UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

INSTITUTO DE ECOLOGIA

MODELOS DE MORFOGENESIS EN *Arabidopsis thaliana*

TESIS QUE PRESENTA

LUIS ANTONIO MENDOZA SIERRA

PARA OBTENER EL GRADO DE

DOCTOR EN CIENCIAS BIOMEDICAS

MEXICO, D. F. ENERO 2000

Esta tesis está dedicada a todos aquellos quienes luchan por transformar a la biología en una *teoría* de los seres vivos

**Eighteenth-century science, following the Newtonian revolution,
has been characterized as developing the sciences of organized
simplicity, nineteenth-century science, via statistical mechanics,
as focusing on disorganized complexity, and twentieth-
and twenty-first-century science as confronting organized complexity.
Nowhere is this confrontation so stark as in biology.
Nowhere are new conceptual tools so deeply needed.**

**Stuart Kauffman
-The Origins of Order, 1993**

INDICE

PROLOGO

Esta tesis es el producto del trabajo realizado durante mi estancia en el Laboratorio de Genética Molecular y Evolución del Instituto de Ecología, UNAM, bajo la dirección de la Dra. Elena Alvarez-Buylla. Durante los poco más de cuatro años en el laboratorio, tuve la oportunidad de desarrollarme en diversos aspectos de mi vida académica. En la presente tesis se encuentran sólo algunos de los resultados obtenidos durante mi doctorado; a saber, aquellos que podían ser incluidos de manera coherente en forma de tres artículos de investigación. Sin embargo, para ahondar en algunos aspectos básicos de mi proyecto de investigación me pareció necesario incluir además dos artículos de revisión, los cuales se escribieron como parte de las actividades académicas del programa del Doctorado en Ciencias Biomédicas.

Los artículos forman la parte medular de la tesis. Sin embargo, para dar continuidad y mostrar la unidad entre los diferentes resultados expuestos en los artículos, escribí un texto principal con las ideas centrales que guiaron mi proyecto de investigación. Por tanto, los artículos se incluyen como apéndices al texto principal para ampliar y precisar la información expuesta en las secciones 2.2, 3, 5.1, 5.2 y 5.3 de la tesis. Debido a su estructura, esta tesis puede leerse por lo menos de tres maneras distintas. La primer forma es leer todo el escrito, incluyendo los apéndices. Si bien este tipo de lectura tiene la mayor coherencia entre las distintas partes, el lector encontrará una gran redundancia de información entre los artículos y el texto principal de la tesis. De hecho, algunos párrafos del texto principal son traducciones literales de algunas secciones de los artículos. La intención del texto principal es servir de guía de las ideas, pero los resultados y su amplia discusión se presentan en cada uno de los artículos. La segunda forma es la lectura exclusiva de los artículos incluidos en los apéndices. De esta manera se reduce considerablemente la redundancia en la información leída. Pero para no perder continuidad, se recomienda leer en el siguiente orden: Apéndices 2, 3, 4, 1 y 5. Así, el lector encontrará en el primer artículo las características generales de las redes de regulación modeladas como sistemas dinámicos discretos. El segundo y el tercer artículos contienen una aplicación de las redes en la forma de un modelo de regulación genética de la floración en *Arabidopsis*. En el cuarto artículo se presenta la biología de la raíz de *Arabidopsis*. Y finalmente, el quinto artículo contiene un modelo de regulación que controla la diferenciación en la epidermis de la raíz. La tercera alternativa de lectura consiste en leer únicamente los artículos de investigación (Apéndices 3, 4 y 5). De esta última manera es posible prestar atención en las ideas centrales y los resultados obtenidos en el proyecto de investigación.

Por último deseo expresar mi agradecimiento a las personas que hicieron posible esta tesis. En primer lugar a Elena Alvarez-Buylla, por haberme aceptado en su grupo de trabajo y por apoyarme a lo largo de mi estancia en su laboratorio. También les debo mucho a Jaime Lagúnez y a Julio Collado, que si bien su papel oficial fue el de miembros de mi comité tutorial, en la vida diaria se desempeñaron como verdaderos amigos y colaboradores, alentándome y ayudándome en todo momento. Merece mención especial mi amigo Denis Thieffry, quien desde el inicio vio el potencial del proyecto, pero además creyó en mi y contribuyó a la proyección internacional de mis resultados. También deseo dar gracias a los miembros de mi jurado: Jesús Aguirre, Elena Alvarez-Buylla, Rafael Barrio, Germinal Cocho, Wilhelm Hansberg, Patricia León y Pedro Miramontes. Ellos se esforzaron por leer y sugerir algunos cambios a ésta tesis, logrando así hacer más sencilla la lectura de la misma.

1 INTRODUCCION

Existen diversos sistemas biológicos, como *Xenopus laevis*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Arabidopsis thaliana* y *Antirrhinum majus*, en los cuales se ha realizado un gran esfuerzo por comprender la regulación de la expresión genética durante el desarrollo. El esfuerzo se ha centrado sobre el estudio de los genes involucrados en la morfogénesis, ya que los efectos fenotípicos resultantes de las mutaciones con ganancia o pérdida de alguna función ayudan a inferir la función de dichos genes durante el desarrollo. En el caso particular de las plantas, el estudio de la morfogénesis tiene una larga tradición. Baste con mencionar el estudio clásico de Johann Wolfgang von Goethe llamado *Versuch die Metamorphose der Pflanzen zu erklären* de 1790, en el cual expone la tesis de que todos los órganos de las plantas no son sino hojas modificadas. Pero además, en los últimos años ha habido un gran impulso al uso de plantas para el análisis de la regulación genética del desarrollo, debido en parte al potencial económico que tiene la manipulación genética de algunos cultivos como el maíz, el jitomate, el algodón, *etc.* El conocimiento de las bases moleculares del desarrollo ha permitido establecer que el control global de la morfogénesis es muy diferente entre las plantas y los animales. A pesar de que las plantas y los animales comparten un sistema molecular básico, tienen mecanismos distintos en la transmisión y el procesamiento de las señales, así como los mecanismos de comunicación celular (Meyerowitz, 1997; Trewavas y Malhó, 1997).

Los estudios moleculares han puesto de manifiesto la dificultad de describir el desarrollo de los organismos como función de la actividad genética. Por una parte, se conoce que un gen puede ser regulado por un gran número de factores de transcripción (Arnone y Davidson, 1997); y por otra parte, un mismo gen puede intervenir en diversos procesos durante el desarrollo. Por tanto, la magnitud de una modificación genotípica no necesariamente corresponde a la magnitud del cambio fenotípico asociado (Moreno, 1994). Una de las razones de la no-linearidad en la relación entre el genotipo y el fenotipo es la existencia de múltiples inter-relaciones entre los genes y sus productos. Dichas regulaciones cruzadas entre diversos grupos de genes forman las llamadas redes de regulación genética. El modelado de las redes de regulación ha mostrado que entre sus propiedades se incluyen a la estabilidad, la redundancia, la homeostasis y la multiestacionariedad[1], entre otras. Dichas propiedades no son intuitivas o evidentes a partir de una inspección visual de las interacciones presentes en una red de regulación. Por tanto, es necesario realizar un análisis de las propiedades dinámicas de tales redes, para así poder predecir el efecto que puede tener la activación o inactivación de un gen en particular sobre los patrones de actividad de la red completa. Ahora bien, si la red de regulación incluye a genes involucrados en el desarrollo, la modificación de los patrones de expresión podría modificar a la morfogénesis del organismo en estudio.

[1] Es decir, la capacidad de un sistema de poseer más de un estado estacionario.

Durante la presente década ha habido un gran esfuerzo por comprender las bases genéticas del desarrollo floral en *Arabidopsis thaliana*, pues ha demostrado ser un sistema biológico experimentalmente adecuado para el estudio de la morfogénesis. Por ello, en la actualidad se cuenta con una gran cantidad de información molecular y morfológica en dicho sistema. De particular relevancia para el tema de esta tesis, es el hecho de que la cantidad de información generada ha permitido la elaboración de algunos esquemas sencillos sobre las bases moleculares del desarrollo floral (Coen, 1991; Coen y Meyerowitz, 1991; Weigel, 1995; Weigel y Meyerowitz, 1993a). En contraste con la abundancia de información molecular, existen aún muchas incógnitas con respecto a la naturaleza de las vías de señalización y las interacciones celulares involucradas en la morfogénesis floral.

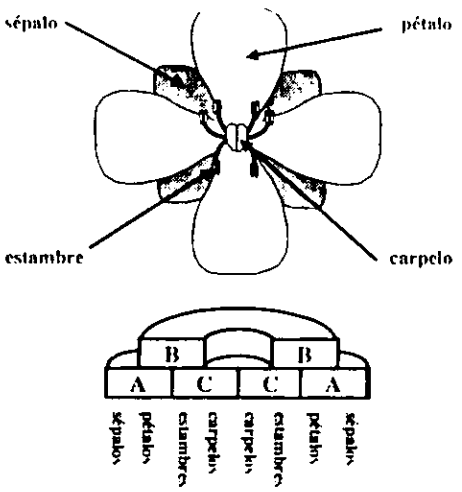
Los estudios sobre el control molecular del desarrollo y la morfogénesis no se han restringido al estudio de la flor. El desarrollo de las hojas y de la raíz también ha sido objeto de estudio, principalmente en aquellos aspectos relacionados al papel de las interacciones celulares en la morfogénesis (Scheres, 1996). Por ejemplo, los resultados experimentales han mostrado que el control de la diferenciación celular en la epidermis y la corteza es muy similar en las hojas, el tallo y la raíz. En esos tejidos, el destino celular es determinado por la posición relativa que guardan las células entre sí, y por lo menos cuatro genes son comunes en el proceso de diferenciación (Benfey, 1999). Sin embargo, la cantidad de información molecular disponible es claramente mayor en los procesos involucrados en el desarrollo floral.

Arabidopsis thaliana es un sistema experimental para el cual se tiene una considerable cantidad de información experimental a distintos niveles y en diversos aspectos del desarrollo. Sin embargo, hasta ahora han sido pocos los esfuerzos por sintetizar dicha información en modelos matemáticos o computacionales. Previo a la realización del proyecto de investigación para esta tesis, solamente se había publicado un modelo combinatorio para describir los patrones globales de expresión genética durante el desarrollo de los órganos flores de *Arabidopsis*. A principios de la década de 1990 se desarrolló el modelo combinatorio ABC (ver adelante), el cual describe la relación existente entre tres regiones de expresión genética con la aparición de los cuatro tipos de órganos florales. Este modelo, sin embargo, no propone un mecanismo a nivel molecular para explicar el establecimiento de tal patrón de expresión.

En esta tesis se presentan modelos dinámicos de regulación genética en dos procesos del desarrollo de *Arabidopsis*: el desarrollo de la flor y la aparición de pelos en la epidermis de la raíz. Los modelos propuestos aquí constituyen las primeras explicaciones en términos de redes de regulación sobre el establecimiento de patrones de expresión genética observados experimentalmente. En el caso particular del modelo de la raíz, además de la regulación genética se combinan algunos aspectos de la transducción de señales y las interacciones celulares. Cabe señalar que, si bien los modelos presentados son muy sencillos, pues utilizan

principalmente variables binarias, ellos tienen un alto valor predictivo.

FIGURA 1
La flor de *Arabidopsis* y el modelo ABC



2 *Arabidopsis thaliana*

Arabidopsis thaliana es una dicotiledonea que pertenece a la familia Brassicaceae, la cual incluye cultivos importantes como la mostaza y la coliflor. Su ciclo de vida completo, incluyendo la germinación de la semilla, la formación de la roseta, el desarrollo del tallo principal, la floración y la maduración de las primeras semillas es de aproximadamente seis semanas. *Arabidopsis* presenta un crecimiento indeterminado[1], con un tallo principal del cual se originan diversas estructuras durante el desarrollo. Después de la germinación y la aparición de los cotiledones, comienza la fase de crecimiento vegetativo. En dicha fase aparecen un par de hojas en cada nodo con una filotaxia decusada[2] y un crecimiento internodal corto, formándose de este modo una roseta. Aproximadamente a las tres semanas de germinación, comienza un crecimiento rápido del tallo, formándose la inflorescencia primaria. En la inflorescencia aparecen hojas con una filotaxia en espiral, y en las axilas de las hojas se originan tallos secundarios. Finalmente, aparecen los primordios florales de manera espiral a lo largo de las inflorescencias. Antes de que las flores abran sus pétalos, éstas se auto-polinizan. Posteriormente, las flores ya maduras se abren, alcanzando su tamaño máximo de aproximadamente 2mm de largo. Finalmente, las flores se marchitan dando lugar a vainas que contienen a las semillas (Bowman, 1994).

Arabidopsis thaliana se utiliza como un organismo experimental para estudios de genética, fisiología y biología del desarrollo, debido a su tamaño pequeño, a su ciclo de vida corto y a su genoma relativamente pequeño (alrededor de 120Mb[3]). Otra de las ventajas es que las semillas de *Arabidopsis* pueden ser germinadas en una caja de petri, facilitando así la búsqueda de mutaciones que alteren la morfología de algún órgano en particular. Dentro de las mutaciones llamativas que ocurren en la planta se encuentran las llamadas homeóticas, las cuales resultan en patrones alterados del desarrollo en donde unos órganos son reemplazados por otros. El análisis del efecto combinado de mutaciones ha

[1] Esto es, que continua a lo largo de la vida del organismo.

[2] Hay 180° entre las hojas de cada nodo, y 90° entre las hojas de nodos adyacentes.

[3] Se estima que *A. thaliana* tiene 23,000 genes, y se espera contar con la secuencia del genoma completo para finales del año 2000 (ver http://genome-www.stanford.edu/Arabidopsis).

permitido el establecimiento parcial de la regulación entre diversos genes y sus productos.

2.1 La arquitectura floral de *Arabidopsis*

Las flores de *Arabidopsis* se componen de cuatro verticilos[1]. El primero (el más externo) está formado por cuatro sépalos, el segundo por cuatro pétalos, el tercero por seis estambres, y el cuarto por dos carpelos fusionados (ver Figura 1; Coen, 1991). Esta arquitectura floral puede ser alterada por la supresión, o sobre-expresión de algunos genes. El efecto morfológico de los distintos mutantes en la floración ha permitido agrupar a los genes del desarrollo floral de *Arabidopsis* en cuatro categorías: 1) genes de floración temprana y tardía (Coupland, 1995), 2) genes de identidad de meristemo (Schultz y Haughn, 1993; Weigel *et al.*, 1992), 3) genes cadastrales (Liu y Meyerowitz, 1995; Sakai *et al.*, 1995; Weigel y Meyerowitz, 1993a), y 4) genes de identidad de órgano (Weigel y Meyerowitz, 1993b). Las mutaciones que afectan a los genes de floración temprana y tardía provocan un cambio en la temporalidad de la aparición de flores, retrasándola o adelantándola, pero sin afectar a la morfología floral. Por su parte, las mutaciones en los genes de identidad del meristemo retrasan la aparición de los meristemos florales, y las flores resultantes presentan alteraciones morfológicas. Las mutaciones en los genes cadastrales alteran los patrones de expresión de los genes de identidad de órgano, modificando el tamaño o número de los órganos florales resultantes. Finalmente, aquellas mutaciones que afectan directamente a los genes de identidad de órgano modifican la fórmula floral.

Se han caracterizado molecularmente a muchos de los genes clasificados en los grupos arriba mencionados, por lo cual se conocen a grandes rasgos las regiones y la temporalidad de su expresión. El establecimiento de las etapas del desarrollo en las cuales se activan los diversos tipos de genes florales, ha permitido que diversos autores propongan algunas de las relaciones de regulación genética que existen entre los diversos grupos de genes florales. Específicamente, se conoce que diversas señales ambientales activan a los genes de floración tardía (Puterill *et al.*, 1995; Weigel y Meyerowitz, 1993a), los cuales regulan la actividad transcripcional de los genes de identidad de meristemo (Coupland, 1995; Weigel y Meyerowitz, 1993a; Ma, 1994). Estos últimos son los que regulan la actividad de los genes cadastrales (Clark *et al.*, 1995; Levin y Meyerowitz, 1995; Liu y Meyerowitz, 1995). Finalmente, tanto los genes de identidad de meristemo como los cadastrales regulan la expresión de los genes de identidad de órgano (Okamuro *et al.*, 1993). No obstante que se conocen las relaciones de regulación transcripcional entre los grupos de genes arriba mencionados, ha sido muy difícil establecer la regulación de ellos en un nivel gen-a-gen. Por ejemplo, existe un artículo en el cual se trataron de reunir las relaciones de regulación conocidas para diversos genes de los grupos ya mencionados (Theißen y Saedler, 1995); sin embargo en dicho trabajo no se establece una distinción clara entre la regulación transcripcional, la regulación

[1] Anillos concéntricos formados cada uno de ellos por los distintos órganos florales.

post-transcripcional y los efectos fenotípicos.

En el caso particular del desarrollo de las flores de *Arabidopsis*, ya se han descrito diversas mutaciones homeóticas que alteran su arquitectura floral. A partir del efecto fenotípico de tales mutaciones se desarrolló un modelo combinatorio llamado ABC (Coen y Meyerowitz, 1991; Meyerowitz, 1994a). El modelo ABC propone que la identidad de los órganos que se encuentran en cada uno de los verticilos florales está determinada por la combinación de tres "actividades" o "funciones" genéticas (A, B y C; Figura 1). En el modelo, los sépalos están determinados por la presencia de la actividad A, los pétalos por una combinación de la actividad A con la B, los estambres por la combinación de las actividades B y C, y los carpelos por la actividad C únicamente. Por último, el modelo ABC indica que las actividades A y C son mutuamente antagonistas.

