 1$, $\tau = 1$ and $\Delta t_p = 1$; see Results and Methods), a temporal pattern identical to that of the Boolean dynamics was obtained, with a sequence of sepal, petal, carpel and stamen. The noise used in both cases was $\eta = 0.03$. Although the Boolean and Glass dynamics need not coincide in general, for the case of the *A. thaliana* GRN, both models provide similar predictions. Simulations show that the order of emergence of the stamen and carpel maxima, as compared to the Boolean model, may depend on the precise values of the kinetic constants.
doi:10.1371/journal.pone.0003626.g006

reported in Figure 6, which shows the evolution in time of the population of cells, and shows how the cells redistribute among the sepal, petal, stamen, and carpel attractors. Similar patterns were recovered for other values of noise (data not shown), ranging between $\eta = 0.005$ and $\eta = 0.1$, as well as for other values of Δt_p .

As can be observed in Figure 6, the results obtained using Glass dynamics are analogous to those obtained for the Boolean model, in that the addition of noise to the dynamics produces the emergence of cell-fate attainment patterns in a population of cells in a specific temporal order. Thus, the use of the Glass model, based on piece-wise linear differential equations, reveals that the time ordering in the emergence of the cell-fate patterns is not an artifact of the synchronous updating in the Boolean model; however, the stamen and carpel peaks are reversed in time between the Boolean and Glass models (Figure 3 vs. Figure 6A).

In real flowers, A genes are first “ON”, followed by the B genes that turn “ON,” thus defining the A (sepal) to AB (petal) transition. This is recovered by both models (Figures 3 and 6), and is observed in real flowers. The C genes then turn “ON,” and hence, the BC (stamens) and C (carpels) configurations are defined at the same time. While the Boolean dynamics predict that the carpel primordia cell fate (C alone) will be attained before that of the

stamen (BC), in the Glass model, these two are reversed (Figure 3 vs. Figure 6A). Interestingly, when this model is simulated to mimic the Boolean model (Figure 6B), both systems recover the same sequence: “Sepal-petal-carpel-stamen” (Figure 3 vs. Figure 6B) and in both cases the time at which stamen and carpel configurations are determined converge as noise levels are increased. Detailed experimental data on the precise spatio-temporal dynamics of the gene activation profiles of cells in the developing flower meristem are needed to test which of the two peaks is observed first in real floral buds. Such data will also be useful to determine which of the two models predicts the most realistic frequency distributions of cell types over time. The latter will be related to the relative sizes of the basins of attraction.

Glass system simulations indicate that the order of appearance of the two peaks (stamen or carpel) may depend on the precise values of the reaction-kinetic constants and degradation times, as well as some other epigenetic processes not taken into consideration in the simple analysis presented here. The important conclusion of both models is that noise in the gene-expression dynamics is necessary and sufficient to qualitatively recover the temporal transitions among the ABC-gene configurations observed during early flower development.

Discussion

Robust morphogenetic patterns that are recreated over the life cycles of individuals from the same species, or even from distantly related species, have led to the prevailing view of development as a deterministic process; however, we have shown here that the stereotypical temporal pattern with which floral organs are determined may result from a stochastic dynamic system associated with a highly non-linear GRN.

This study supports recent work that has concluded that random fluctuations in a system may be important for cell behavior and pattern formation ([14–21]), and contrasts with deterministic and preprogrammed views of development. Intrinsic noise (noise arising from the system itself) has its origin in molecular fluctuations due, for example, to slight modifications in temperature, and in random events due to sampling, given that the number of molecules is not infinite during transcription and translation [16,19,14].

Stochastic implementations of a GRN model as pursued in this study were proposed by C. H. Waddington many years ago ([22]; see review in [23]). He understood development as a complex dynamic system, with genes, proteins, metabolites, and environmental factors constituting complex dynamic networks. The attractors of such networks represent a specific configuration of the system (e.g. cell types). The number, depth, width, and relative position of these attractors are represented by the hills and valleys of his “Epigenetic Landscape” metaphor [22,7]. The study presented here actually explored such an Epigenetic Landscape for the flower organ determination GRN (Figures 1 and 7). Other recent studies have also explored this idea for GRNs [30].

In the case presented here, a GRN generates the overall temporal morphogenetic pattern (Figures 3 and 6) observed during flower development of *Arabidopsis thaliana* [31,32]: A genes are expressed first, followed by B genes, and finally C genes, in a rather broad range of noise magnitudes, and in two different modeling approaches. Therefore, our results provide a possible explanation for the conservation, among many flowering plant species [27,28,31–34], of the temporal transitions of A, B, and C-gene expression, and to some extent, of the observed cell fate attainment patterns.

Our results support the hypothesis that biological systems may not only cope with random perturbations, but that the noise may have been incorporated during evolution in the generation of biological patterns (e.g. [30,35–37]). Central to the constructive role of noise is the existence of non-linear dynamic systems [38] that converge to robust attractors for a range of noise magnitudes. Stochastic implementations of GRNs, such as the one presented here, may guide predictions of actual noise magnitudes experienced in biological systems.

Nevertheless, deterministic signals or inducers of flower development cannot be dismissed. Indeed, our results hold when focusing on the attractors corresponding to the four types of floral organ primordia. However, if all of the attractors (including I1–I4) are considered, and the system is initialized in one of the inflorescence basins, the system hardly ever transits into the floral basins when small noise levels are used, or else it directly jumps to one of the largest basins (stamens1 or carpels) when larger magnitudes of noise are simulated. These results enable us to speculate on the role of reported non-random inducing signals in the transition from cell fates in the inflorescence meristem to those in the flower meristem. Genes such as *FLOWERING LOCUS T*,

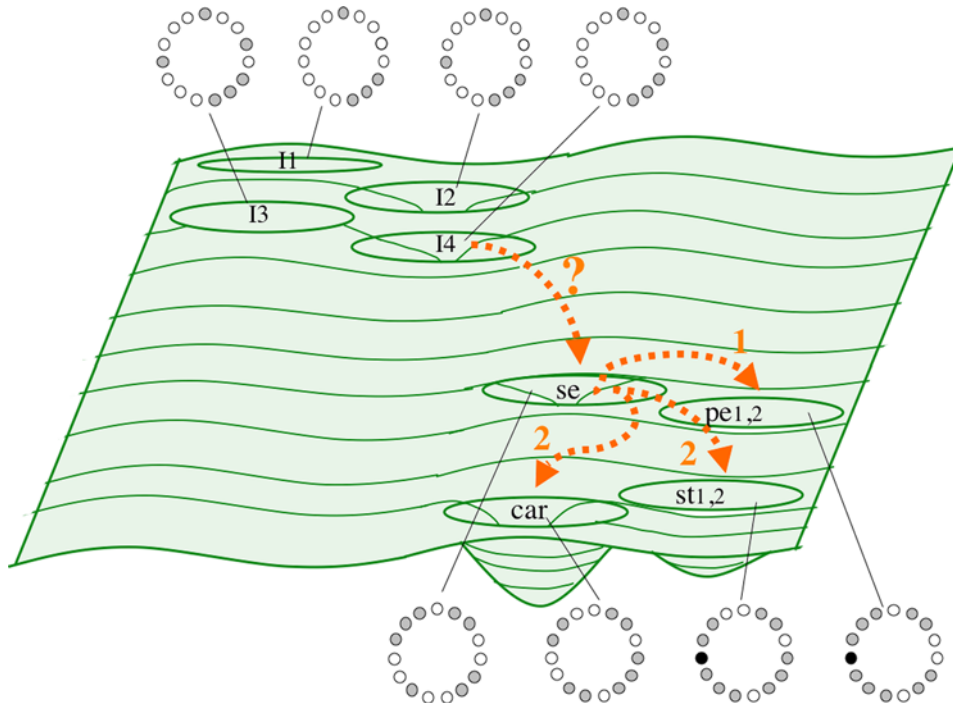


Figure 7. Schematic representation of the epigenetic landscape generated by a stochastic exploration of the GRN for flower development. This schematic landscape is equivalent to the Epigenetic Landscape proposed by C.H. Waddington (1957). Basins comprise the cell genetic configurations that lead to attractors (in this case, gene arrays characteristic of floral organ primordial cell-types: Sepals, petals, stamens, and carpels. See Figure 1 and Discussion). Each cell fate is associated to the GRN configuration corresponding to each of the attractors. The arrows represent transitions among attractors. The transition from inflorescence to sepal attractor might be biased or determined by an inducer. The numbers associated to the arrows represent the sequence of transitions among attractors: From sepals to petals, and then to carpels and stamens. doi:10.1371/journal.pone.0003626.g007

SUPPRESSOR OF OVEREXPRESSION OF CO 1, or *CONSTANS* (see [39] for a review) could constitute or mediate such signals.

The type of model put forward here will enable the predictions of the real magnitudes of stochastic fluctuations once such deterministic biasing signals are considered. They will also be useful to test what mutations may cause alterations in the epigenetic landscape and alter the temporal order with which attractors are visited. Such models will guide the search of genetic alterations underlying atypical morphogenetic patterns during the evolution of flowering plant species [40].