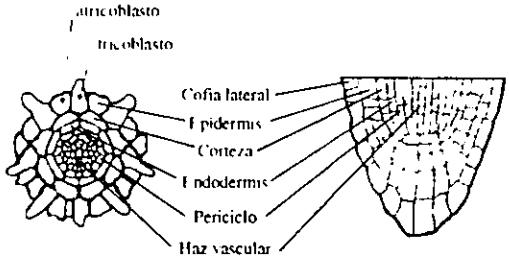
Se han asignado actividades transcripcionales de genes específicos a cada una de las tres funciones arriba mencionadas. Los genes *APETALA1* (*AP1*) y *APETALA2* (*AP2*) son los responsables de la actividad A, los genes *APETALA3* (*AP3*) y *PISTILLATA* (*PI*) son los responsables de la actividad B, en tanto que *AGAMOUS* (*AG*) es el único gen al que se le ha asignado la actividad C (Meyerowitz 1994a). Con la identificación de los genes anteriores a cada una de las actividades mencionadas, el modelo ABC puede utilizarse no sólo para describir a las flores de las plantas silvestres, sino también a las flores de las plantas a las cuales se les ha alterado de algún modo la expresión de los genes mencionados. Esto es, si se elimina la actividad A al mutar el gen *AP1*, la expresión del gen *AG* se expande hasta cubrir la región en donde se expresaba normalmente la actividad A. Como resultado, las flores en desarrollo presentan un patrón de activación tal que desarrollarán (de afuera hacia adentro) carpelos, estambres, estambres y carpelos (ver Figura 1 del Apéndice 3). Por otro lado, al presentarse una mutación en algún gen de la actividad B, las flores se desarrollan presentando sépalos, sépalos, carpelos y carpelos. Finalmente, al mutarse el gen de actividad C, la actividad A toma su sitio produciendo flores conteniendo sépalos y pétalos de manera reiterada. Este grupo de mutaciones homeóticas es tan bien descrito por el modelo ABC, que aunque algunos estudios sugieren pequeñas modificaciones al modelo (Ma, 1994; Ray *et al.*, 1994), éste sigue siendo ampliamente utilizado para describir la morfología floral de plantas mutantes (Coen y Meyerowitz, 1991; Meyerowitz, 1994a; Ma, 1994; Meyerowitz, 1994b) y plantas transgénicas (Krizek y Meyerowitz, 1996).

Existen dos tipos de mutaciones homeóticas florales descritas que no captura el modelo ABC. Una es la causada por la mutación en el gen *SUPERMAN* (plantas *sup*[1]), que produce un aumento en el número de estambres con respecto al tipo silvestre (Sakai *et al.*, 1995), además de una reducción en el número y tamaño de los carpelos. La otra transformación homeótica se

[1] En la literatura de *Arabidopsis*, el nombre de los genes se escribe con letras itálicas mayúsculas (por ejemplo *SUP*). Para indicar que un gene está mutado, se escribe su nombre con letras itálicas minúsculas (*sup*). Finalmente, a las proteínas de dichos genes se les indica con letra mayúsculas sin italizar (SUP).

presenta cuando se muta el gen *BELL1* (*BEL1*). Estas plantas presentan carpelos en vez de óvulos (Ray *et al.*, 1994), de tal manera que se tienen carpelos dentro de carpelos, etc. Finalmente, existen algunos otros cambios homeóticos en las flores de *Arabidopsis* que pueden explicarse como efectos indirectos sobre la actividad transcripcional de los genes de identidad de órgano (los incluidos en el modelo ABC). Por ejemplo, las plantas con la mutación *leafy* (*lfy*) presentan un retraso temporal en la aparición de los meristemos florales. Además, las flores de estas plantas presentan casi siempre una fórmula de sépalos, sépalos, carpelos y carpelos (Weigel *et al.*, 1992), lo cual se atribuye a una disminución de la actividad de los genes *AP3* y *PI*, responsables de la función B (Jack *et al.*, 1994; Parcy *et al.*, 1998).

FIGURA 2
La raíz principal



2.2 La raíz principal de *Arabidopsis*

La raíz de *Arabidopsis thaliana* es un sistema experimental que permite hacer estudios sobre la integración de aspectos celulares y moleculares durante el desarrollo y la morfogénesis. La raíz tiene una arquitectura regular, un tamaño pequeño y además es transparente. Estas características han permitido la combinación de estudios genéticos, moleculares, de muerte celular selectiva por láser, y de microinyección de colorantes para estudiar los mecanismos de señalización involucrados en la especificación celular (Scheres, 1996; 1997). Otra característica importante es que las células vegetales no migran. Esta inmovilidad es la razón por la cual las señales que puedan indicar la posición celular son determinantes para los procesos de diferenciación celular y la formación de patrones en las plantas.

La arquitectura radial de la raíz primaria (o principal, Figura 2) de *Arabidopsis* se establece durante la embriogénesis (Golberg *et al.*; 1994; Laux y Jürgens, 1997). En esta sección, sin embargo, se describe la arquitectura de la raíz madura. Para una descripción más detallada, así como el origen embrionario de los tejidos que se mencionan véase el Apéndice 1. La raíz madura está formada por tejidos concéntricas, que de afuera hacia adentro son: la cofia lateral, la epidermis, la corteza, la endodermis, el periciclo, y el haz vascular. Estos tejidos están formados por columnas (o filas) de células. Cada fila es mantenida por la elongación y la subsecuente división anticlinal[1] de células primordiales, conocidas como las células iniciales. Existen cuatro tipos de células iniciales: de epidermis y cofia lateral, de corteza y endodermis, de tejido vascular y por último de cofia (también llamada cofia de columela, o columela). Las células iniciales crecen y se dividen. Como resultado de la división se obtienen dos células, pero sólo una de ellas seguirá siendo célula inicial; a la otra célula se le llama célula hija. Las células hijas son

[1] Perpendicular al eje longitudinal de la planta.

capaces de dividirse y diferenciarse en alguno de los tipos celulares ya mencionados que conforman a la raíz madura.

El crecimiento de la raíz es el resultado de la combinación de la expansión y la división celulares que se da en el meristemo (Beemster y Baskin, 1998). El meristemo de la raíz contiene a las células iniciales ya mencionadas, y su continua división no sólo origina el crecimiento de la raíz, sino que mantiene la arquitectura originada en la embriogénesis. Una diferencia importante de la función del meristemo de raíz con la del meristemo apical, es que éste último produce diversos tipos de estructuras (meristemos secundarios, hojas, primordios florales) durante las diferentes etapas del desarrollo. Sin embargo, a pesar del crecimiento continuo, se pueden reconocer tres regiones parcialmente sobrepuestas a lo largo de la raíz: el meristemo ubicado en la punta de la misma, la zona de elongación inmediatamente arriba del meristemo, y por último la zona de diferenciación. La zona del meristemo ocupa cerca de 250 µm de la raíz y está cubierta por la cofia y la cofia lateral. De estos tejidos se van desprendiendo las células más viejas conforme crece la raíz. Por su parte, la zona de elongación ocupa una longitud aproximada de 250 µm. Finalmente, la zona de diferenciación es la región más apical, y en donde los tejidos se diferencian a su estado maduro (Dolan *et al.*, 1993; 1994).

El meristemo de la raíz está formado no sólo por las células iniciales ya mencionadas; además tiene un centro mitóticamente inactivo llamado centro quiescente. Las células iniciales rodean al centro quiescente, haciendo contacto con él directamente en la parte superior, lateral e inferior. Las células iniciales de la cofia forman una capa de cerca de veinte células que se dividen periclinalmente[1] produciendo la cofia madura. En una región más interior, pero en contacto con el centro quiescente, se encuentran las células iniciales de la epidermis, cuyas divisiones originan tanto a la epidermis como a la cofia lateral maduras. Por su parte, las células iniciales de corteza dan lugar a la formación de la corteza y la endodermis. El patrón de división secuencial de todas las células iniciales origina la formación de espirales a lo largo de la raíz (Baum y Rost, 1996; Rost *et al.*, 1996).

La epidermis de la raíz no es un tejido homogéneo, ya que está formada por dos tipos celulares: células con una proyección llamada pelo radicular (los tricoblastos) y células sin pelo (los atricoblastos). Si bien la presencia del pelo radicular es la diferencia más notoria entre las células de la epidermis, no es la única. Morfológicamente, los tricoblastos son más cortos y tienen un citoplasma más denso que el de los atricoblastos (Dolan *et al.*, 1994). Notablemente, la localización relativa de estos dos tipos celulares en la epidermis es altamente predecible y depende del contacto establecido con las células de corteza. La epidermis está formada en promedio por 19 filas de células, en tanto que la corteza casi siempre tiene ocho filas. Esta diferencia numérica entre las células epidermales y corticales origina que algunas células de la epidermis hagan contacto con dos células de corteza, y otras sólo con una. Las células de la epidermis que hacen

[1] Paralelo a la superficie de la raíz.

contacto con dos células de la corteza se diferencian en tricoblastos. En cambio las células epidermales en contacto con una sola célula de corteza se diferencian en atricoblastos. La razón del número de filas de tricoblastos a atricoblastos se mantiene constante durante el crecimiento de la raíz, y tiene un valor entre 0.5 a 1, variando de planta a planta.

En el meristemo de la raíz se producen las nuevas células para el crecimiento de la misma. Dicho crecimiento se da mediante un proceso que requiere de un adecuado balance entre la proliferación y la diferenciación de los distintos tipos celulares, manteniendo el patrón radial originado durante la embriogénesis. Sin embargo, a pesar del patrón fijo de la división y la diferenciación celulares ya mencionados, las células meristemáticas de la raíz tienen la capacidad de alterar su patrón de diferenciación dependiendo de las células con las que estén en contacto. Se han realizado experimentos de muerte celular dirigida con láser (Van der Berg *et al.*, 1995; 1997; 1998), los cuales indican que existe una información posicional que actúa continuamente para mantener la identidad celular. Sin embargo, no se conoce la naturaleza molecular de los procesos de señalización para el mantenimiento del destino celular.

Por último, en contraste con la raíz primaria, las raíces secundarias se originan de meristemos formados post-embrionalmente a partir de células fundadoras en el pericielo (Dolan *et al.*, 1993; Sussex *et al.*, 1995). La mayor parte de las raíces laterales emergen de la raíz primaria, y algunas pocas se originan de otras raíces laterales, pero a pesar de su origen distinto, las raíces laterales tienen una arquitectura idéntica a la de la raíz primaria. Esta similitud sugiere que el mismo grupo de genes está involucrado tanto en el desarrollo de las raíces primarias como de las laterales (Laskowski *et al.*, 1995; Malamy y Benfey, 1997a; 1997b).

3 REDES DE REGULACION GENETICA

Las técnicas modernas de la biología molecular han originado un aumento en la velocidad a la cual se descubren genes y se les determina su secuencia primaria. Sin embargo, es necesario hacer uso de estudios genéticos, bioquímicos y fisiológicos para identificar las funciones de dichos genes. Una primera consecuencia de la diferencia metodológica entre la identificación de un gen y el conocimiento de su función, es el hecho de que la velocidad a la cual se describen las vías de regulación es mucho más lenta que la del descubrimiento de nuevos genes y sus productos. Para el tema de esta tesis, son de particular interés aquellos casos en los que se describen genes que codifican a factores de transcripción. Tales factores son proteínas que se unen a las secuencias de regulación de otros genes, controlando sus niveles de transcripción. Pero además, si esos factores regulan a genes que a su vez codifican para otros factores de transcripción, entonces se crea una interdependencia entre grupos de genes formando una red de regulación genética (Kauffman, 1991; 1993; Zuckerkandl, 1994). Como resultado del establecimiento de estas redes, es posible observar la expresión coordinada de diversos grupos de genes. Si bien estas ideas son comúnmente aceptadas por los biólogos, normalmente no se conocen las

contacto con dos células de la corteza se diferencian en tricoblastos. En cambio las células epidermales en contacto con una sola célula de corteza se diferencian en atricoblastos. La razón del número de filas de tricoblastos a atricoblastos se mantiene constante durante el crecimiento de la raíz, y tiene un valor entre 0.5 a 1, variando de planta a planta.

En el meristemo de la raíz se producen las nuevas células para el crecimiento de la misma. Dicho crecimiento se da mediante un proceso que requiere de un adecuado balance entre la proliferación y la diferenciación de los distintos tipos celulares, manteniendo el patrón radial originado durante la embriogénesis. Sin embargo, a pesar del patrón fijo de la división y la diferenciación celulares ya mencionados, las células meristemáticas de la raíz tienen la capacidad de alterar su patrón de diferenciación dependiendo de las células con las que estén en contacto. Se han realizado experimentos de muerte celular dirigida con láser (Van der Berg *et al.*, 1995; 1997; 1998), los cuales indican que existe una información posicional que actúa continuamente para mantener la identidad celular. Sin embargo, no se conoce la naturaleza molecular de los procesos de señalización para el mantenimiento del destino celular.

Por último, en contraste con la raíz primaria, las raíces secundarias se originan de meristemos formados post-embrionalmente a partir de células fundadoras en el pericicio (Dolan *et al.*, 1993; Sussex *et al.*, 1995). La mayor parte de las raíces laterales emergen de la raíz primaria, y algunas pocas se originan de otras raíces laterales, pero a pesar de su origen distinto, las raíces laterales tienen una arquitectura idéntica a la de la raíz primaria. Esta similitud sugiere que el mismo grupo de genes está involucrado tanto en el desarrollo de las raíces primarias como de las laterales (Laskowski *et al.*, 1995; Malamy y Benfey, 1997a; 1997b).

3 REDES DE REGULACION GENETICA

Las técnicas modernas de la biología molecular han originado un aumento en la velocidad a la cual se descubren genes y se les determina su secuencia primaria. Sin embargo, es necesario hacer uso de estudios genéticos, bioquímicos y fisiológicos para identificar las funciones de dichos genes. Una primera consecuencia de la diferencia metodológica entre la identificación de un gen y el conocimiento de su función, es el hecho de que la velocidad a la cual se describen las vías de regulación es mucho más lenta que la del descubrimiento de nuevos genes y sus productos. Para el tema de esta tesis, son de particular interés aquellos casos en los que se describen genes que codifican a factores de transcripción. Tales factores son proteínas que se unen a las secuencias de regulación de otros genes, controlando sus niveles de transcripción. Pero además, si esos factores regulan a genes que a su vez codifican para otros factores de transcripción, entonces se crea una interdependencia entre grupos de genes formando una red de regulación genética (Kauffman, 1991; 1993; Zuckerkandl, 1994). Como resultado del establecimiento de estas redes, es posible observar la expresión coordinada de diversos grupos de genes. Si bien estas ideas son comúnmente aceptadas por los biólogos, normalmente no se conocen las

propiedades dinámicas globales de dichas redes de regulación.

Es una práctica común describir una relación de regulación transcripcional utilizando flechas (A→B) para indicar activaciones, y flechas despuntadas (A⊣B) para indicar represiones. Dicha representación gráfica es muy útil para sintetizar y comprender a la maquinaria global de transcripción de un organismo. Sin embargo, el conocimiento de la conectividad en una red de regulación no es suficiente para determinar el comportamiento dinámico de dicha red. Existen algunas propiedades que no pueden ser deducidas únicamente a partir de la inspección visual de las interacciones en la red. Por ejemplo, no es posible conocer cuantos estados estacionarios de activación puede tener una red determinada, ni si esos estados son estables o no. Para conocer éstas y otras propiedades, es necesario incorporar las tasas de transcripción de cada uno de los genes como función de los factores de transcripción unidos a las secuencias reguladoras de los mismos. Al incorporar dicha información, una red de regulación genética puede ser descrita en términos de los sistemas dinámicos. La tendencia para modelar una red como un sistema dinámico es la de utilizar un sistema de ecuaciones diferenciales acopladas, en donde cada variable dependiente representa la tasa transcripcional de cada gen modelado. La mayor parte de las veces, sin embargo, existe muy poca información cuantitativa experimental para poder fijar el valor de los parámetros en el sistema de ecuaciones. De hecho, una gran mayoría de los resultados experimentales publicados sólo contiene información cualitativa sobre los patrones de expresión espacial y temporal de los genes en estudio.

Ahora bien, una característica importante de las redes de regulación, a diferencia de una cascada o una jerarquía, es que presentan la posibilidad de establecer interacciones de retroalimentación, directa o indirecta. Es decir, el producto de cierto gen puede regular directa o indirectamente la transcripción de él mismo. La presencia de relaciones de retroalimentación permite al sistema alcanzar no sólo patrones fijos de actividad, sino también patrones periódicos de activación. Si bien las propiedades dinámicas de las redes de regulaciones dependen en gran medida del tipo de ecuaciones utilizadas, incluso las redes aparentemente sencillas tienen un comportamiento dinámico global complejo, que incluye la auto-organización, la estabilidad, la redundancia, la periodicidad, la multiestacionariedad y la homeostasis (Somogyi y Sniegoski, 1996; Thieffry *et al.*, 1995; Thomas *et al.*, 1995). El Apéndice 2 es un artículo de revisión que presenta con más profundidad las propiedades dinámicas de las redes de regulación modeladas como sistemas discretos. En él se resumen algunas de las aplicaciones que han tenido dichos sistemas dinámicos.

Existen diversos trabajos teóricos en los cuales se exploran las características de las redes de regulación hipotéticas (Clarke *et al.*, 1993; Kauffman, 1969; 1993; Wagner 1996). A pesar de la abundancia de trabajos teóricos sobre las redes de regulación, la laboriosa obtención de resultados experimentales hace difícil poner a prueba la utilidad de los formalismos de las redes de regulación como sistemas

dinámicos. Hasta ahora se han publicado modelos de la regulación genética principalmente para *Drosophila* (Bodnar, 1997; Burstein, 1995; Reinitz y Sharp, 1995), aunque cada vez es más abundante la información para otros sistemas, como por ejemplo el erizo de mar (Arnone y Davidson, 1997; Davidson *et al.*, 1998; Yuh *et al.*, 1998). El mayor esfuerzo en los modelos de redes de regulación se ha centrado en la mosca debido a la gran cantidad de datos que se tienen sobre la expresión de los grupos de los genes maternos, los gap, los pair-rule, los de polaridad de segmento y los homeóticos. Sin embargo, cabe señalar que en *Drosophila* normalmente se modela la expresión genética durante el blastodermo, en la cual el embrión es un sincisio, en donde hay muchos núcleos pero no hay células como tales (Hunding y Engelhardt, 1995; Mjolsness *et al.*, 1991; Reinitz y Sharp, 1995; Spirov, 1996). En esta tesis se presentan dos modelos de regulación genética: el primero para el desarrollo floral (Apéndices 3 y 4), y el segundo para la diferenciación celular de la epidermis de la raíz en *Arabidopsis thaliana*.

4 ANALISIS LOGICO DE CIRCUITOS DE RETROALIMENTACION

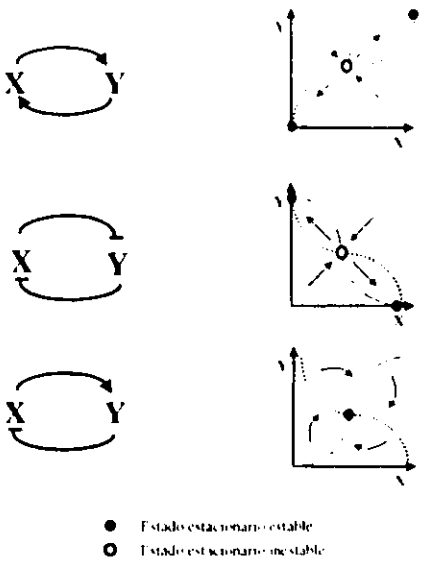
Dentro de las redes de regulación es posible encontrar la presencia de circuitos de retroalimentación, cuya importancia ha sido ampliamente reconocida por los biólogos experimentales y teóricos desde hace por lo menos tres décadas (ver por ejemplo, Monod y Jacob, 1961; Rosen, 1968). Los circuitos de retroalimentación son cadenas circulares de interacciones, de tal manera que cada elemento dentro del circuito influencia su propio nivel de actividad. En muchas ocasiones es posible asignar sin ambigüedad un signo a cada una de las interacciones entre los diferentes elementos de un circuito. Cuando la influencia de un elemento sobre el siguiente es la de aumentar la activación, entonces se considera una interacción positiva. Inversamente, una interacción negativa es aquella que reduce la actividad del elemento en consideración. Cuando se pueden especificar sin ambigüedad los signos a cada una de las interacciones en un circuito en particular, dicho circuito también puede ser clasificado como positivo o negativo. El signo de un circuito se determina por el número de interacciones negativas que lo constituyen. Si dicho número es non entonces el circuito completo es negativo, de otra manera es considerado como positivo. Los circuitos positivos y negativos poseen diferentes propiedades dinámicas, y por lo tanto, confieren distintas capacidades a los sistemas biológicos que los contienen (Thomas *et al.*, 1995). Desde un punto de vista dinámico, los circuitos de retroalimentación positivos tienen como efecto la multiestacionariedad, en tanto que los circuitos de retroalimentación negativos generan oscilaciones sostenidas o amortiguadas (Figura 3). Biológicamente, los circuitos positivos son necesarios para lograr la diferenciación, entendida en el presente contexto como patrones alternativos de expresión genética. Por su parte, los circuitos negativos son necesarios para generar homeostasis. Las anteriores propiedades han sido formalmente demostradas en un contexto general (Gouzé, 1998; Plahte *et al.*, 1995; Snoussi, 1998).

dinámicos. Hasta ahora se han publicado modelos de la regulación genética principalmente para *Drosophila* (Bodnar, 1997; Burstein, 1995; Reinitz y Sharp, 1995), aunque cada vez es más abundante la información para otros sistemas, como por ejemplo el erizo de mar (Arnone y Davidson, 1997; Davidson *et al.*, 1998; Yuh *et al.*, 1998). El mayor esfuerzo en los modelos de redes de regulación se ha centrado en la mosca debido a la gran cantidad de datos que se tienen sobre la expresión de los grupos de los genes maternos, los gap, los pair-rule, los de polaridad de segmento y los homeóticos. Sin embargo, cabe señalar que en *Drosophila* normalmente se modela la expresión genética durante el blastodermo, en la cual el embrión es un sincisio, en donde hay muchos núcleos pero no hay células como tales (Hunding y Engelhardt, 1995; Mjolsness *et al.*, 1991; Reinitz y Sharp, 1995; Spirov, 1996). En esta tesis se presentan dos modelos de regulación genética: el primero para el desarrollo floral (Apéndices 3 y 4), y el segundo para la diferenciación celular de la epidermis de la raíz en *Arabidopsis thaliana*.

4 ANALISIS LOGICO DE CIRCUITOS DE RETROALIMENTACION

Dentro de las redes de regulación es posible encontrar la presencia de circuitos de retroalimentación, cuya importancia ha sido ampliamente reconocida por los biólogos experimentales y teóricos desde hace por lo menos tres décadas (ver por ejemplo, Monod y Jacob, 1961; Rosen, 1968). Los circuitos de retroalimentación son cadenas circulares de interacciones, de tal manera que cada elemento dentro del circuito influencia su propio nivel de actividad. En muchas ocasiones es posible asignar sin ambigüedad un signo a cada una de las interacciones entre los diferentes elementos de un circuito. Cuando la influencia de un elemento sobre el siguiente es la de aumentar la activación, entonces se considera una interacción positiva. Inversamente, una interacción negativa es aquella que reduce la actividad del elemento en consideración. Cuando se pueden especificar sin ambigüedad los signos a cada una de las interacciones en un circuito en particular, dicho circuito también puede ser clasificado como positivo o negativo. El signo de un circuito se determina por el número de interacciones negativas que lo constituyen. Si dicho número es non entonces el circuito completo es negativo, de otra manera es considerado como positivo. Los circuitos positivos y negativos poseen diferentes propiedades dinámicas, y por lo tanto, confieren distintas capacidades a los sistemas biológicos que los contienen (Thomas *et al.*, 1995). Desde un punto de vista dinámico, los circuitos de retroalimentación positivos tienen como efecto la multiestacionariedad, en tanto que los circuitos de retroalimentación negativos generan oscilaciones sostenidas o amortiguadas (Figura 3). Biológicamente, los circuitos positivos son necesarios para lograr la diferenciación, entendida en el presente contexto como patrones alternativos de expresión genética. Por su parte, los circuitos negativos son necesarios para generar homeostasis. Las anteriores propiedades han sido formalmente demostradas en un contexto general (Gouzé, 1998; Plahte *et al.*, 1995; Snoussi, 1998).

FIGURA 3
Dinámicas de los circuitos



El efecto dinámico de un circuito no dependerá solamente de su existencia, sino también de los pesos relativos de cada una de las interacciones que lo conforman. Por tanto, se considera que existen circuitos funcionales y no funcionales. Un circuito positivo es funcional si de hecho genera multiestacionaridad, en tanto que un circuito negativo es funcional si de hecho genera homeostasis. Cuando un circuito es funcional, existe un estado estacionario llamado característico (Snoussi y Thomas, 1993), el cual se encuentra en los valores umbrales de las interacciones del circuito. En un circuito positivo el estado característico siempre es inestable, siendo típicamente un punto silla en dos dimensiones que se encuentra sobre una separatriz[1]. En cambio, el estado característico de un circuito negativo puede ser tanto estable como inestable, formando típicamente un foco en dos dimensiones. Una característica del análisis lógico, es que cuando se conocen las interacciones de una red de regulación, pero no los pesos relativos de las interacciones, es posible calcular las restricciones paramétricas que hacen funcional a cualquier circuito, o a un grupo de circuitos al mismo tiempo. Por supuesto, es también posible encontrar que algunos circuitos no pueden ser funcionales simultáneamente con otros. Por tanto, el análisis lógico de circuitos reduce el estudio de toda una red de regulación al problema más sencillo de estudiar únicamente los circuitos de retroalimentación, y su papel en la dinámica global de la red.

Usando el análisis lógico de los circuitos de retroalimentación es posible disociar una red compleja en un conjunto bien definido de circuitos de retroalimentación. También se puede analizar el papel dinámico de cada circuito individual, y finalmente se puede mantener un control completo de la forma en el que se interconectan los circuitos. Es importante hacer notar que el análisis lógico se puede utilizar sin importar que la actividad genética sea representada como una variable continua (una sigmoide) o una variable discreta (funciones escalón). De hecho, se ha mostrado que si primero se estudia un modelo con variables discretas, se reduce enormemente la dificultad de un análisis exhaustivo de ese mismo modelo pero que utilice variables continuas (ver por ejemplo, Muraille *et al.*, 1996). Brevemente, en un análisis de circuitos se asigna una variable lógica que representa el nivel de activación de cada gen. El conjunto de valores que pueden tomar esas variables es $\{0, s^{(1)}, 1, s^{(2)}, 2, \ldots s^{(n)}, n\}$, en donde $s^{(n)}$ representa el valor

[1] El conjunto de los estados estacionarios que separan a distintas cuencas de atracción.

umbral que separa a los estados *i-1* e *i*. Con el uso de estas funciones en escalón es posible conocer si un elemento de la red se encuentra debajo, en o arriba del umbral de activación. Adicionalmente, se introducen parámetros lógicos (marcados como 'K' e índices apropiados) para calificar el peso de cada interacción, o combinación de interacciones, en la expresión de un gen regulado (Snoussi, 1989; ver Apéndice 4).

En el contexto del formalismo lógico, una red genética de *n* elementos puede ser descrita por dos matrices, la primera siendo una matriz (*nxn*) conteniendo los signos (y eventualmente los umbrales) de todas las interacciones, y la segunda matriz ($nx2^n$) conteniendo los valores de los parámetros lógicos. En este formalismo, el estado del sistema está definido por un vector de dimensión *n*. Cada vez que el vector y su imagen (*i.e.*, el vector formado por los valores de las funciones correspondientes) son iguales, hay un estado estacionario del sistema. Debido a que se toman en cuenta explícitamente los valores umbrales, se pueden distinguir entre estados lógicos *regulares* y *singulares*, donde los últimos incluyen uno o más valores umbrales. Debido al hecho de que el número de estados lógicos es bien definido y finito, se pueden derivar todos los posibles estados, calcular el valor de las funciones correspondientes, y así identificar todos los estados estacionarios del sistema.

5 LOS MODELOS DE *Arabidopsis*

Hasta ahora no se conoce completamente la naturaleza del mapa del genotipo al fenotipo de ningún organismo. Sin embargo, poco a poco se presentan modelos para describir procesos de diferenciación para algunos subsistemas biológicos específicos. Afortunadamente, ahora se cuenta con una gran cantidad de información que relaciona a la actividad genética con varios procesos de morfogénesis en *Arabidopsis thaliana*. En las siguientes páginas se resumen dos modelos de regulación genética, uno sobre el control de la floración, y otro sobre el proceso de diferenciación en la epidermis de la raíz de *Arabidopsis*, los cuales se desarrollan en esta tesis.

5.1 El modelo NET sobre la regulación genética de la floración

En el Apéndice 3 se incluye el primer artículo de investigación que contiene los resultados obtenidos en mi proyecto de doctorado. El artículo presenta una red de regulación genética que incluye 11 genes involucrados en la regulación de la floración en *Arabidopsis*. La red incorpora a los siguientes genes: *EMBRYONIC FLOWER 1* (*EMF1*), *TERMINAL FLOWER 1* (*TFL1*), *LEAFY* (*LFY*), *APETALA1* (*AP1*), *CAULIFLOWER* (*CAL*), *LEUNIG* (*LUG*), *UNUSUAL FLORAL ORGANS* (*UFO*), *AGAMOUS* (*AG*), *APETALA3* (*AP3*), *PISTILLATA* (*PI*) y *SUPERMAN* (*SUP*); ver Figura 4. En el artículo del Apéndice 3 se describen los efectos fenotípicos de las mutaciones en los genes mencionados. Las interacciones de regulación de la transcripción de los genes mencionados fueron inferidas a partir de sus patrones de expresión reportados en mutantes con perdida de función, algunos experimentos de sobre-expresión, e incluso información morfológica de las flores en plantas mutantes. Una vez establecida la red de regulación, se construyó un sistema dinámico capaz de

umbral que separa a los estados *i-1* e *i*. Con el uso de estas funciones en escalón es posible conocer si un elemento de la red se encuentra debajo, en o arriba del umbral de activación. Adicionalmente, se introducen parámetros lógicos (marcados como 'K' e índices apropiados) para calificar el peso de cada interacción, o combinación de interacciones, en la expresión de un gen regulado (Snoussi, 1989; ver Apéndice 4).

En el contexto del formalismo lógico, una red genética de *n* elementos puede ser descrita por dos matrices, la primera siendo una matriz (*nxn*) conteniendo los signos (y eventualmente los umbrales) de todas las interacciones, y la segunda matriz ($nx2^n$) conteniendo los valores de los parámetros lógicos. En este formalismo, el estado del sistema está definido por un vector de dimensión *n*. Cada vez que el vector y su imagen (*i.e.*, el vector formado por los valores de las funciones correspondientes) son iguales, hay un estado estacionario del sistema. Debido a que se toman en cuenta explícitamente los valores umbrales, se pueden distinguir entre estados lógicos *regulares* y *singulares*, donde los últimos incluyen uno o más valores umbrales. Debido al hecho de que el número de estados lógicos es bien definido y finito, se pueden derivar todos los posibles estados, calcular el valor de las funciones correspondientes, y así identificar todos los estados estacionarios del sistema.

5 LOS MODELOS DE *Arabidopsis*

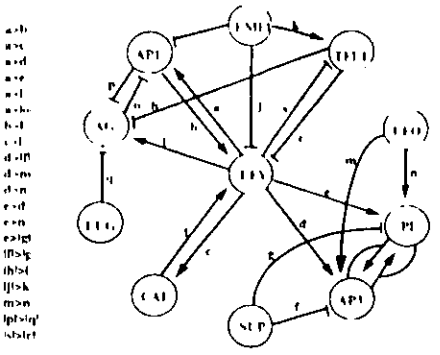
Hasta ahora no se conoce completamente la naturaleza del mapa del genotipo al fenotipo de ningún organismo. Sin embargo, poco a poco se presentan modelos para describir procesos de diferenciación para algunos subsistemas biológicos específicos. Afortunadamente, ahora se cuenta con una gran cantidad de información que relaciona a la actividad genética con varios procesos de morfogénesis en *Arabidopsis thaliana*. En las siguientes páginas se resumen dos modelos de regulación genética, uno sobre el control de la floración, y otro sobre el proceso de diferenciación en la epidermis de la raíz de *Arabidopsis*, los cuales se desarrollan en esta tesis.

5.1 El modelo NET sobre la regulación genética de la floración

En el Apéndice 3 se incluye el primer artículo de investigación que contiene los resultados obtenidos en mi proyecto de doctorado. El artículo presenta una red de regulación genética que incluye 11 genes involucrados en la regulación de la floración en *Arabidopsis*. La red incorpora a los siguientes genes: *EMBRYONIC FLOWER 1* (*EMF1*), *TERMINAL FLOWER 1* (*TFL1*), *LEAFY* (*LFY*), *APETALA1* (*AP1*), *CAULIFLOWER* (*CAL*), *LEUNIG* (*LUG*), *UNUSUAL FLORAL ORGANS* (*UFO*), *AGAMOUS* (*AG*), *APETALA3* (*AP3*), *PISTILLATA* (*PI*) y *SUPERMAN* (*SUP*); ver Figura 4. En el artículo del Apéndice 3 se describen los efectos fenotípicos de las mutaciones en los genes mencionados. Las interacciones de regulación de la transcripción de los genes mencionados fueron inferidas a partir de sus patrones de expresión reportados en mutantes con perdida de función, algunos experimentos de sobre-expresión, e incluso información morfológica de las flores en plantas mutantes. Una vez establecida la red de regulación, se construyó un sistema dinámico capaz de

describir los patrones de activación genéticos presentes en las flores maduras de *Arabidopsis*. El sistema dinámico está formado por ecuaciones en diferencias, y en él se utilizan variables binarias (con valores de 0 y 1) para describir la actividad de cada uno de los genes modelados. El sistema dinámico fue bautizado como el modelo NET.

FIGURA 4
El modelo NET



Un examen exhaustivo de las dinámicas del modelo NET mostró la existencia de seis atractores[1], siendo todos ellos puntos fijos. Esto es, partiendo de cualquier estado inicial de activación, tarde o temprano el sistema llega a alguno de seis estados estacionarios de activación. De los seis atractores obtenidos, cinco de ellos describen los estados de expresión observados *in vivo*; el sexto atractor es una predicción del modelo. El primer atractor consiste en la activación exclusiva de *AP1*, el gen de función A en el modelo ABC. En el segundo atractor, sólo se encuentran activos los genes *AP1*, *AP3* y *PI*, que corresponde a la presencia de las funciones A y B. En el tercer atractor se encuentran activos únicamente los genes *AG*, *AP3* y *PI*, es decir los genes de las funciones B y C. El cuarto atractor corresponde a la activación exclusiva de *AG*, el gen de la función C. Estos primeros cuatro puntos fijos de actividad tienen una correspondencia inmediata con las regiones de actividad propuestas por el modelo ABC. En el quinto atractor del modelo NET se presenta exclusivamente la expresión de los genes *EMF1* y *TFL1*, que son genes de inhibición floral y por tanto se encuentran activos en las células que no formarán parte de las flores. Dicho atractor, por lo tanto, representa el estado de no-floración en *Arabidopsis*. Finalmente, el sexto atractor del modelo NET está formado por la expresión estacionaria de los genes *EMF1*, *TFL1*, *AP3* y *PI*, que representa la activación simultanea de la actividad B y los genes de inhibición floral. Si bien este último atractor no se observa en condiciones normales en *Arabidopsis*, el modelo predice que éste puede ser inducido experimentalmente.

El modelo NET presenta diferencias notables con respecto a los anteriores esfuerzos de presentar la red de regulación que controla la floración. La primera y más importante diferencia es que el modelo NET es un sistema dinámico. Este es el primer modelo que estudia los patrones posibles de activación del grupo de genes ya descrito. Segundo, el modelo NET es la

[1] Un atractor es un estado estacionario, formado por una sucesión de estados de la red. En una red de estados discretos sólo hay dos tipos de atractores: los puntos fijos y los ciclos. Para mayor detalle ver el Apéndice 2.

primera integración sobre la regulación genética existente entre 11 de los principales genes que afectan la floración en *Arabidopsis*. Ya se mencionó que muchas de las interacciones del modelo habían sido propuestas con anterioridad, pero no se había presentado hasta ahora una síntesis de las relaciones de regulación transcripcional. El único caso similar anterior es el propuesto por Theißen y Saedler (1995). Sin embargo, dicha propuesta no distingue claramente entre efectos los fenotípicos y la regulación transcripcional, además de que no es un modelo dinámico. Tercero, el modelo NET presenta por primera vez la importancia, o pesos relativos, de muchas de las interacciones. La información cualitativa de dichos valores se infirió de la información experimental publicada. Sin embargo, para elaborar al sistema dinámico hubo que asignarles valores numéricos específicos a esos valores. Cuarto, los estados estacionarios de activación del sistema tienen un referente morfológico claro, debido a que entre los 11 genes incorporados se encuentran aquellos que conforman al modelo combinatorio ABC. Quinto, el modelo NET es la primera explicación dinámica de como es posible lograr los cuatro estados de activación presentes en el modelo ABC, además del estado de no-floración. Y sexto, el modelo predice por primera vez un estado estacionario estable de activación genética.