One possible interpretation of our model is to assume that, once most cells have attained a certain attractor within a primordium, these are canalized to develop into a particular organ type. One possible explanation for this is that noise does not drive the cells out of each configuration once a certain proportion of them attain an attractor, or that the noise is “frozen” at some point, maybe because irreversible differentiation or synchronization events take place. We may speculate that, in the developmental system we have studied, non-autonomous cell function of key transcription factors [41–44] could play a relevant role in this process, as it could effectively freeze the stochastic fluctuations or synchronize the configuration of the cells within a primordium, and thus, contribute to the formation of the observed spatio-temporal patterns. We could further speculate that the activity of pre-patterning genes (e.g., *WUSCHEL* or *UNUSUAL FLORAL ORGANS*; [43,45,46]) may play important roles during spatio-temporal pattern formation.

Models such as those presented here enable novel predictions about the genetic regulation of cell differentiation and morphogenetic patterns. For example, the stochastic GRN dynamic system eventually attains a stationary distribution of attractor probabilities. The distribution reflects the probability of the cells being in each attractor, and may be interpreted as the proportion of primordial cells fixed to each GRN configuration. In the floral organ specification network, such proportions would correspond to the regions within the floral meristem with A, A+B, B+C, and C function configurations; however, this distribution may only be observed at the very early stages of the partitioning of the floral bud into four concentric rings. This event occurs before cells committed to a certain cell-type start further differentiation and acquire distinct division and elongation rates; hence, the final amount of cells in a certain organ or organ primordium would not necessarily coincide with that predicted by the models presented herein.

Another prediction derived from this model states that the carpel attractor appears either before (Figure 3) or after (Figure 6A) that of stamens. This prediction does not contradict the fact that, in most plants, carpels are the last organs to be fully formed because, again, cells have different division and elongation rates after cell-type differentiation, and therefore, the order in which organogenesis takes place may not match the sequence in which organ primordia cells are determined during early flower development, before the primordia actually emerge.

The discussion above suggests that models that incorporate GRN associated to cellular growth and proliferation, as well as spatial aspects of the system presented here, will eventually be needed to understand the dynamics by which cells attain their fate and proliferate in the floral spatio-temporal domain. In this paper, we have restricted ourselves to exploring the temporal patterns of cell-fate establishment early in flower development, assuming that cells differentiate independently of one another; however, in real organisms, cell-cell communication, cellular dynamics, domain geometry, and growth or mechanical interactions, are all likely to alter the proportion of cells across space and time that are set aside for each type in early flower development [10].

Kauffman’s Boolean model for cell differentiation has been criticized because it is said to oversimplify the gene regulatory interactions and the way activation states of all genes are updated (synchronously in Kauffman’s proposal); however, Boolean GRN models grounded in experimental data have been able to recover observed multi-gene expression arrays characteristic of certain cell types in several biological systems [2,3,7,36]. These results suggest that the logic of regulation considered in Boolean networks suffices to qualitatively reproduce the dynamics of biological GRNs. Furthermore, theoretical studies have suggested that the details of the kinetic functions are not relevant in determining the system’s attractors. In particular, updating schemes do not seem to affect the number and identity of fixed-point attractors [47], as is the case of the attractors recovered in the network used here.

Given that the identity of the attractors and the temporal sequence in which these were attained are the same (Figures 3A–C vs Figure 6B) or very similar (Figures 3A–C vs Figure 6A) using Boolean and Glass dynamics, this study reveals that the time ordering in the emergence of cell-fate patterns is not an artifact of synchronous updating in the Boolean model; however, the sizes of the basins of attraction differ between the two models. In Glass dynamics, the basins corresponding to stamen and carpel primordia cells are smaller, and those of sepals and petals are larger (Table 2); hence, the proportion of cells at each fate along time predicted by the Glass and Boolean dynamics differ, which suggests that the updating schemes might be relevant to recovering the actual temporal cell population dynamics in biological systems. Experimental data on the temporal fluctuations of primordial cells with different multi-gene expression arrays will test which of the two systems and updating hypotheses better reproduces the real system.

Eventual formalizations of stochastic multicellular GRN dynamics in explicit spatial domains may require “hybrid” approximations that enable large computational explorations, and allow, for instance, the explicit incorporation of developmental processes into models of network or phenotypic evolution [48], or the study of the epigenetic landscapes that emerge from GRN related to complex diseases, such as cancer [9].

In conclusion, we put forward a stochastic approach to model the Boolean and continuous dynamics of an experimentally-based GRN, and thus, take Waddington’s Epigenetic Landscapes into a specific biological framework: Flower organ specification in *Arabidopsis thaliana*. The theoretical framework of this proposal could also be useful for studying the behavior of other networks, including, for instance, ecological, epidemiological, immunological, engineering, or social networks. Finally, our results emphasize that complex networks and stochastic processes are central to understanding the biological development and emergence, as well as the stability, of morphogenetic patterns.

Methods

Construction of phenogram of attractors

We obtained six phenograms by estimating the Manhattan distance index to infer the relationships among the 10 attractors for the 15-gene system. This index was obtained by comparing the vectors of zeros and ones of each attractor. We then used the clustering method by the unweighted pair-group method with arithmetic average (UPGMA) to group the attractors. We obtained six different phenograms, with which we constructed a strict consensus that kept the branches that were recovered in all of the six phenograms. In Figure 2, the consensus phenogram is shown below the attractors ordered along the X and Y-axes of the heat map, corresponding to the Similarity Matrix.

Implementation of noise in the GRN model

Boolean case. The GRN has 15 elements; two of them (LUG and CLF) are constitutively expressed in the flower meristem, and thus, their activation states were fixed to 1. The transition probabilities among attractors in the Boolean GRN implementation were obtained by introducing noise to the updating logical rules in 10,000 realizations for each possible configuration of the system. The analyses of the Boolean model were performed with the “Atalia” software, which is publically available (<http://www.ecologia.unam.mx/achaos/Atalia/atalia.htm>).

Another equivalent method to obtain the Markov matrix entries would be to follow the system’s trajectory for every possible initial configuration. For certain levels of noise, the system never remains at a particular basin, and it is hard to determine when to stop the computation for the corresponding initial condition. Nonetheless, we performed a similar type of simulation in order to mimic that of the Glass system. We selected a random configuration from those in the “sepal” basin. Each gene was updated according to its true table, except that with a certain probability (0.01 and 0.03), the rule was violated, and if the true table predicted that a state should be “1,” it was set to “0,” and vice versa. The new basin was registered, and this procedure was continued for 140 iterations. 80,000 such realizations were obtained (i.e., 80,000 randomly chosen configurations from the “sepal” basin were chosen).

Glass system. The model is explained in the Results section. We numerically integrated the set of differential equations (5) and (6) using the Euler method with an integration step $dt=0.01$. The results do not change by choosing smaller values of dt ; however, if we take $dt=\tau=\Delta t_p=1$, then the continuous model given in equations (5) and (6) becomes completely equivalent to the Boolean model given in Eq. (2). The results for this latter case are shown in Figure 6B. In order to recover the temporal sequence, in which attractors (cell-fate) were attained in the *A. thaliana* network using Glass dynamics with noise, we followed transitions for 140

time-steps, starting with a population of 80,000 cells (configurations from the “sepal” basin of attraction), in which each gene was independently chosen not to be updated according to its logical functions (set to “1” if the predicted value was “0,” and vice versa), with a probability $\eta=0.03$; hence, the non-perturbed genes evolved in time according to Eq. (5), while the perturbed genes evolved following Eq. (6). After a time interval $\Delta t_p=2.5$ for Figures 6A, and 1 for Figure 6B, a new set of perturbed genes in the entire population was chosen again, and so on until 140 iterations were completed. Qualitatively similar results were obtained for a noise of 0.01. The code for the Glass system simulations was developed in JAVA, and is available upon request.

Acknowledgments

We thank J. Diaz, G. Fonseca and S. Kauffman for useful discussions, and J. Bowman for permission to use his Scanning Electron Micrography in Figure 1. We thank R. V. Pérez-Ruiz and D. Romo for technical and logistical assistance, respectively. Suggestions from two anonymous referees, as well as those from the Scientific Editor, greatly improved previous versions of this study. We remain responsible for its final contents.

Author Contributions

Conceived and designed the experiments: ERAB AC MA PPL. Analyzed the data: ERAB AC MA MB YCP CES DHF RBL DM GJES PPL. Wrote the paper: ERAB MA MB RBL PPL. Overall project design, coordination and integration: ERAB. Implemented and planned simulations of Boolean Dynamics: ERAB. Ran simulations of Boolean Dynamics: ERAB AC. Designed simulations for Boolean Dynamics: AC. Conceptual contributions: AC MB CES DM GJES PPL. Figures: AC MB. Ran and designed simulations for the Glass Dynamics: MA. Singular value decomposition calculations: YCP. Programming and ran simulations for the Glass Dynamics: DHF. Programming for Boolean Dynamics: DM. Simulations: GJES. Revised paper: GJES PPL.