Finalmente, uno de los objetivos esenciales al elaborar modelos es el de poder hacer predicciones que sirvan de guía para la realización de experimentos futuros. Ya se mencionó previamente que muchas de las interacciones de regulación del modelo NET habían sido propuestas con anterioridad. Otras, sin embargo, fueron hechas por primera vez en este modelo, a saber: la activación de *EMF1* sobre *TFL1*, la inhibición de *EMF1* sobre *AP1*, la inhibición de *EMF1* sobre *LFY*, la inhibición de *TFL1* sobre *AG*, y la activación de *LFY* sobre *AG* (ver Figura 4). Notablemente, algunos autores habían propuesto que *LFY* inhibe a *AG* (Weigel y Meyerowitz, 1993b), pero nosotros utilizamos evidencias experimentales de diversos tipos para proponer que se trataba de una activación y no una inhibición. Resultados experimentales recientes por un grupo independiente a nosotros (Parcy *et al.*, 1998), confirmó que *LFY* activa a *AG* como nosotros lo predijimos por primera vez. Además, existe evidencia preliminar en favor de nuestra predicción acerca de la inhibición de *EMF1* sobre *AP1* (Aubert, 1998). Otra predicción del modelo que aún queda por confirmar es la concerniente a que la activación simultánea de *EMF1*, *TFL1*, *AP3* y *PI* es estable. Para terminar, cabe mencionar que los resultados obtenidos del modelo NET son particulares a las ecuaciones utilizadas. Esto es, en general, si se cambian los pesos de las interacciones, los niveles de activación, o el método de solución de las ecuaciones en diferencia, se obtendrán resultados distintos. Sin embargo, el análisis formal que se explica en la siguiente sección muestra que los estados estacionarios obtenidos en este modelo son robustos, de tal manera que los resultados obtenidos son propiedad de la conectividad del modelo, más que de los parámetros particulares utilizados.

5.2 Análisis de circuitos del modelo NET

Una de las críticas más comunes que presentan los biólogos experimentales ante un sistema dinámico

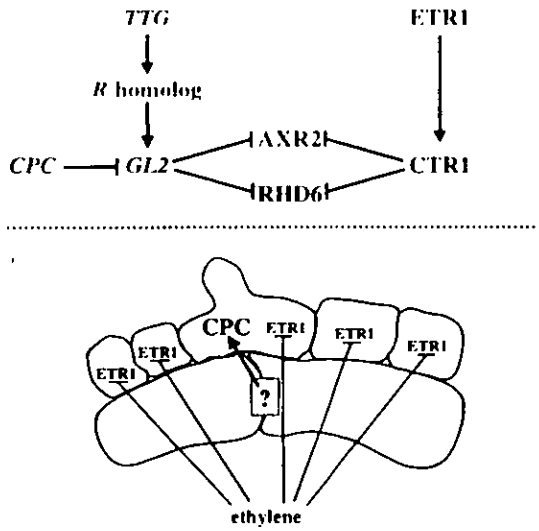
que describe algún proceso biológico, es que los parámetros utilizados en las ecuaciones se ajustan de manera *ad hoc* para obtener los resultados deseados. Sin embargo, existen diversos estudios sobre la estabilidad de los sistemas dinámicos para demostrar la robustez de los resultados obtenidos de los modelos. Para tal fin, escogimos el método de análisis de los circuitos de retroalimentación desarrollado por el grupo de Thomas (1991). Como se describió en una sección anterior, la ventaja de dicho análisis es que se pueden estudiar las implicaciones dinámicas de las interacciones presentadas en el modelo, mostrando que los resultados son altamente independientes del tipo de ecuaciones utilizadas para elaborar el sistema dinámico. El Apéndice 4 es un artículo en donde se hace un análisis formal del papel de los circuitos de retroalimentación en el modelo NET, presentado en la sección anterior.

Es importante recalcar que las metodologías utilizadas en los artículos incluidos como los Apéndices 3 y 4 son diferentes e independientes entre sí. Sin embargo, los resultados obtenidos en ambos estudios son comparables uno a uno. En el caso del análisis de circuitos, éste reveló que la red de regulación genética presenta seis estados de activación. Nuevamente, cuatro de esos estados corresponden a los patrones de activación presentes en los órganos florales, y establecidos por las actividades genéticas combinatorias del modelo ABC. Otro estado estable corresponde al de no-floración; y un último estado que no se encuentra en las plantas silvestres pero que es posible inducir por métodos experimentales. Además, se obtuvieron resultados novedosos que cuales se enumeran enseguida.

Primero, los circuitos establecidos entre los genes *AP1* y *AG* por una parte, y *AP3* y *PI* por otra, son suficientes para dar lugar a la actividad genética propuesta en el modelo ABC. Segundo, los resultados indican que, funcionalmente hablando, la red de regulación genética puede considerarse como compuesta por dos módulos, uno formado por los genes *TFL1*, *LFY*, *AP1* y *AG*, y el otro por los genes *AP3* y *PI*. Pero además, el análisis de circuitos indica que el gen *LFY* es el responsable de la comunicación entre los dos módulos del modelo NET completo. Esta función de *LFY* no había sido reconocida anteriormente. Tercero, el modelo ABC describe los patrones de activación observado en los órganos florales maduros, pero no describe cómo se establece el paso entre el estado genético de no-floración a los de la aparición de las flores. El grupo de parámetros utilizados en el análisis de circuitos permite elaborar una proposición acerca de las condiciones necesarias para pasar del estado inicial de no-floración hacia los distintos estados florales. Cuarto, el análisis de circuitos predice la existencia de un regulador de la actividad del gen *LFY* que no es ninguno de los que se presentan en la red de regulación. El papel de dicho regulador sería el de promover la actividad transcripcional de *LFY* en los primeros estadios de la floración. Quinto, al igual que el artículo anterior, se predice la existencia de un estado estable de activación que no se encuentra en las plantas silvestres. El estado es el de la activación estable simultánea de *EMF1*, *TFL1*, *AP3* y *PI*. La explicación para este estado es que

AP3 y *PI* forman un módulo de la red de regulación, por lo cual una vez activo se queda en ese estado independientemente de lo que ocurra con la actividad de los otros genes de la red. Una manera de confirmar dicha predicción sería, por ejemplo, inyectando en una zona vegetativa de *Arabidopsis* una cantidad suficiente de las proteínas AP3 y PI con las cuales se iniciaría la transcripción de sus mismos genes y se cerraría el circuito positivo de autorregulación. Sexto y último, al igual que lo expresado en el artículo anterior, se propone que los circuitos presentes en este modelo son suficientes para explicar la aparición de los patrones estables de activación de los genes de identidad de órgano observados en las flores de *Arabidopsis*.

FIGURA 5
Modelo de los pelos radiculares



Como se puede apreciar, los resultados del análisis formal no sólo dan sustento a los resultados obtenidos con el sistema dinámico propuesto para el modelo NET, sino que también aporta nuevas predicciones. Tomando en cuenta los dos artículos, el modelado de la red de regulación genética que controla la floración aporta tres tipos de predicciones: 1) la existencia de relaciones de regulación aún no reconocidas entre genes ya descritos, 2) la existencia de por lo menos un gen aún no caracterizado, y 3) la existencia de por lo menos un patrón estacionario de activación genética. Las implicaciones, así como las maneras de confirmar las predicciones son discutidas con más detalle en los Apéndices 3 y 4.

5.3 El modelo de aparición de pelos radiculares

La cantidad creciente de información acerca de los mecanismos responsables de la diferenciación celular y la morfogénesis está originando la aparición de múltiples relaciones de regulación transcripcional en el desarrollo de la raíz de *Arabidopsis*. El Apéndice 5 es un artículo que presenta un modelo de red para describir la diferenciación celular durante el desarrollo de la epidermis de la raíz en *Arabidopsis*, ver Figura 5. En los siguientes párrafos se resumen las principales características de dicho modelo.

La epidermis de la raíz de *Arabidopsis* es un sistema excepcionalmente simple para estudiar a los mecanismos celulares y genéticos que determinan un patrón relativamente sencillo de diferenciación celular. La epidermis de la raíz esta formada por filas alternadas de células que presentan una proyección celular denominada *pelo* (tricoblastos), y otras células sin pelo (atricoblastos). En este sistema, el destino celular está determinado por la localización relativa de las células de la

epidermis con respecto a las células de la corteza (Scheres, 1997). Experimentos de muerte celular dirigida por láser han mostrado que el proceso de diferenciación depende del contacto directo con células ya diferenciadas (Van der Berg *et al.*, 1995; 1997). Por otra parte, la inyección de colorantes muestran que las células meristemáticas de raíz tienen citoplasmas acoplados (Duckett *et al.*, 1994), lo cual refuerza la hipótesis del contacto celular como un factor importante en la diferenciación celular. Con respecto a las moléculas involucradas, se ha identificado la participación de por lo menos cuatro factores de transcripción (ver adelante) en el desarrollo de los pelos radiculares. Finalmente, también se ha identificado a la ruta de transducción en respuesta al etileno como una de las vías de señalización involucradas en el proceso (Scheres y Wolkenfelt, 1998). A pesar de las moléculas y los procesos ya identificados, no se conoce la forma en la que se integran los distintos factores para dar lugar al patrón de diferenciación observado en la epidermis de la raíz.

Previo al modelo aquí desarrollado, no se ha desarrollado ningún otro modelo para proponer un mecanismo de regulación genética responsable de controlar el patrón de diferenciación observado en la epidermis de la raíz de *Arabidopsis*. La arquitectura de la red de regulación propuesta incluye dos vías paralelas. Una vía formada por los factores de transcripción *TRANSPARENT TESTA GLABRA* (*TTG*), el homólogo del gen *R* de maíz, *GLABRA 2* (*GL2*) y *CAPRICE* (*CPC*). Y la otra vía formada por las proteínas de transducción de señales ETHYLENE RESISTANT 1 (ETR1) y CONSTITUTIVE TRIPLE RESPONSE 1 (CTR1). Ambas vías regulan independientemente la actividad de dos genes: *AUXIN RESISTANT 2* (*AXR2*) y *ROOT HAIR DEFECTIVE 6* (*RHD6*), quienes a su vez controlan el desarrollo de los pelos radiculares. La red de regulación se implementó como un sistema dinámico, en donde la actividad de cada molécula se representó como una variable discreta con dos o tres estados. Además, se incluyó la información espacial observada en la epidermis en la forma de distintos estados iniciales de activación de la variable que representa la actividad de *CPC*. Finalmente, el hecho de incluir en el modelo al receptor de etileno (ETR1), permitió simular los efectos farmacológicos de aumentos o disminuciones en las concentraciones de etileno durante el desarrollo de la raíz.

Al igual que con los modelos de la flor, se hizo un estudio de los estados de activación a los cuales llega el sistema dinámico partiendo de todos los estados iniciales posibles. Como resultado, el modelo presenta seis estados estables de activación. El siguiente paso consistió en traducir un patrón de activación en un destino celular. Para ello se ideó una tabla en la cual se relacionan los distintos niveles de activación de *AXR2* y *RHD6* con el porcentaje de células que desarrollan pelo en una fila de células de epidermis. Se eligieron tanto a *AXR2* como *RHD6* pues son los dos elementos que se encuentran más abajo en las relaciones de regulación en nuestro modelo. Además, las mutantes dobles *axr2 rhd6* no presentan ningún pelo radicular, por lo cual es de esperar que en estado silvestre tanto *AXR2* como *RHD6* tienen un papel central en la aparición de los pelos en la epidermis. Con base en esta interpretación

morfológica de los patrones de activación, es claro que de los seis estados estacionarios de activación que alcanza el modelo, dos de ellos corresponden a los patrones de activación presentes en los tricoblastos y los atricoblastos maduros en condiciones normales, respectivamente. Los otro cuatro estados corresponden a patrones de actividad que disminuirían o aumentarían la relación de atricoblastos a tricoblastos debido a cambios en la disponibilidad del etileno.

El modelo propuesto para la raíz contrasta de manera importante en dos aspectos con los modelos para el desarrollo de las flores en *Arabidopsis*. Primero, en la red de regulación de la floración se utilizó al modelo ABC de la morfogénesis floral para traducir los estados estables de activación genética en patrones morfológicos. Debido a que no existe un modelo morfogenético similar para la formación de los pelos en la raíz, se usaron a los estados de activación de los dos genes más abajo en la vía de regulación (*RHD6* y *AXR2*) para establecer el destino celular hacia tricoblastos o atricoblastos. Segundo, los modelos de la flor fueron tratados como sistemas dinámicos autónomos en los cuales no había interacción con ninguna señal extracelular. En el modelo de raíz, sin embargo, se postula la existencia de una señal externa, vía la respuesta de los elementos CPC y ETR1 del modelo. Dichas respuestas son incorporadas fijando los estados iniciales de esos dos elementos en la red, de tal manera que la aparición de los pelos está determinada por la posición de las células de la epidermis en relación con las células de la corteza. Dicho en términos físicos, se propone la existencia de un campo.

Nuestro modelo tiene dos supuestos básicos con respecto a las señales involucradas en la determinación del destino celular, pero cabe mencionar que ambas están basadas en datos experimentales. La primera suposición es que el contacto de las células de la epidermis con una pared longitudinal anticlinal de las células de corteza es determinante para el desarrollo de los pelos (Scheres, 1996; 1997). En tanto que la segunda suposición es que el destino celular por omisión es el de atricoblasto (Berger *et al.*, 1998). Dicha información fue utilizada para establecer que el estado inicial del elemento CPC depende de la posición relativa de la célula de epidermis con respecto a las células de corteza. Esta suposición implica que una señal aún no caracterizada activa a CPC en las células que posteriormente se convierten en tricoblastos. La evidencia proveniente de experimentos de muerte celular dirigida (ver Apéndice 5), hace muy improbable la suposición de la existencia de una molécula difusible como la señal que activa CPC. Por tanto, es posible que un componente no difusible de la pared celular, como por ejemplo una glicoproteína, pudiera estar involucrada en dicho mecanismo. Un requisito indispensable para este componente putativo es que se encuentre exclusivamente en las paredes longitudinales anticlinales de las células de corteza. Cabe señalar que este tipo de patrón altamente restringido a un tipo particular de pared celular ya ha sido documentado para carbohidratos en la raíz de *Arabidopsis* (Freshour *et al.*, 1996). Por otra parte, en el modelo supusimos que el etileno es una señal disponible de manera uniforme para todas las células de la epidermis. Por tanto, el etileno no sería el portador de

ninguna información posicional durante la especificación celular. En contraste, algunos esquemas previamente publicados suponen que el etileno, o su precursor directo ACC, se difunde sólo a través de las paredes longitudinales anticlinales de las células de corteza, induciendo el desarrollo de pelos radiculares en los lugares observados (Scheres, 1997; Tanimoto, 1995). Sin embargo, existen resultados experimentales que no concuerdan con la hipótesis del etileno como señal posicional (ver discusión en el artículo del Apéndice 5).

Otros autores han propuesto representaciones esquemáticas estáticas de las interacciones genéticas involucradas en la determinación del tipo celular en la epidermis de *Arabidopsis* (Scheres y Wolkenfelt, 1998; Schiefelbein, 1998; Schiefelbein *et al.* 1997). Nosotros proponemos la misma arquitectura que ellos con respecto a las interacciones entre los genes que codifican factores de transcripción (*CPC*, *TTG*, *GL2*, y el homólogo de *R*). Sin embargo, a diferencia de los demás, nosotros proponemos que la vía de respuesta del etileno es independiente de la vía que incluye a *GL2*, aunque proponemos que ambas convergen en la regulación negativa de *AXR2* y *RHD6*. Nuestro modelo, además de describir muchos fenotipos de manera similar a otros esquemas presentados anteriormente, describe adecuadamente el efecto de la perdida de función de la familia del receptor de etileno, específicamente el mutante cuádruple *etr1 etr2 ein4 ers2*; fenotipo que no puede ser explicado por los esquemas publicados anteriormente (como por ejemplo, Tanimoto *et al.*, 1995).

El modelo arriba descrito es el primer modelo dinámico de su tipo, y constituye una nueva proposición de las interacciones de regulación entre los genes y las proteínas que controlan el desarrollo de los pelos radiculares. Notablemente, el modelo establece que la vía de transducción del etileno es paralela del grupo de los factores de transcripción que controlan la aparición de pelos. Pero además, es un modelo dinámico que puede ser utilizado para describir y predecir patrones estables de expresión, en condiciones silvestres, mutantes o bajo tratamientos farmacológicos de la raíz. La concordancia entre los resultados publicados y los simulados sugiere que este primer modelo dinámico captura muchos de los elementos importantes involucrados en la determinación genética de la diferenciación en la epidermis. Sin embargo, el modelo es aún un primer esfuerzo y necesita ser aumentado y mejorado. Por ejemplo, se conoce que el gen *ROOT HAIRLESS 1* (*RHL1*) está involucrado en la formación de los pelos radiculares (Schneider *et al.*, 1998), pero no se incluyó en el modelo pues aún no existen los resultados necesarios para inferir sus relaciones de regulación. Más aún, los datos existentes sugieren que existen genes no caracterizados todavía que intervienen en la red propuesta. Específicamente, el fenotipo del doble mutante *cpc ttg* sugiere que *GL2* tiene un nivel de expresión basal en la ausencia tanto de la inhibición por *CPC*, como de la activación por la vía *TTG/R*. Lo anterior sugiere que hay por lo menos un activador de *GL2* por descubrir.

6 CONCLUSIONES Y NOTAS FINALES

En la actualidad existen proyectos de secuenciación de genomas completos para 75 procariontes, 10 eucariontes unicelulares y 10 multicelulares. Esta explosión de información genética está causando un incremento en la necesidad de metodologías para la anotación genómica, la detección de patrones, la identificación de genes, la comparación de secuencias, etc. Además, el rápido avance en tecnologías como los microarreglos o DNA-chips hace inminente la aparición de una gran cantidad de información sobre la expresión de grupos de genes. Toda esta disponibilidad de información permite prever la aparición de un número importante de redes de regulación genética. Es por ello indispensable contar con metodologías formales para el análisis de las propiedades tanto generales como particulares de dichas redes. Desde hace algunas décadas han habido estudios teóricos sobre las propiedades de redes de estados discretos (Glass, 1975; Kauffman, 1969; Thomas, 1973). Sin embargo, sólo hasta hace poco tiempo comenzaron a aparecer modelos de redes de regulación de sistemas biológicos en particular. Con mucho, *Drosophila melanogaster* es el sistema mejor estudiado desde el punto de vista de las redes de regulación genética. La complejidad de los modelos, así como la tecnología para la adquisición de información elaborados para la mosca han alcanzado gran sofisticación. Como resultado se han presentado modelos que permiten la descripción, con gran resolución espacial y temporal, de la formación de patrones de expresión de diversos grupos de genes (ver por ejemplo Kosman y Reinitz, 1998). Pero además de los modelos para *Drosophila*, el estudio de otros sistemas biológicos está avanzando lentamente. Sin embargo, el resultado central de esta tesis fue el establecimiento de modelos dinámicos de las redes de regulación genética que controlan dos procesos de morfogénesis en *Arabidopsis*. Dichos modelos son los primeros de su tipo para un sistema vegetal.