References

- Von Dassow G, Meir E, Munro EM, Odell GM (2000) The segment polarity network is a robust developmental module. *Nature* 406: 188–193.
- Albert R, Othmer HG (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.
- Espinosa-Soto C, Padilla-Longoria P, Alvarez-Buylla ER (2004) A gene regulatory network model for cell-fate determination during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles. *Plant Cell* 16: 2923–39.
- Kauffman SA (1993) *The Origins of Order: Self-organization and Selection in Evolution*. Oxford: Oxford University Press.
- Davidson E, Rast JP, Oliveri P, Ransick A, Caestani C, et al. (2002) A genomic regulatory network for development. *Science* 295: 1669–1678.
- Acar M, Becskei A, van Oudenaarden A (2005) Enhancement of cellular memory by reducing stochastic transitions. *Nature* 435: 228–32.
- Huang S, Eichler G, Bar-Yam Y, Ingber DE (2005) Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys Rev Lett* 94: 128701.
- Lipniacki T, Paszek P, Brasier AR, Luxon BA, Kimmel M (2006) Stochastic regulation in early immune response. *Biophys J* 90: 725–42.
- Huang S, Ingber DE (2007) A non-genetic basis for cancer progression and metastasis: self-organizing attractors in cell regulatory networks. *Breast Dis* 26: 27–54.
- Scheres B (2001) Plant cell identity. The role of position and lineage. *Plant Physiol* 125: 112–114.
- Rudall P (1987) *Anatomy of flowering plants: An introduction to structure and development*. Edward Arnold, USA.
- Chaos A, Aldana M, Espinosa-Soto C, Garcia-Ponce B, Garay-Arroyo A, Alvarez-Buylla ER (2006) From genes to Flower patterns and evolution: dynamic models of gene regulatory networks. *Plant Growth Regul* 25: 278–289.
- Coen E, Meyerowitz E (1991) The war of the whorls: genetic interactions controlling flower development. *Nature* 353: 31–37.
- Hume DA (2000) Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. *Blood* 96: 2323–2328.
- Fedoroff N, Fontana W (2002) Genetic networks. Small numbers of big molecules. *Science* 297: 1129–31.
- Paulsson J (2004) Summing up the noise in gene networks. *Nature* 427: 415–8.
- Martínez-Arias A, Hayward P (2006) Filtering transcriptional noise during development: concepts and mechanisms. *Nature Reviews (Genetics)* 7: 34–44.
- Wang J, Huang B, Xia X, Sun Z (2006) Funneled landscapes leads to robustness of cell networks: yeast cell cycle. *PLoS Computational Biology* 2(11): e147.
- Wang J, Huang B, Xia X, Sun Z (2006) Funneled landscapes leads to robustness of cellular networks: MAPK signal transduction. *Biophysical Journal*: Biophysical Letters doi: 10.1529/biophysj.106.086777: L54–L56.
- Han B, Wang J (2007) Quantifying robustness and dissipation cost of yeast cell cycle network: the funneled energy landscape perspectives. *Biophysical Journal* 92: 3755–3763.
- Kaneko K (2007) Evolution of robustness to noise and mutation in gene expression dynamics. *PLoS ONE* 2: e434.
- Waddington CH (1957) *The Strategy of the Genes*. London: Geo Allen & Unwin.
- Slack JM (2002) Conrad Hal Waddington: the last Renaissance biologist? *Nat Rev Genet* 3: 889–95.
- Sasai M, Wolynes PG (2003) Stochastic gene expression as a many-body problem. *Proc Natl Acad Sci U S A* 100: 2374–9.
- Aldrich J (1997) R.A. Fisher and the making of maximum likelihood 1912–1922. *Statist Sci* 12: 162–176.
- Huang K (1990) *Statistical Mechanics*. Wiley, John & Sons.
- Yu H, Ito T, Zhao Y, Peng J, Kumar P, Meyerowitz EM (2004) Floral homeotic genes are targets of gibberellin signaling in flower development. *Proc Natl Acad Sci USA* 101: 7827–7832.
- Bombliès K, Dagenais N, Weigel D (1999) Redundant enhancers mediate transcriptional repression of AGAMOUS by APETALA2. *Dev Biol* 216: 260–264.
- Glass L (1975) Classification of biological networks by their qualitative dynamics. *J Theor Biol* 54: 85–107.
- Wang Z, Hou Z, Xin H, Zhang Z (2006) Engineered internal noise stochastic resonator in a gene network: A model study. *Biophys Chem* 125: 281–285.
- Bowman J, ed. (1994) *Arabidopsis*. An Atlas of Morphology and Development. U.S.A.: Springer-Verlag.

32. Schmid M, Uhlenhaut NH, Godard F, Demar M, Bressan R, et al. (2003) Dissection of floral induction pathways using global expression analysis. *Development* 130: 6001–12.
33. Whipple CJ, Ciceri P, Padilla CM, Ambrose BA, Bandong SL, Schmidt RJ (2004) Conservation of B-class floral homeotic gene function between maize and *Arabidopsis*. *Development* 131: 6083–91.
34. Adam H, Jouannic S, Morcillo F, Verdeil JL, Duval Y, Tregear JW (2007) Determination of flower structure in *Elaeis guineensis*: do palms use the same homeotic genes as other species? *Ann Bot* 100: 1–12.
35. Li F, Long T, Lu Y, Ouyang Q, Tang C (2004) The yeast cell-cycle network is robustly designed. *Proc Natl Acad Sci U S A* 101: 4781–4786.
36. Escudero C (2006) Stochastic resonance due to internal noise in reaction kinetics. *Phys Rev E Stat Nonlin Soft Matter Phys* 74: 010103.
37. Sigal A, Milo R, Cohen A, Geva-Zatorsky N, Klein Y, et al. (2006) Variability and memory of protein levels in human cells. *Nature* 444: 643–6.
38. Gammaitoni L, Hanggi P, Jung P, Marchesoni F (1998) Stochastic Resonance. *Rev Mod Phys* 70: 233.
39. Baurle I, Dean C (2006) The timing of developmental transitions in plants. *Cell* 125: 655–64.
40. Ambrose BA, Espinosa-Matías S, Vázquez-Santana S, Vergara-Silva F, Martínez E, Márquez-Guzmán J, Álvarez-Buylla ER (2006) Comparative developmental series of the Mexican Triurids support a euanthial interpretation for the unusual reproductive axes of *Lacandonia schismatica* (Triuridaceae). *Am J Bot* 93: 15–35.
41. Lucas WJ (1995) Plasmodesmata: intercellular channels for macromolecular transport in plants. *Curr Opin Cell Biol* 7: 673–80.
42. Sessions A, Yanofsky MF, Weigel D (2000) Cell-cell signaling and movement by the floral transcription factors *LEFTY* and *APETALAI*. *Science* 289: 779–781.
43. Parcy F, Nilsson O, Busch MA, Lee I, Weigel D (1998) A genetic framework for floral patterning. *Nature* 395: 561–566.
44. Wu X, Dinnyen J, Crawford K, Rhee Y, Citovsky V, Zambryski P, Weigel D (2003) Modes of intercellular transcription factor movement in the *Arabidopsis* apex. *Development* 130: 3735–3745.
45. Lohmann J, Hong RL, Hobe M, Busch MA, Parcy F, et al. (1995) Selective trafficking of *KNOTTED1* homeodomain protein and its mRNA through plasmodesmata. *Science* 270: 1980–1983.
46. Lenhard M, Bohnert A, Jürgens G, Laux T (2001) Termination of stem cell maintenance in *Arabidopsis* floral meristems by interactions between *WUSCHEL* and *AGAMOUS*. *Cell* 105: 805–14.
47. Gershenson C (2004) Updating Schemes in Random Boolean Networks: Do They Really Matter? In: Pollack J, Bedau M, Husbands P, Ikegami T, Watson RA, eds. *Artificial Life IX, Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*. MIT Press. pp 238–243.
48. Álvarez-Buylla ER, Benítez M, Dávila EB, Chaos A, Espinosa-Soto C, Padilla-Longoria P (2007) Gene regulatory network models for plant development. *Curr Opin Plant Biol* 10: 83–91.

Capítulo 4

Conclusiones generales y perspectivas

Como hemos visto en los capítulos previos, el modelo de redes nos proporciona un mejor acercamiento para ligar el comportamiento y la dinámica de los sistemas genéticos, lo que se denomina como el genotipo, con las características que se generan de su funcionamiento, a lo que se conoce como el fenotipo. Además, suministra un sustento teórico y metodológico sobre el cual se pueden poner a prueba hipótesis sobre la evolución y el desarrollo de los seres, partes fundamentales para comprender cómo es que se comportan en el tiempo y en el espacio. La manera en que incorpora la información de cada uno de los genes produce que los resultados no sean aditivos ni lineales, y mucho menos predecibles, lo cual descubre que los sistemas biológicos se comportan de manera compleja y que su estudio es mucho más difícil de lo esperado anteriormente, es decir, que para comprender a una criatura no basta con conocer sus elementos y suponer que se comportará como la suma de ellos, sino que su comportamiento integrado tendrá características insospechadas si se aborda bajo un esquema reduccionista.

Sin embargo, al abordar el estudio de los sistemas biológicos bajo este paradigma se enfrentan problemas de tipo metodológico serios. El espacio omega de redes pequeñísimas es inmenso, por ejemplo, una red binaria con 15 nodos, como la analizada en este estudio, tiene un espacio omega de $2^{15}=32,767$ configuraciones. Dado que el espacio omega se incrementa exponencialmente según el número de nodos, resulta intratable estudiar sistemas con muchos nodos, ya que la velocidad y el tipo de cómputo de los que disponemos son muy limitados para este tipo de problemas. Sin duda el advenimiento de computadoras más veloces y con procesamiento en paralelo ayudará para expandir un poco la frontera del estudio de redes mayores. La invención de algoritmos más eficaces contribuirá a la misma tarea. Por ejemplo, el hombre puede tener unos 40,000 genes, el espacio omega de este sistema es de $2^{40,000}$ ó $10^{12,000}$. Nada más para tener una idea de esta magnitud, el número de partículas del universo conocido es de 10^{80} . Si se abstrae a los seres vivos como ensambles de redes, la forma de actuar se ve claramente, lo que se necesita es un escalpelo que sepa dónde cortar a las redes que se han unido para conformar a un organismo. Algo similar a lo que se hizo cuando se empezó a estudiar el cuerpo humano, primero se disecaron e identificaron sus partes o sistemas y luego se estudiaron por separado. Determinar qué tipo de redes son susceptibles de unirse a otras será fundamental para comprender la vida. Probablemente dentro de estas redes habrá nodos que sirvan como lazos potenciales por medio de los cuales se ligan con nodos de otras redes y así originen un ensamble de redes. La caracterización de estos nodos, los nodos eslabones, si es que en ellos reside la capacidad de ensamble, será necesaria para desmembrar a los organismos en sistemas tratables de acuerdo a las capacidades tecnológicas e intelectuales de las que disponemos. Mediante este escalpelo que desenreda la maraña podremos separar a estas congregaciones gigantescas de genes llamadas seres vivos.

Se han realizado algunos esfuerzos para poder distinguir estos módulos, pero el acercamiento ha sido más bien desde el punto de vista de la sincronización, lo cual identifica a los módulos de acuerdo con un comportamiento repetitivo. Otros lo hacen

tomando en cuenta el número de las conexiones que tienen los nodos entre ellos, ya que supone que la identidad de los módulos está relacionada con nubes de conexiones más densas. El problema con estos métodos, en el mejor de los casos, es que sólo toman en cuenta la función de los elementos, cuando es sabido que la función distorsiona la historia real, es decir, la evolución.

Al suponer que dos organismos muy simples, cada uno conformado por una red, se fusionan, la restricción básica que tendrán será la de conservar la mayoría de sus sumideros originales. Además, si se crean sumideros nuevos, un evento de simbiosis habrá nacido. En este ejemplo se ve claramente la restricción evolutiva al unir redes. Es necesario mantener la mayoría de los sumideros originales, entendidos como características en el organismo, al juntar redes.

De la misma forma será necesario desarrollar análisis que se interpreten fácilmente y rápidamente debido a la cantidad avasalladora de datos que surgen tan sólo cuando se calcula la dinámica de un sistema genético pequeño. Con el programa *Atalia* que se presenta aquí y con el cual se realizaron todas las simulaciones, es posible generar análisis visuales novedosos sobre el efecto de las mutaciones en el sistema, la estabilidad de los sumideros y hacer exploraciones sobre el morfoespacio, entre otros más, de manera sencilla, visual e intuitiva.

Las investigaciones sobre el desarrollo cuentan ahora con modelos formales sobre los cuales se pueden empezar a modelar las configuraciones que se suceden durante las etapas de formación de los organismos. Por ejemplo, en el caso de *A. thaliana* queda la pregunta para explorar que se relaciona con el tamaño de las cuencas de atracción. Recordemos que de los diez sumideros, tres de ellos corresponden a los estambres y a los carpelos, que son las estructuras reproductoras de la planta cuya importancia evolutiva es crucial. Las cuencas de atracción de estas estructuras son inmensas, si se suman el tamaño es aproximadamente de 90% del espacio omega. Es interesante esta observación, de la cual surge la pregunta siguiente ¿es mera coincidencia que las cuencas de atracción de las estructuras más importantes para la planta sean abrumadoramente mayores que todas las demás? Tal parece, bajo un modelo determinista, que el tamaño de las cuencas no es trascendental si existe una señal determinista y robusta que dirige el sistema a través de todos los sumideros. Sin embargo, si la dinámica que observamos es el resultado colateral e indirecto de fluctuaciones estocásticas sobre las reglas de actualización, entonces las cuencas de atracción de las estructuras reproductoras podrían estar relacionadas con efectos de selección natural, ya que son gigantescas comparándolas con las de las demás estructuras florales. El sondeo de las configuraciones transitorias y su significado en el modelo de redes genéticas binarias queda todavía pendiente.

Como se mencionó en el Capítulo 2, será también necesario investigar si el tamaño o la forma de las cuencas de atracción tienen relevancia y cuál es su conexión con la actualización del sistema. La visión, aparentemente más realista, de que los genes cambien de estado de forma asíncrona, deberá ser explorada con muchísimo más detalle; no obstante, si las redes son lo suficientemente robustas, los tamaños y las estructuras de las cuencas de atracción podrían pasar a segundo plano, entonces lo que se debería de encontrar es la fuente de esta resistencia intrínseca del sistema.

Otra línea de investigación que puede dar resultados sobre lo anterior consiste en estudiar las restricciones propias de una red en particular. La identificación de la propiedad que acota la dinámica del sistema sin duda será importantísima para comprender los ensambles de redes. Se requieren estudios para saber qué tanto la topología de una red y/o

las reglas restringen ciertas combinaciones de sumideros. De esta forma se podrá ligar a una discusión vieja y generar exploraciones muy prometedoras: la discusión sobre los morfoespacios. Descubriendo restricciones de acuerdo a la topología, a las reglas o a otras circunstancias se podrán mapear sobre el morfoespacio teórico las posibilidades reales sobre las cuales la selección natural podrá actuar. Tanto los espacios ocupados del morfoespacio como los desocupados, como en el caso de las conchas, podrán ser explicados y calculados proveyéndonos de una visión evolutiva profunda. Esta disputa teórica personalizada últimamente por Dawkins por un lado, y por Gould por el otro (Sterelny 2001), podrá tener el inicio de una respuesta basada en métodos experimentales. Recordemos la querrela. Dawkins arguye que el campo de acción de la selección natural es inmenso, ilimitado (*Escalando el monte improbable* [*Climbing mount improbable*] 1996). Hablando en términos del morfoespacio quiere decir que cualquier configuración puede ser un sumidero si se dan las mutaciones correctas, no hay restricciones históricas que restrinjan algunas configuraciones de ser sumideros. Bajo el ejemplo de la red floral de *A. thaliana*, la idea doquiniana asegura que no importa cómo se haya formado la flor, ni cuál haya sido su camino evolutivo, con algunas mutaciones, cualquier configuración podría ser un sumidero. La postura antagonista (Gould y Lewontin 1979), señala que a medida que una población de organismos evoluciona, se van seleccionando características, sumideros en términos de este estudio, se van seleccionando interacciones en una red, entonces automáticamente quedan otras posibilidades exentas de poder ser sumideros. Ejemplificando con el caso de la flor, la posición gouldiana afirmarí que debido a la topología y a las reglas lógicas que conforman la red de *A. thaliana*, que han sido resultado de la evolución, quedarán restringidas automáticamente muchas de las configuraciones de la posibilidad de ser sumideros. Esta pregunta no poseía un escenario formal para empezar a contestarla antes contar con redes fundamentadas en datos experimentales para sistemas biológicos particulares.

De aquí surge otra pregunta básica que tiene que ver justamente con la formación de las redes biológicas. ¿En realidad la selección natural ha escogido la estructura de estas redes o, en cambio, esta organización se debe a productos colaterales de ensamble? Si la selección natural dirige exclusivamente el proceso de construcción de las redes biológicas, habría elementos constructivos más abundantes que otros.

Ciertos estudios han medido si dentro de las redes biológicas hay estructuras recurrentes más allá de los que se esperaría por el azar (Milo *et al.* 2004). Se ha encontrado que en estas redes, además de otras de diversa índole, presentan motivos representados con mayor y con menor frecuencias de los esperado, es por ello que se ha sugerido que cuando un motivo es más abundante en cierto tipo de red se debe a que brinda alguna ventaja adaptativa para el organismo, de manera contraria cuando se le halla con menor proporción de los esperado, es muy probable que la selección natural haya actuado en su contra para eliminarlos.

Sin embargo, puede haber otra explicación. Que haya más motivos que otros en las redes no necesariamente implica un proceso de elección dirigido por la selección natural, si no que puede deberse a que es un efecto colateral de la manera en que las redes se construyen, serían pechinas moleculares. Valiéndose de la duplicación de motivos y divergencia de funciones Kuo *et al.* (2006) crean una red en la cual hay abundancia de unos motivos y escasez de otros. Se basan en estos resultados para afirmar que la estructura de las redes biológicas se puede deber en parte a la forma en que se van creando en lugar de

haber sido moldeadas nada más por la selección natural. Sin duda queda mucho trabajo que hacer para explorar esta idea y llegar a unas conclusiones sólidas.

Si la organización de las redes no se debe a la selección natural y son de alguna manera pechinas, se tendrán evidencias muy importantes, que junto con los demás resultados presentados aquí y en los trabajos citados, para comenzar a remendar la síntesis moderna, ya que el papel que representa la selección natural podría caer a un segundo plano muy detrás de la estabilidad de los sistemas complejos y de las características que de ellos emanan.

Ciertas ideas que ha evadido responder la síntesis moderna, como los ejemplos de los monstruos esperanzados, de los cuales tenemos un ejemplar en México que mora en la selva lacandona, y la ortogénesis, podrían tener explicación bajo un nuevo maridaje: la auto organización y la selección natural, siendo la auto organización la que dispone el menú biológico y la selección natural quien elige el vino de acuerdo a los alimentos.

Apéndices

Apéndice 1. Poema «Vientos del pueblo».

Miguel Hernández nació en Orihuela, España, en 1910. De pequeño era pastor de cabras, oficio que desempeñó a pesar de las recomendaciones de sus profesores hacia su padre de que le costeara una carrera dado su talento e inteligencia grandes. A fines de 1931 viaja a Madrid con una maleta llena de versos y la lectura de todos los clásicos que cayeron en sus manos. Tiempo después, Hernández se consagra como el más hondo y original poeta de la generación del 36. Grandeza que trunca el franquismo. El poeta morirá en la cárcel en 1942 (Introducción de J.L. Cano en: Hernández. Poemas, 1973).