Existen aún pocos estudios en los cuales se puedan integrar de manera clara tanto los mecanismos celulares como los moleculares para explicar la formación de patrones durante el desarrollo de algún organismo. Sin embargo, las plantas constituyen sistemas muy útiles para el desarrollo de modelos de regulación involucrados en la morfogénesis porque no existe la migración celular. La acumulación de resultados experimentales sobre la arquitectura celular y la regulación genética en diferentes órganos de *Arabidopsis thaliana* han hecho de este organismo un sistema ideal para elaborar modelos sobre la regulación genética de la morfogénesis. Para el trabajo desarrollado en esta tesis, la primera aproximación fue la de describir al grupo de genes involucrados en algún proceso de morfogénesis, y estudiarlos como un sistema dinámico cuyos estados estables de activación corresponden a los patrones de actividad genética observados experimentalmente. El siguiente paso fue el de imaginar a una célula como una red de regulación, donde el estado de activación de la red determina el destino celular. En el futuro, cuando se incorporen más detalles a dichos modelos, será posible hacer predicciones sobre el posible efecto que tendría la duplicación, eliminación o sobre-expresión de uno o

varios genes sobre el desarrollo de la planta. La utilidad de este tipo modelos es doble, por un lado sirven de guía para la realización de experimentos determinados, reduciendo las búsquedas azarosas de mutantes de interés. Por otro lado, este tipo de modelos sistematiza el conocimiento experimental y abre el camino hacia el análisis formal de los resultados experimentales.

El estudio de las redes de regulación genética como sistemas dinámicos permite obtener resultados que no son evidentes. Es una práctica común que cuando los estudios genéticos y bioquímicos sugieren una secuencia de eventos de regulación, se utilicen flechas para representar las interacciones entre las macromoléculas. Dicha representación gráfica es necesaria, pero no suficiente, para entender el comportamiento dinámico de la vía de regulación bajo estudio. Por ejemplo, ¿por qué se observan ciertos patrones de actividad y no otros? ¿Por qué ciertos patrones son transitorios y otros estacionarios? ¿Cuantos estados estacionarios son posibles para una red determinada? ¿Los estados observados son todos los posibles? ¿Cuales interacciones son necesarias para obtener un patrón de actividad necesario? ¿Existe algún tipo de redundancia funcional que no dependa de la homología en la secuencia primaria de los genes? Las respuestas a estas, y muchas otras preguntas, no son evidentes y requieren de un estudio de la dinámica de la activación transcripcional. En particular, el análisis de las redes de regulación genética como sistemas dinámicos provee de una herramienta adecuada para el análisis integrativo de una gran cantidad de información genética disponible. Más aún, los modelos de redes no sólo son útiles para la síntesis de información sino que además permiten la elaboración de predicciones que no se pueden elaborar utilizando la concepción de la regulación genética como un sistema estático y jerárquico. Por ejemplo, la predicción de interacciones genéticas, o de estados estables de expresión, son guías útiles para los biólogos experimentalistas para continuar con el análisis molecular de ciertos organismos modelo. Es claro, sin embargo, que el estudio de las redes de regulación no agota el tema de los modelos de la morfogénesis. En particular, es indispensable incorporar la información espacial a los modelos, de tal manera que se puedan describir los cambios de expresión asociados a la diferenciación celular en distintas regiones de un organismo.

El tipo de modelos presentados aquí para *Arabidopsis* permite un estudio completo del comportamiento dinámico de algunos subsistemas de dicha planta. Dicho estudio provee de información importante para elaborar modelos continuos conforme se obtengan datos cuantitativos de la dinámica de activación de cada uno de los genes involucrados. Sin embargo, es importante tener en mente que cualquier proceso de modelado implica ciertos problemas. Un problema importante al elaborar las redes de regulación de cualquier tipo es la inferencia correcta de las interacciones genéticas propuestas. El proceso de inferencia involucra un análisis concienzudo de una gran cantidad de información experimental publicada por diversos laboratorios. Este paso incluye por lo menos dos posibles fuentes de error. El primero tiene que ver con el proceso de discriminación, ¿cual

información es relevante y cual es superflua? No es evidente si un resultado experimental reportado se relaciona con un proceso de regulación, o a una respuesta molecular debida a un proceso no previsto por el modelador. Una solución ideal al problema sería la formación de grupos interdisciplinarios, en los cuales colaboraran tanto expertos en el proceso de modelado como expertos en el sistema biológico bajo estudio. Sin embargo, ésta no es una tendencia común; usualmente el modelador es quien recolecta, analiza e incorpora la información generada por los diferentes grupos experimentales en el mundo. Dicha necesidad en la forma de trabajo trae consigo el segundo problema; usualmente en diferentes laboratorios se trabajan con diferentes metodologías, subespecies animales, ecotipos vegetales, etc. Entonces, ¿cómo distinguir entre diferencias de resultados debidas a procesos de regulación, o a diferencias en las metodologías de adquisición de datos? En este aspecto no hay una respuesta sencilla. El modelador tiene que establecer un balance entre el conocimiento del sistema experimental y la intuición que le da la experiencia. Por supuesto, es muy importante mantener el contacto con los propios biólogos experimentales para poder valorar la información y las suposiciones que se hacen durante el desarrollo del modelo.

El problema de la inferencia correcta de las relaciones de regulación, sobre todo cuando se cuenta con mucha información, ha despertado el interés por el desarrollo de algoritmos para la deducción automatizada de redes genéticas, a partir de patrones temporales de transcripción. Sin embargo, muchos de los algoritmos desarrollados están basados en la suposición de que las redes de regulación se comportan como redes binarias. Por ejemplo, Liang *et al.* (1998) propone un algoritmo llamado REVEAL para la inferencia de la red partiendo de tablas de transición de estados, las cuales pueden corresponder a series temporales de expresión. El algoritmo usa un análisis de información mutua para reducir la cantidad de puntos de información necesarios para la inferencia. A pesar de la eficiencia de los algoritmos propuestos hasta ahora, su implementación en laboratorios requiere de una alta precisión en la adquisición de datos de expresión genética, lo cual es muy difícil. Sin embargo, avances recientes como el ensayo RT-PCR permite mediciones simultaneas de un gran número de genes en tiempos diversos. Un notable ejemplo de esta clase de tecnología es la matriz de expresión genética de Wen *et al.* (1998), la cual contiene la expresión de 112 genes en al sistema nervioso central de rata. A partir de esa información es posible, en principio, inferir las relaciones de regulación transcripcional de grandes grupos de genes.

La convergencia de los avances teóricos y experimentales tiene un impacto importante en la comprensión de los procesos biológicos. Por ejemplo, se sabe que todas las células de un organismo multicelular tienen el mismo genotipo, y por lo tanto la misma red de regulación dentro de cada una de ellas. Sin embargo, los patrones de expresión y las respuestas genéticas difieren entre los distintos tipos celulares. Es decir que la respuesta a los cambios ambientales y las señales celulares son particulares de cada tipo celular. Esta respuesta diferencial, a pesar del material genético común, puede ser entendida con el uso del concepto de las redes de regulación

genética como sistemas dinámicos. Como se expresó anteriormente, la red puede estar en diferentes estados de equilibrio, y cada uno de esos estados estables de expresión puede interpretarse como el perfil genético característico de cada tipo celular. Pero también existen aspectos prácticos en el uso de modelos de regulación genética. La reciente disponibilidad de grandes recursos financieros para establecer análisis genéticos de múltiples especies, especialmente en sistemas con un alto potencial económico como los cultivos, atrae el interés de muchos científicos (Bennetzen *et al.*, 1999). Los alcances técnicos para estudiar la expresión genética de genomas completos están estableciendo la pauta para la formación de grandes grupos de investigación para describir patrones de expresión genética. El entendimiento de los procesos básicos es necesario para el control del desarrollo de un organismo con importancia económica. Basta pensar, por ejemplo, en el control del proceso de floración: su control podría llevar a elevar la producción de algodón, frutas, etc., con el consiguiente impacto económico.

Todas las consideraciones anteriores apuntan a la importancia que tienen, y pueden tener, los modelos de redes de regulación. Existe una vasta literatura en la que se exploran las características generales de las redes de regulación, como por ejemplo su capacidad de evolución (Clarke *et al.*, 1993) y redundancia (Wagner, 1996). Pero los modelos desarrollados para sistemas biológicos específicos ayudan a validar o rechazar algunas de las suposiciones biológicas utilizadas con los enfoques no formales. Estas validaciones finalmente tendrían el valor de que los modelos de regulación sean más aceptados por la comunidad de biólogos interesados en la regulación genética. Además, la formalización permite la síntesis de la información experimental generada para distintos niveles de organización. El modelo de raíz presentado no es más que un esbozo de lo que pueden ser modelos que integren la regulación transcripcional con la comunicación celular. Existen razones que permiten afirmar que la interacción entre diversos niveles de organización tienen como resultado la disminución de los grados de libertad del sistema en cuestión (Goodwin *et al.*, 1993). En el caso en particular del desarrollo, las interacciones entre dos niveles de organización resultarán en la formación de patrones morfogenéticos más robustos que los obtenidos si se modelaran los dos niveles de organización por separado. Este tipo de ideas ha impulsado el desarrollo de propuestas para utilizar modelos híbridos. En particular, Mjolsness *et al.* (1991) propusieron un sistema constituido por una red neuronal acoplada a una gramática, con la intención de modelar la diferenciación celular. Sin embargo, los mismos autores reconocen que no cuentan con sistemas biológicos a los cuales se pueda aplicar esta metodología. Nuevamente, el modelo de raíz presentado en esta tesis pudiera servir como punta de lanza para el desarrollo de modelos que integren el nivel molecular de regulación genética y el nivel celular.

7 BIBLIOGRAFIA

Arnone, M.I.; and Davidson, E.H. (1997). The hardwiring of development: organization and function of genomic

genética como sistemas dinámicos. Como se expresó anteriormente, la red puede estar en diferentes estados de equilibrio, y cada uno de esos estados estables de expresión puede interpretarse como el perfil genético característico de cada tipo celular. Pero también existen aspectos prácticos en el uso de modelos de regulación genética. La reciente disponibilidad de grandes recursos financieros para establecer análisis genéticos de múltiples especies, especialmente en sistemas con un alto potencial económico como los cultivos, atrae el interés de muchos científicos (Bennetzen *et al.*, 1999). Los alcances técnicos para estudiar la expresión genética de genomas completos están estableciendo la pauta para la formación de grandes grupos de investigación para describir patrones de expresión genética. El entendimiento de los procesos básicos es necesario para el control del desarrollo de un organismo con importancia económica. Basta pensar, por ejemplo, en el control del proceso de floración: su control podría llevar a elevar la producción de algodón, frutas, etc., con el consiguiente impacto económico.

Todas las consideraciones anteriores apuntan a la importancia que tienen, y pueden tener, los modelos de redes de regulación. Existe una vasta literatura en la que se exploran las características generales de las redes de regulación, como por ejemplo su capacidad de evolución (Clarke *et al.*, 1993) y redundancia (Wagner, 1996). Pero los modelos desarrollados para sistemas biológicos específicos ayudan a validar o rechazar algunas de las suposiciones biológicas utilizadas con los enfoques no formales. Estas validaciones finalmente tendrían el valor de que los modelos de regulación sean más aceptados por la comunidad de biólogos interesados en la regulación genética. Además, la formalización permite la síntesis de la información experimental generada para distintos niveles de organización. El modelo de raíz presentado no es más que un esbozo de lo que pueden ser modelos que integren la regulación transcripcional con la comunicación celular. Existen razones que permiten afirmar que la interacción entre diversos niveles de organización tienen como resultado la disminución de los grados de libertad del sistema en cuestión (Goodwin *et al.*, 1993). En el caso en particular del desarrollo, las interacciones entre dos niveles de organización resultarán en la formación de patrones morfogenéticos más robustos que los obtenidos si se modelaran los dos niveles de organización por separado. Este tipo de ideas ha impulsado el desarrollo de propuestas para utilizar modelos híbridos. En particular, Mjolsness *et al.* (1991) propusieron un sistema constituido por una red neuronal acoplada a una gramática, con la intención de modelar la diferenciación celular. Sin embargo, los mismos autores reconocen que no cuentan con sistemas biológicos a los cuales se pueda aplicar esta metodología. Nuevamente, el modelo de raíz presentado en esta tesis pudiera servir como punta de lanza para el desarrollo de modelos que integren el nivel molecular de regulación genética y el nivel celular.

7 BIBLIOGRAFIA

Arnone, M.I.; and Davidson, E.H. (1997). The hardwiring of development: organization and function of genomic

regulatory systems. *Development* **124**: 1851-1864.

Aubert, D. (1998). EMF genes regulate shoot development in *Arabidopsis*. *Proceedings of the 9th international conference on Arabidopsis research*. University of Wisconsin, Madison. p156.

Baum, S.F.; and Rost, T.L. (1996). Root apical organization in *Arabidopsis thaliana* 1. Root cap and protoderm. *Protoplasma* **192**: 178-188.

Beemster, G.T.S.; and Baskin, T.I. (1998). Analysis of cell division and elongation underlying the developmental acceleration of root growth in *Arabidopsis* thaliana. *Plant Physiol.* **116**: 1515-1526.

Benfey, P.N. (1999). Is the shoot a root with a view? *Curr. Op. Plant Biol.* **2**: 39-43.

Bennetzen, J.L. (1999). Plant genomics takes root, branches out. *Trends Genet.* **15**: 85-87.

Berger, F.; Haseloff, J.; Schiefelbein, J. & Dolan, L. (1998). Positional information in root epidermis is defined during embryogenesis and acts in domains with strict boundaries. *Curr. Biol.* **8**: 421-430.

Bodnar, J.W. (1997). Programming the *Drosophila* embryo. *J. Theor. Biol.* **188**: 391-445.

Bowman, J. (1994). *Arabidopsis*: An atlas of morphology and development. Springer-Verlag.

Burstein, Z. (1995). A network model of developmental gene hierarchy. *J. theor. Biol.* **174**: 1-11.

Clark, S.E.; Running, M.P.; and Meyerowitz, E.M. (1995). CLAVATA3 is a specific regulator of shoot and floral meristem development affecting the same process as CLAVATA1. *Development* **121**: 2057-2067.

Clarke, B.; Mittenthal, J.E.; and Senn, M. (1993). A model for the evolution of networks of genes. *J. theor. Biol.* **165**: 269-289.

Coen, E.S. (1991). The role of homeotic genes in flower development and evolution. *Annu. Rev. Plant Physiol Plant Mol. Biol.* **42**: 241-279.

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

Coupland, G. (1995). Genetic and enviromental control of flowering time in *Arabidopsis*. *Trends Genet.* **11**: 393-397.

Davidson, E.H.; Cameron, R.A. and Ransick, A. (1998). Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. *Development* **125**: 3269-3290.

Dolan, L.; Janmaat, K.; Willemsen, V.; Linstead, P.; Poethig, S.; Roberts, K.; and Scheres, B. (1993) .Cellular organisation of the *Arabidopsis* thaliana root. *Development* **119**: 71-84.

Dolan, L.; Duckett, C.M.; Grierson, C.; Linstead, P.; Schneider, K.; Lawson, E.; Dean, C.; Poethig, S.; and Roberts, K. (1994). Clonal relationships and cell patterning in the root epidermis of *Arabidopsis*. *Development* **120**: 2465-2474.

Duckett, C.M.; Oparka, K.J.; Prior, D.A.M.; Dolan, L.; and Roberts, K. (1994). Dye-coupling in the root epidermis of *Arabidopsis* is progressively reduced during development. *Development* **120**: 3247-3255.

Freshour, G.; Clay, R.P.; Fuller, M.S.; Albersheim, P.; Darvill, A.G. and Hahn, M.G. (1996). Developmental and tissue-specific structural alterations of the cell-wall polysaccharides of

Arabidopsis thaliana roots. *Plant Physiol.* **110**: 1413-1429.

Glass, L. (1975). Classification of biological networks by their qualitative dynamics. *J. theor. Biol.* **54**: 85-107.

Golberg, R.B.; De Pavia, G.; and Yadegari, R. (1994). Plant embryogenesis: zygote to seed. *Science* **266**: 605-614.

Goodwin, B. C.; Kauffman, S. A.; and Murray, J. D. (1993). Is morphogenesis an intrinsically robust process? *J. theor. Biol.* **163**: 135-144.

Gouzé, J.L. (1998). Positive and negative circuits in dynamical systems. *J. Biol. Systems* **6**: 11-15.

Hunding, A.; and Engelhardt, R. (1995). Early biological morphogenesis and nonlinear dynamics. *J. theor. Biol.* **173**: 401-413.

Jack, T.; Fox, G.L.; and Meyerowitz, E.M. (1994). *Arabidopsis* homeotic gene APETALA3 ectopic expression; transcriptional and posttranscriptional regulation determine floral organ identity. *Cell* **76**: 703-716.

Kauffman, S.A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *J. theor. Biol.* **22**: 437-467.

Kauffman, S.A. (1991). Antichaos and adaptation. *Sci. Am.* **265**: 64-71.

Kauffman, S.A. (1993). *The origins of order: self-organization and selection in evolution.* Oxford University Press.

Kosman, D.; and Reinitz, J. (1998). Automated assay of gene expression at cellular resolution. *Pacific Symposium on Biocomputing* **3**: 6-17.

Krizek, B.A.; and Meyerowitz, E.M. (1996). The *Arabidopsis* homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. *Development* **122**: 11-22.

Laskowski, M.J.; Williams, M.E.; Nusbaum, H.C.; and Sussex, I.M. (1995). Formation of lateral root meristem is a two-stage process. *Development* **121**: 3303-3310.