El poema *Vientos del pueblo* fue publicado por vez primera en la revista madrileña *El mono azul*, (número 9, Madrid, 22 de octubre de 1936) por el bando republicano durante la Guerra Civil Española. Dicha revista servía como medio de comunicación para los milicianos y, sobre todo, como un aliento para aquellos que defendían los valores de la República en contra del fascismo. Algunos colaboradores de la revista fueron: Rafael Alberti, Antonio Machado, André Malraux y Pablo Neruda. Posteriormente, el poema apareció en el libro de poemas *Viento del pueblo* (1936-1937).

Vientos del pueblo¹

Vientos del pueblo me llevan,
vientos del pueblo me arrastran,
me esparcen el corazón
y me aventan la garganta.

Los bueyes doblan la frente,
impotentemente mansa,
delante de los castigos:
los leones la levantan
y al mismo tiempo castigan
con su clamorosa zarpa.

No soy de un pueblo de bueyes,
que soy de un pueblo que embargan
yacimientos de leones,
desfiladeros de águilas
y cordilleras de toros
con el orgullo en el asta.

Nunca medraron los bueyes

¹ Texto obtenido del disco editado por Gravure Universelle: *Espagne. El hombre nuevo cantando*. De la colección: *Le chant du monde. Le nouveau chansonnier international*.

en los páramos de España.

¿Quién habló de echar un yugo
sobre el cuello de esta raza?

¿Quién ha puesto al huracán
jamás ni yugos ni trabas,
ni quién al rayo detuvo
prisionero en una jaula?

Asturianos de braveza,
vascos de piedra blindada,
valencianos de alegría
y castellanos de alma,
labrados como la tierra
y airoso como las alas;
andaluces de relámpagos,
nacidos entre guitarras
y forjados en los yunques
torrenciales de las lágrimas;
extremeños de centeno,
gallegos de lluvia y calma,
catalanes de firmeza,
aragoneses de casta,
murcianos de dinamita
frutalmente propagada,
leoneses, navarros, dueños
del hambre, el sudor y el hacha,
reyes de la minería,
señores de la labranza,
hombres que entre las raíces,
como raíces gallardas,
vais de la vida a la muerte,
vais de la nada a la nada:
yugos os quieren poner
gentes de la hierba mala,
yugos que habéis de dejar
rotos sobre sus espaldas.

Crepúsculo de los bueyes
está despuntando el alba.

Los bueyes mueren vestidos
de humildad y olor de cuadra:
las águilas, los leones
y los toros de arrogancia,
y detrás de ellos, el cielo
ni se enturbia ni se acaba.

La agonía de los bueyes
tiene pequeña la cara,
la del animal varón
toda la creación agranda.

Si me muero, que me muera
con la cabeza muy alta.
Muerto y veinte veces muerto,
la boca contra la grama,
tendré apretado los dientes
y decidida la barba.

Cantando espero a la muerte
que hay ruiseñores que cantan
encima de los fusiles
y en medio de las batallas.

El ejemplo que se muestra de una red estática en la Figura 3 corresponde a este poema. Los nodos representan solamente a las palabras diferentes del poema y las conexiones indican si dos palabras están contiguas en el texto. La red se construyó siguiendo el algoritmo general siguiente:

- 1-Se pone un nodo representando la palabra primera del poema.
- 2-Se pone otro nodo que representa la palabra segunda del poema.
- 3-Se traza una flecha desde la palabra primera a la segunda.
- 4-Se lee la palabra siguiente, si es diferente a las anteriores, se traza una flecha desde el nodo último hacia un nodo nuevo. Por el contrario, si la palabra ya está representada por un nodo, se traza la flecha desde el nodo último dibujado hacia el nodo que ya existe y que simboliza a la palabra repetida.
- 5- Se repite el paso 4 hasta terminar.

Posteriormente se registran medidas diversas de conectividad de los nodos. De esta forma se puede identificar cuáles son las palabras que funcionan como concentradores y así generalizar las propiedades del escrito desde este acercamiento. La Tabla 9 presenta las cuatro palabras más conectadas. El análisis se realizó con el programa *Enne* escrito por el autor y que está a disposición en <http://www.ecologia.unam.mx/~achaos/Enne/enne.htm>. La red de la Figura 3 se realizó con el programa UCINET.

Tabla 1. Palabras con más conexiones (concentradores) del poema «Vientos del pueblo».

Palabra	Conexiones
De	61
La	32
Y	28
Los	17

Apéndice 2 «ATALIA»

Las simulaciones hechas en los trabajos anteriores con redes discretas fueron realizadas con el programa *Atalia* escrito por el autor. El nombre de *Atalia* es una especie de acrónimo de la *A. thaliana*.

Atalia tiene un sitio en la Internet y se piensa en distribuirlo de forma gratuita como un instrumento de investigación y de docencia. La versión actual disponible es la 1. La página se encuentra por el momento en la dirección siguiente: <http://www.ecologia.unam.mx/~achaos/Atalia/atalia.htm>. Otro programa que hace algunos análisis similares los de *Atalia* es el DDLab escrito por Andrew Wuensche. Sin embargo, se optó por escribir uno nuevo debido a que el DDLab es más lento, no trabaja bajo un sistema de ventanas y se ejecuta en sistemas de línea de comandos, como DOS y Linux, haciéndolo extremadamente anti amigable.

De manera sucinta se presentan a continuación algunas de las capacidades de *Atalia* además de las expuestas en el Capítulo 2. El ejemplo que se tomó para hacerlas es el de la red floral de *A. thaliana*, eje conductor de todo el trabajo.

Despliegue de la red introducida

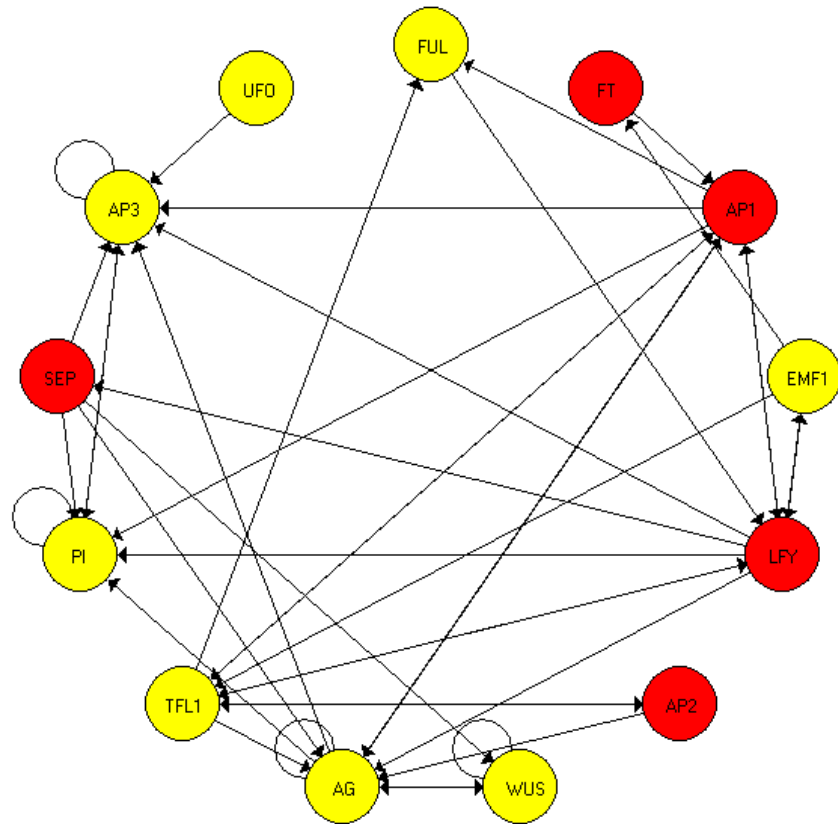
En la Figura 14 se presenta la red de la red floral de *A. thaliana*. Además es posible mostrar la configuración de cierto sumidero, por ejemplo, este caso corresponde al sumidero de los sépalos cuya configuración es: 0110110000100. El color amarillo significa que el nodo está inactivo y el rojo que está activo.

Despliegue de los sumideros de la red

La red floral de *A. thaliana* tiene diez sumideros y cada uno de ellos posee una cuenca de atracción (Figura 15). Las hay pequeñas como la primera de los pétalos, cuya cuenca es de ocho configuraciones; medianas, como la de los sépalos, que está conformada por 152 configuraciones; hasta las inmensas, compuestas por más de tres mil configuraciones, como son las de los carpelos y una de los estambres.

Estabilidad de un sumidero

Es posible que al cambiar un nodo de una configuración correspondiente a un sumidero se pueda caer en la ruta de otro sumidero, pero también podría caerse en una configuración que pertenece al mismo sumidero. Cuando la mayoría de los cambios en la configuración original nos llevan a otro sumidero se dice que el sumidero es muy inestable. Se muestra en la Figura 16 la estabilidad del sumidero de los sépalos a una distancia Manhattan=1. Gracias al esquema se puede saber rápidamente cuáles son los nodos que una vez alterados nos llevarán a otro sumidero y cuáles no. En esa figura se aprecia de color verde a aquellos genes que no provocarán un cambio de sumidero si se les cambia. Por el contrario, si se altera el estado de los genes iluminados en magenta la configuración resultante el sistema desembocará en otro sumidero, de igual forma pasará con el gen en granate. Por lo tanto, la estabilidad de este sumidero es de 7/13.



Atalia

Figura 1. La red floral de *A. thaliana*. Se muestra la configuración del sumidero de los sépalos, los genes en rojo están activos, los amarillos inactivos.

Espacios de las formas sobre las configuraciones genéticas (Espacios FCG)

El espacio omega de *A. thaliana* es difícil de analizar principalmente porque no se disponen de métodos formales para hacerlo debido a que no está muy claro que es lo que hay que estudiar. Cuando esto sucede es necesario echar un vistazo a todos los resultados, una manera de hacerlo es la que se propone aquí utilizando una especie de tapetes o mosaicos que registran información básica y la muestran de forma colorida y simple. Con este método es posible descubrir ciertos diseños que de otra manera estarían ocultos para nuestros ojos. La interpretación de ellos queda todavía por estudiarse con detenimiento, pero eso no invalida la utilidad del método, al contrario prueba su utilidad descubriendo situaciones que de otra forma permanecerían ocultas.

La Figura 17 muestra el destino de todas las configuraciones del espacio omega para la red floral de *A. thaliana*. Cada tesela representa una configuración y están ordenadas a partir del extremo superior izquierdo, que corresponde a la configuración 0000000000, hasta el extremo inferior derecho que es la configuración 1111111111. Este mosaico constituye el mapeo del morfoespacio sobre las configuraciones genéticas, lo que origina un espacio diferente del morfoespacio de Raup, porque incorpora al espacio de las formas el espacio de las configuraciones genéticas, creando un espacio más rico en información. Se propone aquí llamar a esta representación nueva como el *espacio de las formas sobre las configuraciones genéticas* (espacio FCG). Como se aprecia, la mayoría de las configuraciones llegan a dos sumideros, el amarillo y el magenta, que representan los sumideros de los estambres y de los carpelos. Los demás colores representan a los otros ocho sumideros. De esta forma es muy fácil apreciar los tamaños de las cuencas de atracción y saber que un cambio en el último gen, *UFO*, siempre provocará un cambio de sumidero.

En la Figura 18 se representa la distancia Manhattan que hay entre una configuración particular y su destino o sumidero.

En la Figura 19 se representa la distancia topológica que existe entre una configuración dada y su destino o sumidero. Curiosamente las distancias se agrupan formando diseños muy claros y no se aprecia una relación entre distancias topológicas y distancias Manhattan.

Análisis de mutantes

Si se pretende hacer un análisis de mutantes, los datos cuantitativos serán importantes, pero es posible también generar otros resultados de forma cualitativa que aportarán también mucha información al estudio, además de que permiten desplegar la información de una manera clara, concisa y elegante.

En este ejemplo se simuló una mutación en un gen del grupo A, el *API*. La Figura 20 presenta la red floral de *A. thaliana* en donde se señalan los genes ABC. Cuando se muta un gen A el resultado que se espera obtener es una flor sin sépalos ni pétalos. Sin embargo, ¿cómo se altera la dinámica del sistema?, ¿y los destinos del espacio omega?, ¿se crean sumideros nuevos? Para contestar estas preguntas se debería comparar el espacio omega original con el mutante. Aquí se presenta un método con el cual esta comparación se realiza a través de los mosaicos o tapetes presentados anteriormente en la cual se hacen operaciones con ambos tapetes y se da como resultado un tapete en donde se aprecian visualmente estos cambios.

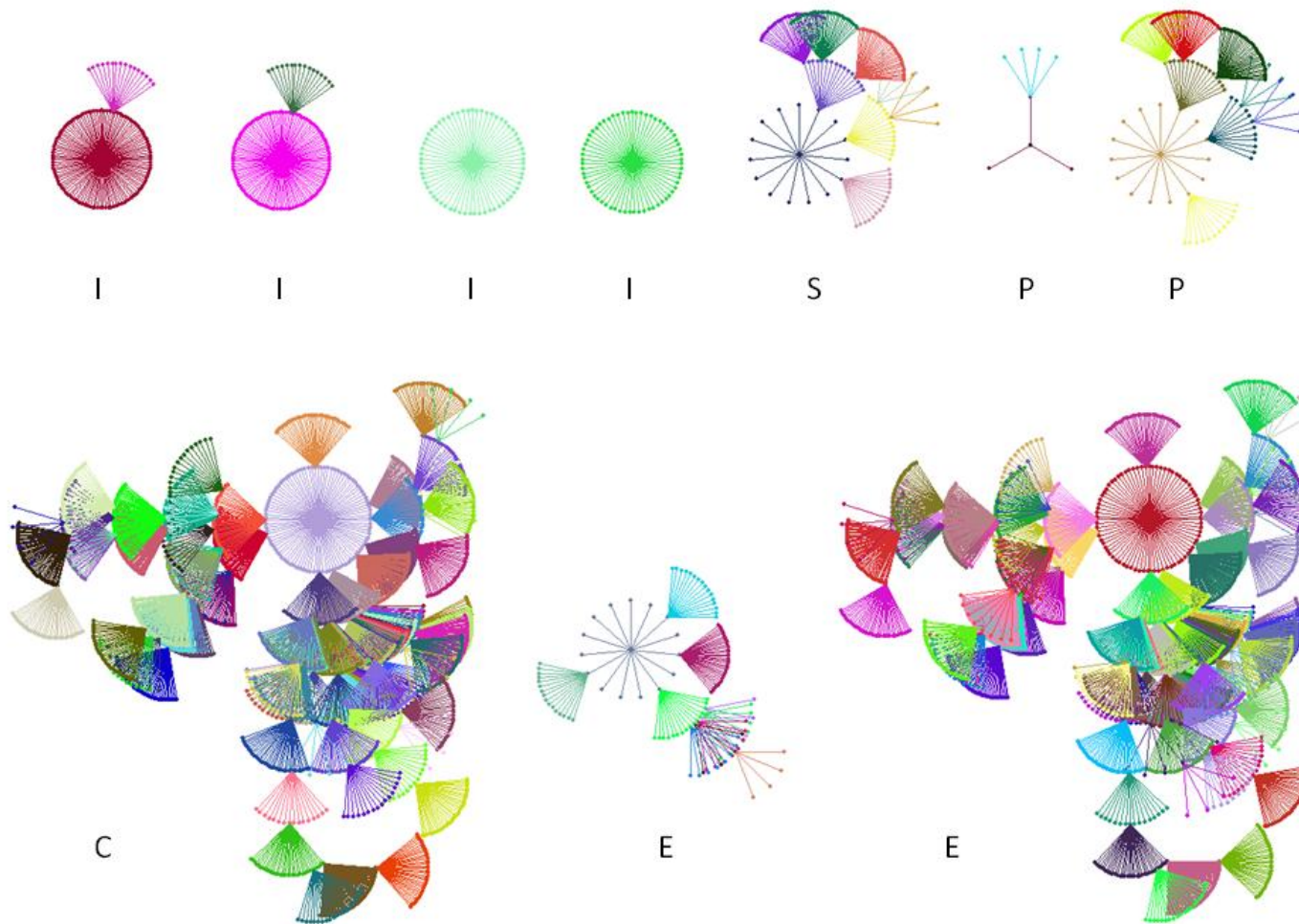


Figura 2. Se presentan las 10 cuencas de tracción de la red de la *A. thaliana*. En el centro de cada una de ellas se encuentra la configuración en equilibrio, el sumidero. Los colores están asignados al azar y solamente tienen valor estético. I: inflorescencia, S: sépalo, P: pétalo, C: carpelo y E: estambre.

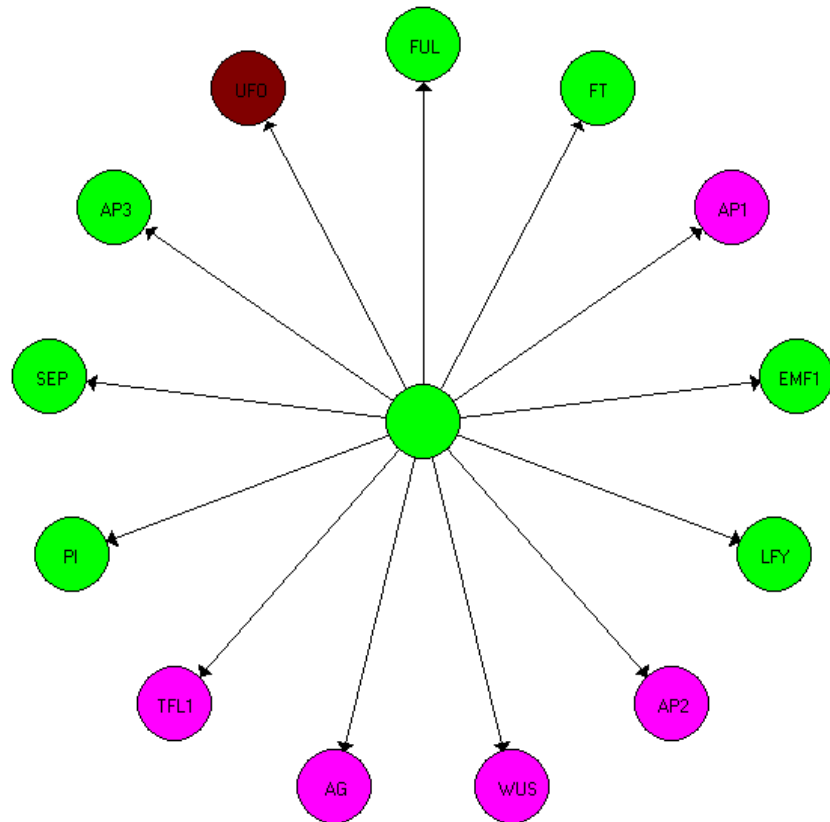
El tapete que se obtiene de hacer la operación comparativa entre el tapete silvestre y el del mutante *API* se muestra en la Figura 21. El amarillo indica que para esa configuración no hubo cambio de destino, es decir, se llegó al mismo sumidero. El morado indica que esa configuración sí cambió de destino, pero llegó a un sumidero que ya existía, por lo tanto no se crearon sumideros nuevos. En caso de que la configuración hubiera llegado a un sumidero nuevo la tesela se hubiera iluminado de rojo indicando este cambio. Rápidamente se percibe que son muy pocas las configuraciones que cambian su destino, lo cual es razonable porque las cuencas de atracción de los sépalos y de los pétalos son pequeñas.