Laux, T.; and Jürgens, G. (1997). Embryogenesis: a new start in life. *Plant Cell* **9**: 989-1000.

Levin, J.Z.; and Meyerowitz, E.M. (1995). UFO: An *Arabidopsis* gene involved in both floral meristem and floral organ development. *Plant Cell* **7**: 529-548.

Liang, S.; Fuhrman, S.; and Somogyi, R. (1998). REVEAL, a general reverse engineering algorithm for inference of genetic network architectures. *Pacific Symposium on Biocomputing* **3**: 18-29.

Liu, Z.; and Meyerowitz, E.M. (1995). LEUNIG regulates AGAMOUS expression in *Arabidopsis* flowers. *Development* **121**: 975-991.

Ma, H. (1994). The unfolding drama of flower development: recent results from genetic and molecular analyses. *Genes & Dev.* **8**: 745-756.

Malamy, J.E.; and Benfey, P.N. (1997a). Organization and cell differentiation in lateral roots of *Arabidopsis thaliana*. *Development* **124**: 33-44.

Malamy, J.E.; and Benfey, P.N. (1997b). Down and out in *Arabidopsis*: the formation oflateral roots. *Trends Plant Sci.* **2**: 390-396.

Meyerowitz, E. M. (1994a). The genetics of flower development. *Sci. Am.* **271**: 40-47.

Meyerowitz, E. M. (1994b). Flower development and evolution: New answers and new questions. *Proc. Natl. Acad. Sci. USA* **91**: 5735-5737.

Meyerowitz, E.M. (1997). Plants and the logic of development. *Genetics* **145**: 5-9.

Mjolsness, E.; Sharp, D. H.; and Reinitz, J. (1991). A connectionist model of development. *J. theor. Biol.* **152**: 429-453.

Monod, J.; and Jacob, F. (1961). General conclusions: telenomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harbor Symp. Quant. Biol.*, **26:** 389-401.

Moreno, G. (1994). Genetic architecture, genetic behavior, and character evolution. *Annu. Rev. Ecol. Syst.* **25**: 31-44.

Muraille, E.M.; Thieffry, D.; Leo, O.; and Kaufman, M. (1996). Toxicity and neuroendocrine regulation of the immune response: a model analysis. *J. theor. Biol.* **183**: 285-305.

Okamuro, J.K.; Der Boer, B.G.W.; and Jofuku, K.D. (1993). Regulation of *Arabidopsis* flower development. *Plant Cell* **5**: 1183-1193.

Parcy, F.; Nilsson, O.; Busch, M.A.; Lee, I.; and Weigel, D. (1998). A genetic framework for floral patterning. *Nature* **395**: 561-566.

Plahte, E.; Mestl, T.; and Omholt, S.W. (1995). Feedback loops, stability and multistationarity in dynamical systems. *J. Biol. Syst.* 3: 409-413.

Puterill, J.; Robson, F.; Lee, K.; Simon, R.; and Coupland, G. (1995). The CONSTANS gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* **80**: 847-857.

Ray, A.; Robinson-Beers, K.; Ray, S.; Baker, S.C.; Lang, J.D.; Preuss, D.; Milligan, S.B.; and Gasser, C.S. (1994). *Arabidopsis* floral homeotic gene BELL (BEL1) controls ovule development through negative regulation of AGAMOUS gene (AG). *Proc. Natl. Acad. Sci. USA.* **91**: 5761-5765.

Reinitz, J.; and Sharp, D. H. (1995). Mechanism of eve stripe formation. *Mech. Dev.* **49**: 133-158.

Rosen, R. (1968). Recent developments in the theory of control and regulation of cellular processes. *Int. Rev. of Cytology* **23**: 25-88.

Rost, T.L.; Baum, S.F.; and Nichol, S. (1996). Root apical organization in *Arabidopsis thaliana* ecotype 'WS' and a comment on root cap structure. *Plant and Soil* **187**: 91-95.

Sakai, H.; Medrano, L.J.; and Meyerowitz, E.M. (1995). Role of SUPERMAN in maintaining *Arabidopsis* floral whorl boundaries. *Nature* **378**: 199-203.

Scheres, B. (1996). Embryo patterning genes and reinforcement cues determine cell fate in the *Arabidopsis thaliana* root. *Seminars Cell & Dev. Biol.* **7**: 857-865.

Scheres, B. (1997). Cell signaling in root development. *Curr. Op. Genet. & Develop.* **7**: 501-506.

Scheres, B. and Wolkenfelt, H. (1998). The *Arabidopsis* root as a model to study plant development. *Plant Physiol. Biochem.* **36**: 21-32.

Schiefelbein, J. (1998). Cell-fate specification in the root epidermis. *Trends Plant Sci.* **3**: 3-4

Schiefelbein, J.W.; Masucci, J.D.; and Wang, H. (1997). Building a root: the control of patterning and morphogenesis during root development. *Plant Cell* **9**: 1089-1098.

Schneider, K.; Mathur, J.; Boudonck, K.; Wells, B.; Dolan, L.; and Roberts, K. (1998). The ROOT HAIRLESS 1 gene encodes a nuclear protein required for root hair initiation in *Arabidopsis*. *Genes Dev.* **12**: 2013-2021.

Schultz, E.A.; and Haughn, G.W. (1993). Genetic analysis of the floral

initiation process (FLIP) in *Arabidopsis*. *Development* **119**: 745-765.

Snoussi, E.H. (1989). Qualitative dynamics of piece-linear differential equations: a discrete mapping approach. *Dyn. Stability Syst.* **4**: 189-207.

Snoussi, E.H. (1998). Necessary conditions for multistationarity and stable periodicity. *J. Biol. Systems* **6**: 3-9.

Snoussi, E.H.; and Thomas, R. (1993). Logical identification of all steady states: the concept of feedback loop characteristic states. *Bull. Math. Biol.* **55**: 973-991.

Somogyi, R.; and Sniegoski, C.A. (1996). Modeling the complexity of genetic networks: understanding multigenic and pleiotropic regulation. *Complexity* **1**: 45-63.

Spirov, A.V. (1996). Self-assemblage of gene nets in evolution via recruiting of new netters. *CSTB bulletin* - **Spring 1996**. (http://dbm.ulb.ac.be/~denis/cstbspring96/CSTB Spring96 2.html)

Sussex, I.M.; Godoy, J.A.; Merk, N.M.; Laskoski, M.J.; Nusbaum, H.C.; Welsch, J.A.; and Williams, M.E. (1995). Cellular and molecular events in a newly organizing lateral root meristem. *Phil. Trans. R. Soc. Lond. B.* **350**: 39-43.

Tanimoto, M.; Roberts, K. and Dolan, L. (1995). Ethylene is a positive regulator of root hair development in *Arabidopsis thaliana*. *Plant J.* **8**: 943-948.

Theiβen, G.; and Saedler, H. (1995). MADS-box genes in plant ontogeny and phylogeny: Hackel's "biogenetic law" revisited. *Current Op. Gen. Dev.* **5**: 628-639.

Thieffry, D.; Snoussi, E.H.; Richelle, J.; and Thomas, R. (1995). Positive loops and differentiation. *J. Biol. Syst.* **3**: 457-466.

Thomas, R. (1973). Boolean formalization of genetic control circuits. *J. theor. Biol.* **42**: 563-585.

Thomas, R. (1991). Regulatory networks seen as asynchronous automata: a logical description. *J. theor. Biol.* **153**: 1-23.

Thomas, R.; Thieffry, D.; and 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.

Trewavas, A.J.; and Malhó, R. (1997). Signal perception and transduction: the origin of the phenotype. *Plant Cell* **9**: 1181-1195.

Van der Berg, C.; Willemsen, V.; Hage, W.; Weisbeek, P.; and Scheres, B. (1995). Cell fate in the *Arabidopsis* root meristem determined by directional signalling. *Nature* **378**: 62-65.

Van der Berg, C.; Willemsen, V.; Hendriks, G.; Weisbeek, P.; and Scheres, B. (1997). Short range control of cell differentiation in the *Arabidopsis* root meristem. *Nature* **390**: 287-289.

Van der Berg, C.; Weisbeek, P.; and Scheres, B. (1998). Cell fate and cell differentiation status in the *Arabidopsis* root. *Planta* **205**: 483-491.

Wagner, A. (1996). Genetic redundancy caused by gene duplication and its evolution in networks of transcriptional regulators. *Biol. Cybern.* **74**: 557-567.

Weigel, D. (1995). The genetics of flower development: from floral induction to ovule morphogenesis. *Annu. Rev. Genetics* **29**: 19-39.

Weigel, D.; Alvarez, J.; Smyth, D.; Yanofsky, M.F.; and Meyerowitz, E.M. (1992). LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* **69**: 843-859.

Weigel, D.; and Meyerowitz, E.M. (1993a). Genetic hierarchy controlling flower development. In *Molecular basis of morphogenesis*. pp 93-107. Wiley-Liss, Inc.

Weigel, D.; and Meyerowitz, E.M. (1993b). Activation of floral homeotic genes in *Arabidopsis*. *Science* **261**: 1723-1726.

Wen, X.; Fuhrman, S.; Michaelis, G.S.; Carr, D.B.; Smith, S.; Barker, J.L.; and Somogyi, R. (1998). Large-scale temporal gene expression mapping of central nervous system development. *Proc. Natl. Acad. Sci. USA*. **95**: 334-339.

Yuh, C.H.; Bolouri, H.; and Davidson, E.H. (1998). Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science* **279**: 1896-1902.

Zuckerkandl, E. (1994). Molecular pathways to parallel evolution: I. Gene nexuses and their morphological correlates. *J. Mol. Evol.* **39**: 661-678.

APENDICE 1

THE ROOT OF *Arabidopsis thaliana* AS A BIOLOGICAL SYSTEM SUITABLE FOR PATTERN FORMATION STUDIES: MOLECULAR GENETICS AND CELL-BIOLOGY APPROACHES

Luis Mendoza and Elena Alvarez-Buylla

Lab. de Genética Molecular y Evolución
Instituto de Ecología, UNAM. Ap. Postal 70-275
Coyoacán, D.F. CP04510. MEXICO

ABSTRACT

The ample use of *Arabidopsis thaliana* as a model system has helped to understand the cellular and molecular mechanisms controlling plant development and morphogenesis. In particular, the *Arabidopsis* primary root presents the convenience that cellular fate can be traced and predicted with accuracy. Classical genetic and molecular analyses have shown that many genes involved is cellular determination act early during root development. Furthermore, laser ablation experiments show that intercellular contacts are determinant to establish cellular fate. Therefore it is clear that developmental pathways arise from a coordinated interplay between molecular and cellular cues. To better understand those two levels of organization, we review most of the mutations that affect the normal development of the primary root in *Arabidopsis*, and some of the results from cellular ablation.

INTRODUCTION

The study of the molecular basis underlying development and morphogenesis has been an active area of biological research since the first morphogens were discovered (the *bicoid* protein is by far the most studied. Driever and Nüsslein-Volhard, 1988). However, most of these studies have used animals as model systems. *Drosophila melanogaster*, *Caenorhabditis elegans*, *Xenopus laevis*, *Mus musculus* being the most thoroughly studied ones. In contrast, plants have become the focus of molecular geneticists only for the last two or three decades. Great progress has been achieved in understanding the genetic and molecular mechanisms underlying plant development, thanks to the use of *Arabidopsis thaliana* as a model system. These studies are showing that plants and animals are very similar in their housekeeping genes, in many aspects of the regulatory networks controlling development and in their biochemical pathways. However, plants and animals differ considerably in the mechanisms of signal acquisition by cells (Trewavas and Malhó, 1997) and the mechanisms of cellular communication (Meyerowitz, 1997). Therefore, the overall control of morphogenesis and pattern formation seems to be quite different in plants and animals. Consequently, the study of a plant model system has become now more relevant.

Flower development has received special attention and many genes involved in this process and their interactions have now been characterized at the genetic and molecular levels (Theiβen and Saedler, 1995; Weigel, 1995). However, there are still many open questions in plant development concerning the signaling pathways, cell-cell interactions and the integration of internal and environmental cues among cells that orchestrate plant development and morphogenesis. To address many of these issues, the root is becoming a useful system. Nevertheless, we still know very little about the molecular nature of the genes involved in

cell differentiation and pattern formation in this plant organ. In contrast with the dozens of root mutants affecting different aspects of development, the sequences of only four of these genes are known up to now. However, it is highly probable that several root-specific genes for which no mutant phenotype is still available will be known soon as the *Arabidopsis* genome sequencing project progresses. Finally, the reverse genetic approaches available now for this plant should enable the pursuit of loss-of-function mutants for these genes in the somewhat near future.

Roots have several advantages over aerial shoots to study the biochemical, molecular, genetic and cellular mechanisms underlying morphogenesis. The advantages are: a) roots have a relatively simple cylindrical form with stereotyped cell differentiation and division patterns, b) in some of the root tissues the number of cells is small and fixed or nearly so, c) the root pattern formation can be traced to the plant embryo, d) main and lateral roots have the same cellular structure and overall architecture, and the lateral organs arise from a cell layer within the primary root rather than from lateral buds with complicated phyllotaxis as in the shoot apex, and e) roots present a continuous development without large shifts in developmental stages as in the aerial part of the plant. Some technical aspects also facilitate now the study of roots from a molecular perspective. The stereotyped structure of the root has attracted the attention of many developmental biologists and embryologists, and there are detailed descriptions of the wild type root cell division patterns that constitute an excellent reference for the analysis of mutants. Also there are a number of enhancer trap lines for roots that are yielding cell type markers, to trace cell types at early developmental stages. As a drawback, however, the elegant simplicity of the root limits the number of obvious developmental mutations, such as the many organ types, size and number, and homeotic mutations available in flowers (see for example Coen, 1991; Weigel 1995; Yanofsky, 1995). These mutations have been very useful to establish genetic and molecular models of pattern formation in flowers.

In this review we describe the wild type root architecture, then we describe the phenotypes of most of the root mutants available up to now, as well as the genes responsible for these mutations in the few cases that this is known. Finally, we review recent cell ablation experiments that are providing cues to the cell interaction mechanisms underlying cell fate determination in the root.

ROOT ARCHITECTURE IN WILD TYPE *Arabidopsis thaliana*

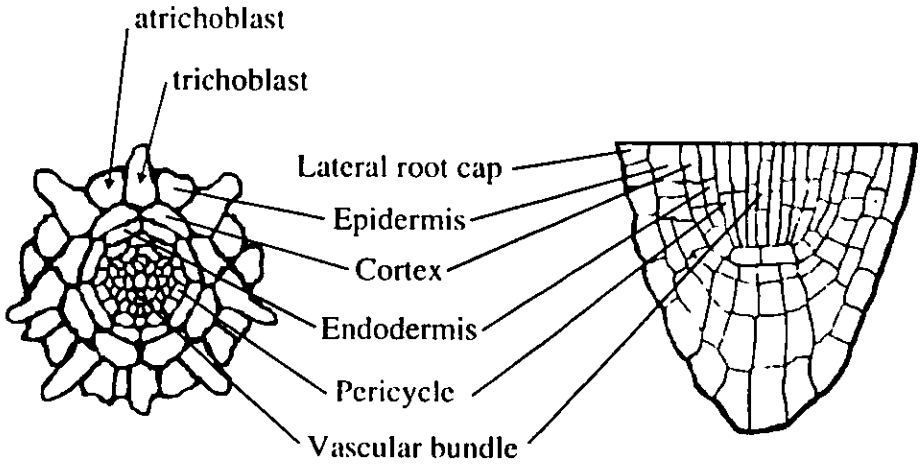
The architecture of the primary root is laid down during embryogenesis. Various previous reviews have considered *Arabidopsis thaliana* embryogenesis (Golberg, De Pavia and Yadegari, 1994; Laux and Jürgens, 1997), but we focus here to those embryogenesis events most relevant for root development. The *Arabidopsis thaliana* zygote expands along the apical-basal axis to later divide asymmetrically producing a small apical and a large basal cell. While the apical cell gives rise to most of the embryo, the basal cell forms the very basal end of the plant. The root originates, therefore, from both apical and basal cells. The quiescent center and columella root cap regions come from the basal cell, while vascular bundle, lateral root cap, cortex and endodermis come from the apical cell (Scheres *et al.*, 1994). The apical cell undergo two consecutive rounds of vertical divisions and the resulting cells then divide horizontally, forming the octant-stage embryo. These embryonic cells divide tangentially to give an inner cell mass and an outer layer of epidermal precursor cells to yield the dermatogen embryonic stage. By this time, the basal cell has divided and its uppermost derivative joins the

aforementioned cell group and becomes the hypophysis. Subsequent cell differentiation and division events result in the production of an inner procambium tissue layer and a middle layer of ground meristem cells. In this globular stage embryo, therefore, the three basic, primordial layers are established: the protoderm, ground tissue and procambium, that will become the epidermal, parenchyma and vascular tissues, respectively. After the globular stage, the cotyledons are specified, the hypocotyl elongates and the root meristem differentiates from the hypophysis at the bottom of the embryo. By this time the embryo acquires the shape of a heart, and the body plan altogether with the tissue layers of the mature embryo are already formed.

experiments have revealed four different types of initials: epidermis/lateral root cap, cortex/endodermis, vascular and columella initials. While initial cells only divide, the resulting daughter cells are able to divide and to differentiate into one of the mature cell types. In contrast to the development of the shoot, root development is continuos without significant developmental phase transitions. This continuous development yields three partially overlapping recognizable regions along the root axis: at the tip of the root is the meristem, immediately above it is located the elongation zone and then the differentiation zone. The meristematic zone occupies the most distal 250 μm of the root, and is formed by small cells covered by the root cap. Adjacent to this region lies the elongation

FIGURE 1

Schematic representation of transversal (left) and longitudinal (right) sections of the *Arabidopsis thaliana* primary root



The *Arabidopsis thaliana* mature root consists of single concentric layers of lateral root cap, epidermis, cortex, endodermis and pericycle, surrounding a vascular bundle. Every cell file is maintained by the elongation and subsequent anticlinal division (perpendicular to the apical-basal axis) of initial cells. Transposon excision

zone spanning about 250 μm also. Finally, the differentiation zone is where cells from all root tissue mature into their final developmental fate (Dolan *et al.*, 1993; Dolan *et al.*, 1994).