Rastreador de configuraciones

Atalia tiene la capacidad de rastrear cualquier configuración hasta que llegue al sumidero que le corresponde, ya sea puntual o cíclico. Presenta los resultados de forma gráfica indicando las configuraciones por las que se transitará hasta llegar al destino final como se muestra en la Figura 22. En ese caso se parte de la configuración 0000000000000000 y se acaba en el sumidero 1100110101100, que es el de los carpelos.

Rastreador del espacio FCG

Otra capacidad de *Atalia* es que permite rastrear qué configuraciones del espacio FCG podrían ser sumideros si se perturba el sistema con mutaciones pequeñas. En la Figura 23 se presenta el espacio omega de la red floral de *A. thaliana*. Las teselas azules representan a los sumideros originales, las rojas son aquellas configuraciones que podrían ser sumideros bajo un régimen de mutaciones particular. Las blancas son las configuraciones que jamás serían sumideros.



Atalia

Figura 3. Estabilidad del sumidero de los sépalos a una distancia Manhattan de 1. Los círculos de color verde indican que cuando se muta al gen que aparece dentro del mismo círculo el destino al que se llegará no cambia, seguirá siendo un sépalo. Los círculos de color magenta indican que la mutación en ese gen producirá que ahora se finalice en carpelos. El círculo en granate señala que un cambio de estado de *UFO* conducirá a que la célula finalice en un pétalo.



Figura 4. Espacio FCG. El color de las teselas indica el destino celular final de la configuración. Amarillo: estambres, magenta: carpelos, azules: inflorescencias, verdes: sépalos y rojos: pétalos.

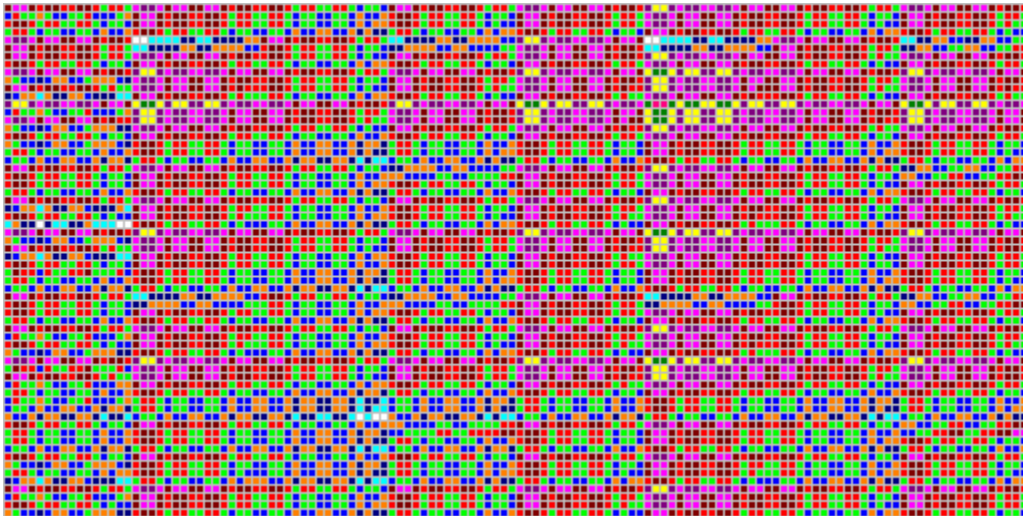


Figura 5. Distancias Manhattan al sumidero. Las distancias van de 0 a 12.

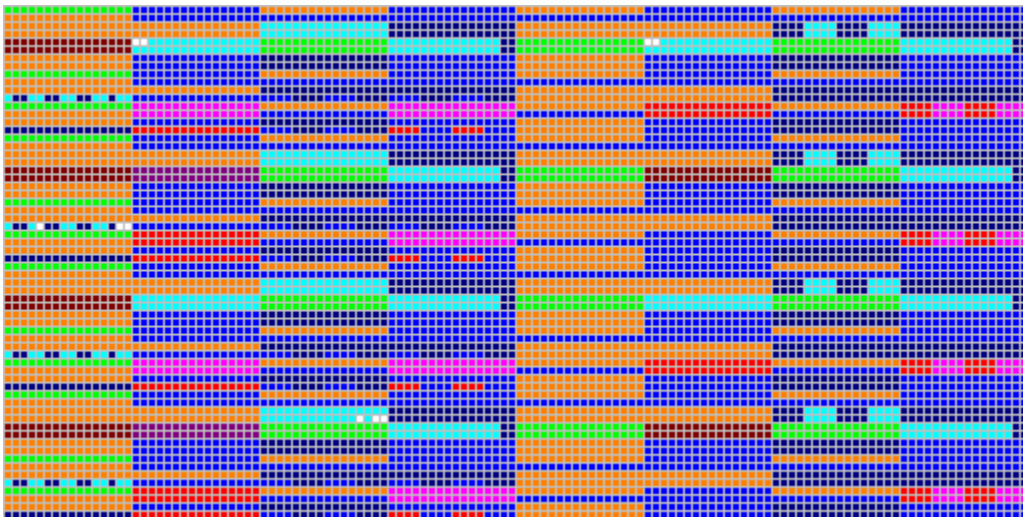
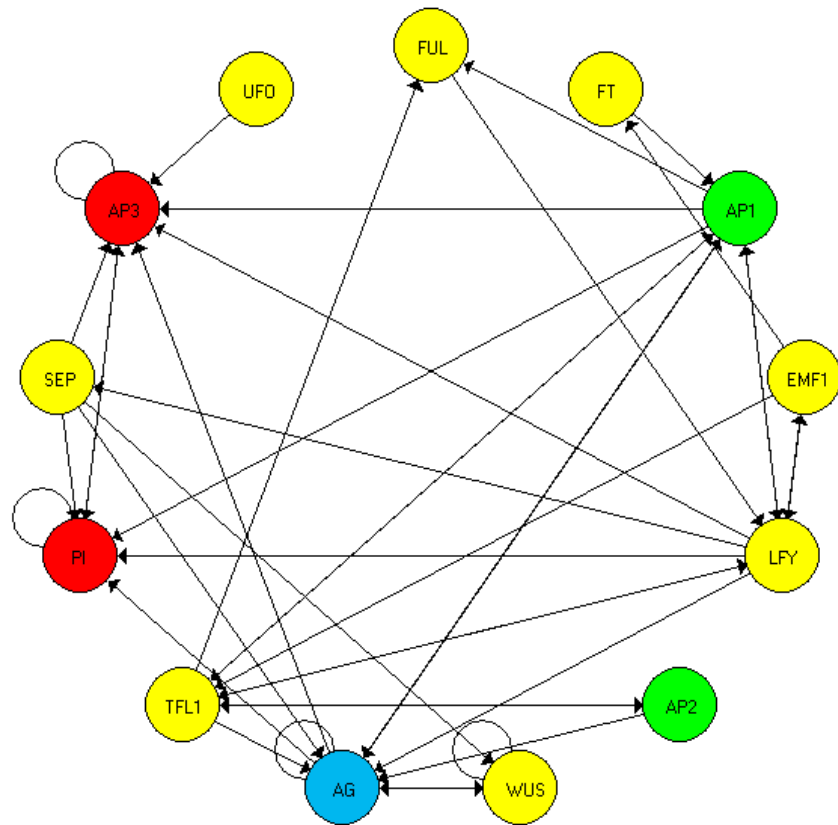


Figura 6. Distancias topológicas al sumidero. Las distancias van de 0 a 9.



Atalia

Figura 7: Red floral de *A. thaliana* donde se señalan cuáles son los genes ABC. En verde se representan a los genes del grupo A, en rojo a los del grupo B y en azul al del grupo C.