The root meristem consists of the initial cells and a mitotically inactive center of four cells called the quiescent center. Initials surround the quiescent center, making

direct contact with it from above, around and below, the vascular, cortex and columella initials respectively. Epidermal initials, on the other hand, are somehow adjacent and below the quiescent center cells (see Figure 1). This meristem generates new cells for the growth and development of the root throughout the plant's life. Above the apical meristem, the rate of cell division decreases while cells enlarge and differentiate into specialized cell types. Therefore, root growth results from a combination of expansion and proliferation. See Beemster and Baskin (1998) for a study on the rates of cell division underlying *Arabidopsis* root development.

The columella initials form a single layer of nearly twenty cells that divide periclinally (parallel to the apical-basal axis) producing the columella root cap. Surrounding these cells, there is a collar of cells that give rise to the epidermis and root cap. These initials divide periclinally in an asynchronic and sequential pattern around the columella initials. This sequence of periclinal divisions in the root cap initials gives rise to two cell layers. The inner cells become more initial cells that maintain the pool that originated them in the first place, while the outer cells become the peripheral root cap. Such pattern of division gives rise to a spiral that surrounds the rest of the root tip (Baum and Rost, 1996; Rost, Baum and Nichol, 1996).

Epidermal initials enlarge, divide and give rise to daughter cells that form the mature epidermis. This tissue is not homogeneous, however, it has two types of cells: hair (trichoblasts) and non-hair (atrichoblasts) cells. Interestingly, the location of these two cell types is highly predictable and it depends on their spatial location with respect to the cortical cells. On average, each epidermal ring consists of 19 cells, which generally surround eight cortical cells. This numerical difference between cortical and epidermal cells causes that some epidermal cells contact only one cortex cell while others contact two of them. The epidermal cells that are in contact with two underlying cortex cells will develop into trichoblasts, while those epidermal cells that contact only one cortical cell become atrichoblasts. The ratio of epidermal to cortical cells is maintained constant during root growth, and both trichoblasts and atrichoblasts form files, with a ratio that ranges between 0.5 and 1 of hair and non-hair files. Trichoblasts are shorter and have denser cytoplasm than atrichoblasts (Dolan *et al.*, 1994).

The cortex and endoderm cells arise from the ground tissue formed during embryogenesis. As it occurs with the epidermal and lateral root cap layers, the most basal cells from the ground tissue behave as initials dividing continually by anticlinal divisions and elongating. This ring of initials is almost always formed by eight cells and this number is settled before the globular stage embryo (Scheres *et al.*, 1995). These cells undergo a few proliferative divisions, giving rise to cortical daughters. Later on, cortical daughters divide periclinally and asymmetrically to give rise to cells of unequal size. The smallest innermost cell becomes an endodermis cell, and the larger outermost one becomes a cortex cell (Scheres, 1997). Recent laser ablation experiments have shown that the differentiation process depends upon contacting already differentiated cells, when this contact is eliminated cortex initial cells only grow and divide anticlinally (Van der Berg *et al.*, 1995).

Finally, in the lower region (the low lower tier) of the globular/triangular embryo, a series of regular periclinal divisions create four inner and four outer cells that constitute the pericycle and the vascular primordium, respectively. Later, the cell number forming the pericycle increases to approximately eight or nine by means of anticlinal divisions, until the final number of initials found in the

seedling root is attained (Scheres *et al.*, 1995).

Cell ablation and dye injection experiments demonstrate that cell-cell interactions are important during root development and morphogenesis. Dye injection experiments have shown that most of the injected cells in the tip zone are symplastically coupled, while very few are coupled in the differentiation zone. The movement of the dye is seen preferentially towards cells of the same type rather than among cells of different tissue type (Duckett *et al.*, 1994).

In contrast to the primary root, secondary roots originate from meristems formed postembryonically from founder cells in the pericycle (Dolan *et al.*, 1993; Sussex *et al.*, 1995). Most of the lateral roots emerge from the primary root and some originate from other lateral roots. This review, however, is restricted to mutants that alter mainly the primary root. But despite their different origin lateral roots have an architecture identical to that of the primary root. This suggests that the same set of genes is involved in development of both lateral and primary roots (Laskowski *et al.*, 1995; Malamy and Benfey, 1997a; Malamy and Benfey, 1997b).

MUTANTS

There is a large collection of mutants that disrupt primary root development. It is estimated that between 15 and 50 genes control root pattern formation in the *Arabidopsis thaliana* embryo (Jürgens *et al.*, 1991). In this section, we will describe briefly the morphological effects of most of the published mutants in the primary root. The chromosomic location of these genes is listed in Table 1. In this review we first discuss the mutants that affect the embryonic root, then those that alter cell division and elongation patterns, and finally we review mutants that disrupt cell fate or patterning during the development of the primary root.

Several mutations that affect root development during embryogenesis have been identified. Mutations in the *GNOM* (*GN*) gene produce zygotes with no apical-basal polarity already from the first cell division. In wild type embryos, after the first asymmetrical cell division, the resulting apical cell is about one-third the size of the basal cell. In contrast, in *gn* mutants the apical cell is only slightly smaller than the basal one, and both cells are smaller than in wild type. Other alterations become evident as zygotic development proceeds. Most *gn* mutants do not show root differentiation at all and these mutants never develop a root meristem, suggesting that *GN* is normally required for proper root formation. This gene affects, however, other aspects of plant development. The strongest mutant phenotypes do not have cotyledons and completely lack an apical-basal axis. In the developing *gn* seedlings, the radial pattern of the epidermal, ground and vascular tissues is present although the latter consists mainly of single cells, rather than of interconnected strands, as in wild type. Despite their radial tissue organization, the pattern of division of these mutants is highly distorted, the cell arrangement is abnormal, and the number of cells at the quadrant and octant stages is larger than in wild type embryos of the same stages. These phenotypes suggest that *GN* is required before the first asymmetric zygotic cell division, when apical-basal polarity first becomes visible, and that it participates in the formation of the root (Mayer, Büttner and Jürgens, 1993).

Mutations in *MONOPTEROS* (*MP*) lack root and hypocotyl, and have an altered number and position of cotyledons. Shoots of these mutants are normal. All strong *mp* mutant seedlings have a conical basal structure formed of large cells with no morphological or tissular characteristics typical of roots. This gene seems, therefore, to be necessary for patterning of basal structures as the hypocotyl, the radical and

root meristems. Abnormalities of *mp* mutants can be traced to the octant stage of the embryo where four, instead of the normal two, embryonic cell tiers are attached to the suspensor. Consequently, *mp* embryos fail to specify the lower tier at the octant stage, resulting in an abnormal cell division pattern (Berleth and Jürgens, 1993). *mp* mutants also show alterations in the embryonic vascular system. Early in embryogenesis, *mp* mutants lack provascular cells in the basal region that gives rise to the hypocotyl and primary root. This gene has been cloned recently and *MP* encodes a protein similar to the *Arabidopsis* ARF1 transcription factor, which binds to auxin-induced genes. *In situ* hybridization experiments show that *MP* mRNA is present in all subepidermal cells in globular embryos, and its mRNA expression is limited to the central domains and the embryo axis later in development at the heart stage (Hardtke and Berleth, 1998). *gn mp* double mutants have seedlings with the *gn* phenotype, suggesting that an intact *GNOM* product is necessary for *MONOPTEROS* activity, thus it is very probable that these two genes are in the same pathway (Mayer *et al.*, 1993).

Mutations in the *FASS* (*FS*) gene produce abnormalities in the whole plant. Seedling shape is altered, although overall body patterning is maintained. *FS* product is probably required throughout development for proper cell-wall orientation. Alterations are evident from the heart embryo stage during which cells have abnormal shapes. Later on, seedlings show an extremely compressed apical-basal axis with an enlarged diameter. The root and root meristem can be recognized. The root tip of *fs* mutants is made of compressed cells, probably belonging to the root cap, that overlap with small and densely cytoplasmic meristematic cells. The quiescent center is apparently absent. These mutants can attain maturity but they are very tiny plants (Torrez-Ruiz and Jürgens, 1994). These mutants are important because they show that morphogenesis in *Arabidopsis* can proceed almost normally even though precise cell patterning is heavily disrupted.

Several mutations that affect the root apical meristem have been identified. In any of the *ROOT MERISTEMLESS* (*RML1* and *RML2*) mutants, the primary root growth stops after germination. In *rml* seedlings root growth is arrested early, they lack a root meristem, and have nodule-like lateral root outgrowth. Shoot development, however, is normal in these mutants. There are some differences between *rml1* and *rml2* plants, while the former produce shorter roots, lateral roots and a curved hypocotyl, the latter have roots larger than those in *rml1*, do not give rise to lateral roots and the hypocotyl is straight. At the cellular level, *rml1* seedlings lack half-sized daughter cells, and maintain the same cellular arrangement as the wild type embryonic root. The arrest in root growth is due to a lack of cell division in *rml1* and to limited division in the *rml2* root tips as evidenced by ^{3}H-thymidine incorporation. Despite the apparent embryonic root morphology, most root apical cells do not remain undifferentiated, mutants have hairs near the root tip evidencing premature cell differentiation. Furthermore, instead of the meristematic cells, mutants present highly vacuolated epidermal and cortical cells, again suggesting that the root apex have undergone terminal differentiation (Cheng, Seeley and Sung, 1995).

Plants with the *hobbit* (*hbt*) mutation have a characteristic short embryonic root with no meristematic activity. The phenotypic effects of this mutation are not restricted to the root, however. The whole seedling has a stout appearance. The root and hypocotyl of these mutants have all the wild type cell types and the normal radial arrangement. But in the root cap of *hbt* mutants there are irregular cell divisions that affect the columella, eliminating in this way the layered structure present in wild type root cap. Another feature of the root of *hbt* plants

is the absence of a recognizable quiescent center, and in the case of strong mutants there are also root hairs in irregular positions. The quiescent center and the columella root cap, including columella initials, originate from the hypophysis in the basal embryonic region. Therefore, it is very probable that the earliest effect of the *hbt* mutation is during the development of the hypophyseal cell region. The short embryonic root is due to a reduction in the number of cells at the root apex, which is caused by the absence or extreme reduction of mitotic activity of the root meristem initials that flank the hypophyseal region at the heart stage of the embryo (Willemsen *et al.*, 1998).

Several genes seem to play important roles in root development by affecting the rate or capacity of cell division, the orientation of cell division planes, the rate and axis of cell elongation, and the asymmetry in the division of initial cells that precedes cell differentiation of daughter cells. *tonneau* (*ton1* and *ton2*) mutants grow as thick small plants. Organs are present in their normal relative positions, but root sections reveal that cell division patterns are completely disordered and no cell files are formed. Root cells of these mutants do not have normal cell elongation and their division planes are also abnormal. These cell

TABLE 1

GENE	LOCATION	REFERENCE
ANR1	Sequence available in GenBank	Zhang and Forde, 1998
COB	Chromosome 5. Position 90 *.	Hauser *et al.*, 1995
CTR1	Top of chromosome 5.	Schneider *et al.*, 1997
CUD	Chromosome 4. Position 76 *.	Hauser *et al.*, 1995
ERH1	Middle of chromosome 5.	Schneider *et al.*, 1997
ERH3	Bottom of chromosome 1.	Schneider *et al.*, 1997
FS	Chromosome 5. At 5cM from *ttg*, and closely linked to *GA3*.	Torres-Ruiz and Jürgens, 1994
GL2	Bottom of chromosome 1. Gene sequence available.	Schneider *et al.*, 1997 Rerie, Feldmann and Marks, 1994
GN	Chromosome 1. At 24cM from *dis-1* and 2cM from *ga-4*.	Mayer *et al.*, 1993
HBT	Chromosome 2. At 15.4cM from marker m246, and 8.2cM from GPA1.	Willemsen *et al.*, 1998
LIT	Chromosome 5. Position 70 *.	Hauser *et al.*, 1995
MP	Chromosome 1. At 25cM from *angustifolia*. Sequence available.	Berleth and Jürgens, 1993
PAS1	Chromosome 3.	Mentioned in Faure *et al.*, 1998
PAS2	Top of chromosome 5. At 2cM of nga249, and 4.1cM of nga151.	Faure *et al.*, 1998
PAS3	Cromosome 1. At 1cM of GAPB.	Faure *et al.*, 1998
POM1	Chromosome 1. Position 5 *.	Hauser *et al.*, 1995
POM2	Chromosome 2. Position 60 *.	Hauser *et al.*, 1995
QUI	Top of chomosome 1.	Hauser *et al.*, 1995. Schneider *et al.*, 1997
RGL3	Middle of chromosome 3.	Schneider *et al.*, 1997
RHD6	Bottom of chromosome 1.	Sneider *et al.*, 1997
RHL1	Middle of chromosome 1.	Schneider *et al.*, 1997
RHL2	Top of chromosome 5.	Schneider *et al.*, 1997
RML1	Middle of chromosome 4.	Cheng *et al.*, 1995
RML2	Lower arm of chromosome 3.	Cheng *et al.*, 1995
SCR	Bottom of chromosome 3, close to *BGL1*. Gene sequence available.	Di Laurenzio *et al.*, 1996
TTG	Top of chromosome 5.	Schneider *et al.*, 1997

* According with the CAPS procedure presented in Konieczny and Ausubel (1993)

division abnormalities suggest that the *ton* mutation affects the microtubules. However, root hair cells, phragmoplasts and spindles are normal in these plants, indicating that microtubules are normal (Traas *et al.*, 1995). Authors of such work suggest, however, that either *TON1* or *TON2* is an allele of the already described *FASS*

Plants bearing the *sabre* (*sab*) mutation have abnormally wide roots and stunted aerial organs. The increase in root diameter is the result of abnormal cell expansion. Cortical cells expand radially instead of longitudinally (Aeschbacher *et al.*, 1995; Benfey *et al.*, 1993). In these mutants the cells of the epidermis and endodermis look somewhat altered, but contrary to cortical cells their expansion axis are normal, along the longitudinal axis. The defect present in *sab* plants is reversible to some extent. In mutants grown in the presence of an inhibitor of ethylene synthesis or action, the radial expansion is reduced, but not eliminated. Nonetheless, the *sab* phenotype cannot be attributed to an excess of ethylene, because *i*) wild type plant grown with ethylene precursors increase the radial growth inespecifically, and *ii*) *sab* plants do not have ectopic root hairs as wild type plants with an excess of ethylene (Aeschbacher *et al.*, 1995).

There are several mutants that affect the size of roots by affecting the degree and axis of elongation of cells in different layers. *Cobra* (*cob*) and *lion's tail* (*lit*) mutants have also thick roots, but the aerial shoot is normal. The expansion of these mutant roots is mainly in the meristematic zone. The elongation and differentiation zones are fairly normal in thickness. The extent to which these mutants' roots are thickened depends on the sucrose concentration in the medium (Benfey *et al.*, 1993), and this effect is not due to the osmolar potential of the sucrose (Hauser, Morikami and Benfey, 1995). Root expansion in *cob* and *lit* is also temperature dependent, and in low temperature the roots are of a normal diameter. The extent of expansion is not homogeneous among the different cell layers. In the *cob* mutant the epidermis shows the greatest degree of expansion. Cell architecture is fairly normal, and because expanded cells are less elongated in the normal direction, the cell volume is not affected as dramatically as the linear expansion (Benfey *et al.*, 1993). On *lit* plants the stele is the region with the greatest expansion, but in this case the resulting cellular volume is smaller than the wild type, only one fourth of the normal (Hauser *et al.*, 1995). In contrast to the effect of other mutations, *cob* and *lit* do not cause significant abnormalities in cell number of epidermal, cortical or endodermal layers (Benfey *et al.*, 1993). *cob* and *lit* are not unique in their morphological effect. Other mutants, such as *pom-pom1* (*pom1*), *quill* (*qui*), *pom-pom2* (*pom2*), *cudgel* (*cud*), also have significantly altered cell elongation with respect to wild type. Like *cob*, in *qui* and *cud* mutants the epidermal cells are the ones with an altered polarity of expansion, but unlike *cob*, the resulting cell volume in these mutants is greater than in wild type plants. *pom1 and pom2* mutants also have epidermal and cortical cells with a larger volume than those of wild type. In *qui* and *pom2* mutants there is a reduced root caused by a reduction in cell elongation. In *cob*, *lit*, *pom1* and *cud* mutants there is also a reduction in the rate of cell division (Hauser *et al.*, 1995).

Some genes are important in root radial patterning by controlling the asymmetric divisions of initial cells. One of these genes is *SCARECROW*. This gene is expressed in the mature root endodermis of wild type plants. Its expression begins very close to, maybe inside, the cortex/endodermis initials (Di Laurenzio *et al.*, 1996; Scheres, 1996). The *SCR* product contains a basic domain reminiscent to that in bZIP transcription factors. Seedlings with the *scr* mutation are phenotypically distinct from the

wild type because they have a shorter root. Moreover, transverse root sections show that in such plants there is one cell layer missing between the pericycle and the epidermis. The only ground tissue layer of *scr* plants retains the casparian strip, which is a marker for endodermis. However, the same layer also has the arabinogalactan epitope, which is normally present both in epidermal and cortex cells. Therefore, it seems that the unique layer between the epidermis and the pericycle of *scr* mutants has a mixed nature presenting biochemical properties of both cortex and endodermis (Di Laurenzio *et al.*, 1996). It is important to stress that at least part of the biochemical differentiation between cortex and endodermis is still maintained in *scr* plants. Therefore the *SCR* gene is somehow involved in the asymmetric division of the cortex daughter, but not in cell differentiation.

An important question in morphogenesis is to what extent is organ formation determined by the rates and patterns of cell division. In order to evaluate the role of the genes controlling the cell cycle in root growth and development, expression of the mitotic cyclin *cyc1At* under the *cdc2aAt* promoter (which confers an expression in apical meristems) was tested in transgenic plants. The *cyc1At* gene is normally expressed only in cells that divide actively, like those in apical meristems or lateral root primordia. The resulting cyclin over-expression gives rise to plants with a marked increase in root growth, caused by an increased number of cells rather than an increase in overall cell size. Despite the abnormal excessive length of these trangenics, the overall root architecture, including the pattern of lateral root formation, is indistinguishable from wild type. This result indicates that the enhanced root growth by stimulation of cell division in meristems does not alter the organization of the root meristem or its size (Doerner *et al.*, 1996).

The rate of cell division has been experimentally manipulated by other non-genetic means. Root length can be increased by treating roots biochemically with glutathione (GSH). This compound causes a significant, dose-dependent increase in the mean root length, which is attributed to an increased number of cells. Treated roots also had a meristematic zone that extended beyond the apex into the elongation zone. A decrease in GSH by biochemical treatment causes a reduction of mitotic figures, although the root keeps its normal overall size. That GSH concentration affects cellular division, is further suggested by the high GSH levels present in actively dividing cells of the apical meristem. Inversely, the quiescent center shows lower GSH levels than neighboring cells (Sánchez-Fernández *et al.*, 1997).