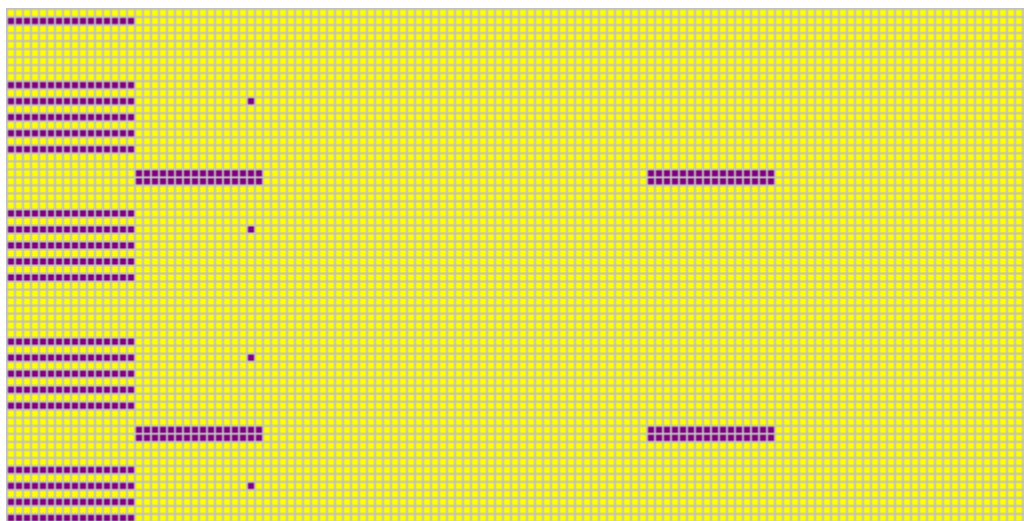


Figura 8. Mosaico comparativo entre el mosaico silvestre y el mutante. Las teselas amarillas muestran que esa configuración llegó al mismo sumidero, las moradas indican que cambiaron de sumidero, pero llegaron a uno que ya existía.

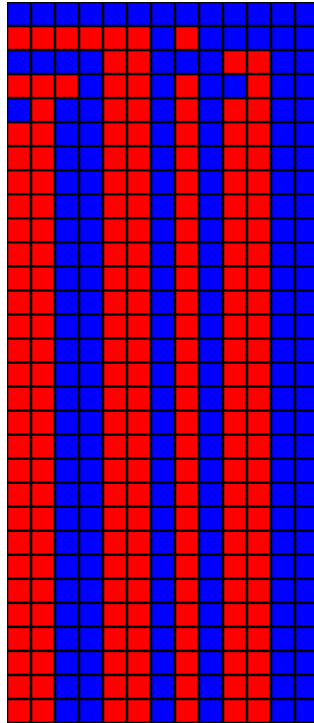


Figura 9. Rastreo de la configuración 00000000000000 hasta su destino final (sumidero) 1100110101100.



Figura 10. Configuraciones que podrían ser sumideros bajo un régimen de mutaciones. Las teselas azules son los sumideros originales, las rojas representan los sumideros nuevos después de la mutación y las blancas son las configuraciones que no serían sumideros.

Referencias generales

Acosta, J. de (1590). *Historia natural y moral de las Indias*. [1962] Fondo de cultura económica. Segunda edición revisada. México.

Albert, R. y Ohtmer, H. 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.

Aldana, M. y Cluzel, P. 2003. A natural class of robust network. *PNAS* 100(15):8710-8714.

Balleza, E., Álvarez-Buylla, E.R., Chaos, Á., Kauffman, S., Shmulevich, I. y Aldana, M. (2008) Critical dynamics in genetic regulatory networks: examples from four kingdoms. *PLoS ONE*. 2008 Jun 18; 3(6):e2456.

Barabasi, A. y Oltvai, Z. 2004. Network biology: understanding the cell's functional organization. *Nature reviews* 5:101-114.

Behzad, M, Chartrand, G. y Lesniak-Foster, L. (1979). *Graphs & digraphs*. Prindle, Weber & Schmidt International series, Boston, Mass.

Brock, T. y Madigan, M. (1991). *Biology of microorganisms*. Sixth edition. Prentice Hall, New Jersey.

Darwin, C. y Wallace, A.R. 1858. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Journal of the proceedings of the Linnean Society, Zoology*, 3:45-62.

Diestel, R. (2000). *Graph theory*. Electronic edition. Springer Verlag, New York.

Hernández, M. (1973). *Poemas*. Rotativa, Plaza y Janés, S.A. Editores. España.

Coen, E.S. y Meyerowitz, E. 1991. The war of the whorls: genetic interactions controlling flower development. *Nature* 353:31-37

Colacino, C. y Grehan, J. (2003). Suppression at the frontiers of evolutionary biology: León Croizat's case. En: M. Mamone (Ed.). *Atti del convegno scienza e democrazia*, Napoli, April 2001. Hevelius, Benevento, Italia.

Craw, R.C., Grehan, J.R. y Heads, M.J. (1999). *Panbiogeography: Tracking the History of Life*. Oxford University Press, New York.

Croizat, L. (1977). Carlos Darwin y sus teorías. *Bol. Acad. Cienc. Fís. Mat. Nat. Caracas* 37 (n° 113):15-90.

- Dawkins, R. 1996. *Climbing mount improbable*. W.W. Norton & Company Ltd. London.
- Depew, D. y Weber, B. (1995) *Darwinism Evolving. Systems Dynamics and the Genealogy of Natural Selection*. Cambridge, MA & London: MIT Press.
- Deutscher, D., Meilijson, I., Schuster, S. y Ruppin E. 2008. Can single knockouts accurately single out gene functions? *BMC Systems Biology* 2008, 2:50.
- Dobzhansky, T., Ayala, F., Stebbins G.L. y Valentine J.W. (1983) *Evolución*. Ediciones Omega, Barcelona
- Espinosa-Soto, C., Padilla-Longoria, P. y Álvarez-Buylla E. 2004. A gene regulatory network model for cell-fate differentiation during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles. *Plant Cell* 16:2923–2939.
- Futuyma, D. (1998). *Evolutionary biology*. Third edition, Sinauer Associates, Sunderland, Massachusetts.
- Goodwin, B. (1994) *How the Leopard Changed its Spots: The Evolution of Complexity*. Scribner, New York.
- Gould, S.J. y Lewontin R.C. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Society of London* B205 581-598.
- Harary, F. (1972). *Graph theory*. Third edition. Addison Wesley Publishing Company, USA.
- Huxley, J. (1942). *Evolution. The modern synthesis*. [1963] (second edition). George Allen and Unwin limited. Great Britain.
- Jacob, F. y Monod, M. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Molec. Biol.* 3:318.
- Kauffman, S.A. 1969. Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor Biol* 22:437–467.
- Kauffman, S.A. (1985). Self-organization, selective adaptation and its limits: a new pattern of inference in evolution and development. En D.J. Depew y B.H. Weber (eds) *Evolution at a Crossroads: The New Biology and the New Philosophy of Science*. MIT Press.
- Kauffman, S.A. 1991. Antichaos and adaptation. *Sci. Amer.* 265(2):64-70.
- Kauffman, S.A. (1993). *The origins of order: Self-organization and selection in evolution*. Oxford: Oxford University Press.

Kuo, P.D., Banzhaf, W. y Leier A. 2006. Network topology and the evolution of dynamics in an artificial genetic regulatory network model created by whole genome duplication and divergence. *BioSystems* 85, 177–200

Lamarck, J.B. (1963). *Zoological philosophy*. An exposition with regard to the natural history of animals. Hafner publishing company, New York.

Lewin, B. (1987). *Genes III*. Third edition. John Wiley & Sons, New York.

Lozoya, X. (1984). *Plantas y luces en México. La Real expedición científica a Nueva España*. Ediciones del serbal, España.

Mayr, E. (1942). *Systematics and the origin of species*. Columbia University Press.

Mayr, E. (1991). *One Long Argument: Charles Darwin and the Genesis of Modern Evolutionary Thought*, Harvard University Press, Cambridge Massachusetts.

Mendoza, L. y Álvarez-Buylla, E. 1998. Dynamics of the genetic regulatory network of *Arabidopsis thaliana* flower morphogenesis. *J. Theor. Biol.* 193, 307-319.

Mendoza, L., Thieffry, D. y Álvarez-Buylla E. 2000. Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis. *Bioinformatics* 15, 593-606.

Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M, y Alon, U. 2004. Superfamilies of Evolved and Designed Networks. *Science* 303:1538-1542.

Nelson, G. 1978. From Candolle to Croizat: Comments on the history of biogeography. *Journal of the History of Biology*, (11)2:269-395.

Newman, M.E.J. (2003). The structure and function of complex networks. *arXiv:cond-mat/0303516v1*

Pelaz, S., Tapia-López, R., Álvarez-Buylla, E.R. y Yanofsky, M.F. (2001) Conversion of leaves into petals in *Arabidopsis*. *Current Biology* 11, 182-184.

Raup, D.M. 1966. Geometric analysis of shell coiling: general problems. *Journal of Paleontology* 40: 1178-1190.

Ridley, M. (2004). *Evolution*. Third edition. Blackwell science, Oxford, UK.

Simpson, J.G. (1944). *Tempo and mode in evolution*. Columbia University Press.

Solé, R. y Goodwin, B. (2000) *Signs of life. How complexity pervades biology*. Basic Books, New York.

Sorensen, S.A. (2004). *Network performance*. Lecture notes for DCNDS Z21. University College London.

- Sterelny, K. (2001). *Dawkins vs Gould. Survival of the fittest*. Icon books UK.
- Strickberger, M. (2000). *Evolution*. Third edition. Jones and Bartlett Publishers International, London, UK
- Von Dassow, G., Meir, E., Munro, E.M. y Odell, G.M. 2000. The segment polarity network is a robust developmental module. *Nature* 406:188-193.
- Vorzimmer, P. 1963. Charles Darwin and blending inheritance. *Isis*, 54:371-390.
- Waddington, C. H. (1940). *Organisers and Genes*, Cambridge University Press.
- Wagner, A. (2007). *Robustness and evolvability in living systems*. Oxford University Press.
- Williams, K.S. y Gilbert, L.E. 1981. Insects as selective agents on plant vegetative morphology: egg mimicry reduces egg laying by butterflies. *Science* 212:467-469.