Relatively few genes that affect root patterning or root cell fate have been identified and studied. Many of the mutants with altered root patterning and cellular differentiation also have altered cell size and number. Some mutants with altered radial patterning in roots are *pinocchio* (*pic*), *gollum* (*glm*) and *wooden leg* (*wol*) with short roots and retarded root growth. *pic* mutants lack a ground tissue layer, *glm* do not have a normal organization in the pericycle and vascular bundle, and *wol* have fewer vascular cells than wild type (Scheres *et al.*, 1995). Mutants 5905 are short, hairy and with a diameter larger than wild type due to both a radial swelling in all tissues and an increased cell number in epidermis, cortex, endodermis and pericycle layers. Plants of the 1767 line have hairy roots, trichoblasts are not always overlying two cortical cells, and overall cellular elongation is reduced. Epidermal, cortical and endodermal cells of 4792 primary roots are radially swollen and present an irregular size and shape. 7133 mutants are very similar to the previous but only their epidermal and cortical cells are swollen. Lastly, 7203 mutants lack the endodermal cell layer (Holding, McKenzie

and Coomber, 1994). It is important to say that all these swollen mutants also present anomalies in their aerial development. A deeper understanding of the mechanism leading to the abnormal phenotype will unravel the normal role of all these genes in root development.

Short-root (*shr*) mutants have particularly short roots, but their aerial parts are relatively normal. Besides being short, the root in *shr* mutants is determinate and it does not have the ability of continuous growth. These mutants also have an increased number of secondary roots that initiate from the hypocotyl. At the cellular level, *shr* plants do not posses the small, densely cytoplasmic cells that are present in the normal elongation and meristematic regions. Furthermore, these plants do not present the same number of cell rings as in roots of wild type plants. The endodermis ring is missing and the stele has fewer cells than in wild type. These results have been obtained by studying the presence of the casparian strip, normally associated with endodermal and some stele cells (Benfey *et al.*, 1993).

Probably, the best understood developmental pathway in roots is that leading to the formation of trichoblasts and atrichoblasts. Mutants in *ROOT HAIRLESS* (*RHL1*, *RHL2* and *RHL3*) genes have short roots and fewer root hairs than in wild type. The effects of *rhl* mutation are not restricted to the root, though, *rhl1* plants are dwarf, never produce flowers, show necrosis and die. Despite the evident dwarfism, the reduction in hair cells is not due to a decrease in cell number, and the different tissue layers are normally established in *rhl* plants. However, epidermal cell walls of these mutants are perfectly smooth without any sign of root hair initiation, indicating that the *RHL* genes are somehow involved in the trichoblast developmental pathway (Schneider *et al.*, 1997).

Roots of *ectopic root hair 1* (*erh1*) mutants are shorter, wider and more hairy than wild type. Despite the radially expanded roots and excess of hairs, cellular root architecture in *erh1* mutants is otherwise identical to wild type. As mentioned above, cells that will bear hairs are distinguished by their dense cytoplasm, but in *erh1* plants some cells overlying two cortical cells (putative future hair cells) do not contain dense cytoplasm, indicating developmental abnormalities (Schneider *et al.*, 1997).

glabra2 (*gl2*) mutants are characterized by the presence of hairs on essentially every root epidermal cell (Schiefelbein, Masucci and Wang, 1997), suggesting that the action of wild type *GL2* is to generate the hairless cellular fate. The *GL2* mRNA and its homeodomain product are preferentially expressed in those atrichoblasts within the meristematic and elongation regions of the root. With the use of GUS staining, it is possible to observe that the earliest *GL2* expressing cells are located outside the core meristem, approximately two or three cells away from the epidermal/lateral root cap initials (Masucci *et al.*, 1996). Furthermore, *GL2* expression is absent or highly reduced in developing trichoblasts, in mature root epidermal cells of either cell type, in the epidermis/lateral root cap initials, or in the other underlying rings of the *Arabidopsis* root. Hairs in *gl2* plants are normal at the morphological level irrespective of their relative position with respect to the underlying cortical cells. Despite the increase in number of hair cells, the total number of cortical and epidermal cell files is equal to those in wild type plants. No other abnormalities in root cell organization were identified in *gl2* mutants (Masucci *et al.*, 1996).

Like *gl2*, *transparent testa glabra* (*ttg*) mutants present root hairs in almost every epidermal cell (Schiefelbein *et al.*, 1997). Developing epidermal cells in such mutants exhibit the normal characteristics of

differentiating root hair cells despite their position. In the elongation region of the root, all epidermal cells display a delay in vacuolation that is normal in differentiating root hair cells of wild type plants. The phenotype of these mutants indicate that the wild type *TTG* gene normally influences the fate of atrichoblasts but does not affect the cellular organization (Galaway *et al.*, 1994). These results indicate that *TTG* and *GL2* act in similar pathways. Since the *ttg* and *gl2* mutants resemble the effect of ethylene/auxin overproduction, it has been suggested that the action mechanism of normal *TTG* and *GL2* products is to negatively regulate the ethylene/auxin pathway, very probably by inhibiting the *AXR2* gene pathway in the developing hairless epidermal cells (Masucci and Schiefelbein, 1996).

Genetic analyses of the genes involved in hair root formation have led to partial models for the genetic and hormonal control of root epidermis development. Briefly, it has been suggested that in atrichoblasts *TTG* and *GL2* are the primary genes transforming the positional information into biochemical pathways. One effector of those genes seems to be *RHD6*. At the same time, in the trichoblasts the ethylene and auxins are supposed to convey signals affecting genes *AUX1*, *ETR1*, and *AXR1* directly. The products of those genes would activate a biochemical pathway ending in the activation of *AXR2*. A proper interaction of signals in trichoblast and atrichoblast would determine the correct appearance of root hair cells in the root (Masucci and Schiefelbein, 1996; Schneider *et al.*, 1997)

Plants with the *constitutive triple response 1* (*ctr1*) mutation are small with short hypocotyls and root, compact inflorescence and a very reduced root system overall. Also, these mutants have ectopic root hairs in epidermal cells overlying only one cortical cell (Dolan *et al.*, 1994). But in contrast to *ttg* and *gl2* plants, *crt1* mutants have only a modest increase in the proportion of root hair cells. The *CTR1* gene encodes a RAF homologue that seems to regulate negatively the ethylene signal transduction pathway (Kieber *et al.*, 1993; Scheres, 1997). Nevertheless, the *ctr1* phenotype cannot be reverted by inhibitors of ethylene biosynthesis or action, suggesting rather that the defect of this mutant is localized somewhere in the ethylene signal transduction pathway. Now, because all alleles of *ctr1* are loss-of-function mutations, and the morphological effect resembles an increase in ethylene, the ethylene response pathway contains at least one negative regulator in wild type *Arabidopsis*. See Ecker (1995) for a preliminary genetic pathway of ethylene response.

The *ethylene overproduction* mutants (*eto1*, *eto2* and *eto3*) have inhibited root and hypocotyl elongation, and a radial swelling of both. If seedlings with the *eto* are grown with either inhibitors of ethylene biosynthesis or antagonists of ethylene action, the aberrant phenotype is no longer observed. The previous result implies that the *eto* phenotype is originated by a defect, in this case an increase, in the function of ethylene biosynthesis (Ecker, 1995).

Pasticcino mutants (*pas1*, *pas2* and *pas3*) are plants with stunted growth, deformed leaves and abnormal root development. While *pas1* and *pas3* have shorter primary roots and very few (if any) secondary roots compared to wild type, *pas2* plants have larger roots than wild type and an increased number of secondary roots. Differences in root length are not accompanied by changes in the radial root pattern, however. The phenotype of *pas* mutants is similar to that caused when there is an unbalanced cytokinin/auxin ratio in the medium. The *pas* phenotype is also very similar to the phenotype caused by *Corynebacterium fascians* or *Rhodococcus fascians* diseases that are correlated with an increased cytokinin concentration. These results suggest that mutations in the *PAS*

genes have altered cytokinin sensitivity. Finally, it is very probable that *pas1*, *pas2* and *pas3* form part of a common pathway as suggested by double-mutant phenotypes (Faure *et al.*, 1998).

There are other gene products that mediate signal transduction pathways that also have an effect on root development. Mutants in the *LONG HYPOCOTYL 5* (*HY5*) gene present excessive cell elongation in root hairs, widely spread lateral roots, and abnormal gravitropism. Interestingly, the normal *HY5* gene product is transported to the nucleus to modulate light stimuli signal transduction (Oyama, Shimura and Okada, 1997).

Finally, cosuppression lines of a MADS-box gene (*ANR1*) suggest that this gene is important for lateral root elongation in response to local nitrate levels in the soil. *ANR1* mRNA is only present in roots, and not in flowers, stems or leaves. Northern analyses suggest that *ANR1* expression is inducible by treatment with NO_3^-, independently of the presence of K^+ or PO_4^{3-}. In wild type plants, local NO_3^- treatment stimulates the development of lateral roots. This seems to be a response dependent of the presence of *ANR1*. However, a ubiquitous high concentration (>10mM) of NO_3^- inhibits lateral root growth, but through an *ANR1*-independent mechanism. Importantly, primary root elongation is insensitive to this treatment. In *ANR1* suppressed lines, but not in control plants, the presence of 10mM to 100mM of NO_3^- significantly inhibited lateral root growth (Zhang and Forde, 1998). The authors of this study propose that *ANR1* is a key gene for root morphological plasticity. This is the first MADS-box gene characterized up to now that is not important in floral morphology in *Arabidopsis* (Weigel, 1995).

LASER ABLATION

Recent laser cell ablation experiments are uncovering some cell interaction mechanisms important in determining cell fate in *Arabidopsis thaliana* root. After cells are killed with a precisely focused laser beam, dead cells are compressed and pushed towards the periphery of the root. The resulting vacant space is invaded by neighboring cells. When quiescent center cells are ablated, the cellular remains are displaced towards the root tip while cells of the proximal vascular bundle occupy the new available space (Van der Berg *et al.*, 1995). The replacement rate, however, is dependent on the number of quiescent center cells ablated (Van der Berg *et al.*, 1997). Invading cells transdifferentiate, at least partially, as evidenced by the disappearance of a vascular marker and the appearance of a root cap marker in the cells that occupy the position of the ablated cells.

As described above, each cortical initial divides asymmetrically to produce one cortex and one endodermis cell. Upon ablation of such initials, some cells of the pericycle invade the available space and continue to divide and generate the pericycle and cortical cell files. Pericycle cells are smaller than cortical cells, and more than one pericycle cell take the place of the ablated cortex initial. This yields a cortical ring with more than eight cells that is the typical number in wild type roots. The invading cells thereafter elongate and divide asymmetrically, like normal cortical initials do, forming the cortical and endodermal files. These new endodermal cells posses a casparian strip that is evidence of the transdifferentiation undergone by these cells once they are in their new location (Van der Berg *et al.*, 1995). In a similar experiment a neighboring cortical cell replaces an ablated epidermis initial, here too, the replacing cell behaves as a normal epidermis initial, growing and dividing anticlinally.

Daughters of cortical initials were ablated before their asymmetric division. If only one cell is ablated, the cortical initial continues to grow and divide while the daughter divides asymmetrically as in wild type. However, if three adjacent cortical daughters are ablated, the underlying middle initial still generates a daughter cell, but this is unable to divide asymmetrically to generate the cortex and endodermis files. Such experiments indicate that the meristem initial cells perpetuate the pre-existing cellular pattern (Van der Berg *et al.*, 1995).

Other cell ablation experiments have been performed to study the control of cell division versus differentiation in root meristems. Columella initials stop their normal growth and division if the touching quiescent center cell is ablated. Furthermore, the columella cells that stop dividing differentiate into daughter cells as suggested by the presence of starch granules specific of daughters and absent in initials (Van der Berg *et al.*, 1997). These results are confirmed in experiments performed in mutants that lack post-embryonic division, suggesting a relative independence of cellular differentiation and division. From these results, the authors suggest that the quiescent center does not directly regulate cell division, put primarily arrests cell differentiation.

Both of the cell ablation studies described in the previous paragraphs suggest that there is a high correlation between the position inside the root and the cellular fate. This relationship does not seem to be restricted to the early stages of development but seems to continue throughout the plant life. It seems that positional signals directing differentiation arise from mature cells to guide initials to their proper cell fate (Van der Berg *et al.*, 1995). But there seems to be another kind of signaling, that accomplished by the quiescent center to maintain touching cells in an undifferentiated (or initial) state (Van der Berg *et al.*, 1997).

FINAL REMARKS

A big gap in the study of development and pattern formation concerns the mechanisms and processes by which molecular signals are integrated and coordinated across cells and cell layers to yield the overall morphogenetic patterns of complex structures. The *Arabidopsis thaliana* root architecture is very suitable for studies on morphogenesis, because of its simple and highly stereotyped layered pattern and in some cases, fixed number of cells per layer. Also, the recent demonstration that in roots the fate of at least some cells is determined by the cell's relative position inside the root structure, provides experimental support for the so-called positional hypothesis (Wolpert, 1994; Wolpert, 1996). These results open the possibility of developing computational models of root morphogenesis, using advanced formal tools such as cellular automata (Ermentrout and Edelstein-Keshet, 1993) or L-systems (Prusinkiewicz and Lindenmayer, 1990). These tools have been widely used to model the *form* of plants, and such implementations assume, precisely, that the future state of an element is a function of the present state of itself and its contacting neighbors. Previously developed models only provide descriptive simulations of pattern formation, however. The cellular-level information available for *Arabidopsis* root already enables the formulation of models that may be used to explore the overall morphogenetic effects of different mechanisms and patterns of cell-cell interactions. For example, many models have assessed the role of positive and negative signals to orchestrate differentiation. These could be applied to roots. Also, the process of cellular communication may be elicited by direct cell-cell contact (Slack, 1993), or by the diffusion of signaling molecules (Kerszberg and Wolpert, 1998; Wolpert, 1969; Wolpert, 1996). The kind of results provided by the laser ablation experiments in *Arabidopsis* regarding cellular signaling for differentiation, makes the *Arabidopsis* root a nice biological system suitable for a joint

theoretical and experimental effort to unravel the cellular and molecular mechanism of morphogenesis. But for roots we still need much more molecular and genetic data to formulate models of the regulatory networks within cells. This type of models have already been developed for genes underlying *Arabidopsis* flower development (Mendoza and Alvarez-Buylla, 1998).

The integrative analysis of many developmental mutants in *Arabidopsis* is enabling the formulation of morphogenetic models in plants. The ABC model for flower morphogenesis, despite its simplicity, has been very robust for describing the overall organ type organization of wild type, mutant and transgenic flower phenotypes (Coen and Meyerowitz, 1991; Ma, 1994; Meyerowitz, 1994, Krizek and Meyerowitz, 1996; Mizukami and Ma, 1992). There are propositions to modify the ABC model (Ma, 1994; Ray *et al.*, 1994), but they suggest only minor changes. Is the *Arabidopsis* root architecture also ruled by an ABC-type model? In order to establish that kind of model, it is necessary to look for homeotic mutants. At the organ level, the possibilities are few in roots, however. Would it be possible to obtain a shoot instead of a root? Or maybe a secondary instead of a primary root? Other possible helpful mutants might be those in which the concentric rings of different cell types are interchanged by altering the cell fates of different cell layers. For example, one plant with an extra normal cortical ring taking the place of the endodermis. As more mutant and molecular data of root development accumulate, we will be able to address these and other questions.

LITERATURE CITED

Aeschbacher RA, Hauser MT, Feldmann KA, Benfey PN. 1995. The *SABRE* gene is required for normal cell expansion in *Arabidopsis*. *Genes and Development*. **9**: 330-340.

Baum SF, Rost TL. 1996. Root apical organization in *Arabidopsis thaliana* 1. Root cap and protoderm. *Protoplasma* **192**: 178-188.

Beemster GTS, Baskin TI. 1998. Analysis of cell division and elongation underlying the developmental acceleration of root growth in *Arabidopsis thaliana*. *Plant Physiology* **116**: 1515-1526.

Benfey PN, Linstead PJ, Roberts K, Schiefelbein JW, Hauser MT, Aeschbacher RA. 1993. Root development in *Arabidopsis*: four mutants with dramatically altered root morphogenesis. *Development* **119**: 57-70.

Berleth T, Jürgens G. 1993. The role of the *monopteros* gene in organising the basal body region of the *Arabidopsis* embryo. *Development* **118**: 575-587.

Cheng JC, Seeley KA, Sung ZR. 1995. *RML1* and *RML2*, *Arabidopsis* genes required for cell proliferation at the root tip. *Plant Physiology* **107**: 365-376.

Coen E. 1991. The role of homeotic genes in flower development and evolution. *Annual Review of Plant Physiology and Plant Molecular Biology* **42**: 241-279.

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

Di Laurenzio L, Wysocka-Diller J, Malamy JE, Pysh L, Helariutta Y, Freshour G, Hahn MG, Feldmann KA, Benfey PN. 1996. The *SCARECROW* gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root. *Cell* **86**: 423-433.

Doerner P, Jørgensen JE, You R, Steppuhn J, Lamb C. 1996. Control of root growth and development by cyclin expression. *Nature* **380**: 520-523.

Dolan L, Duckett CM, Grierson C, Linstead P, Schneider K, Lawson E, Dean C, Poethig S, Roberts K. 1994. Clonal relationships and cell patterning in the root epidermis of *Arabidopsis*. *Development* **120**: 2465-2474.

Dolan L, Janmaat K, Willemsen V, Linstead P, Poethig S, Roberts K, Scheres B. 1993. Cellular organisation of the

Arabidopsis thaliana root. *Development* **119**: 71-84.

Driever W, Nüsslein-Volhard C. 1988. The bicoid protein determines position in the *Drosophila* embryo in a concentration dependent manner. *Cell* **54**: 95-104.

Duckett CM, Oparka KJ, Prior DAM, Dolan L, Roberts K. 1994. Dye-coupling in the root epidermis of *Arabidopsis* is progressively reduced during development. *Development* **120**: 3247-3255.

Ecker JR. 1995. The ethylene signal transduction pathway in plants. *Science* **268**: 667-675.

Ermentrout GB, Edelstein-Keshet L. 1993. Cellular automata approaches to biological modeling. *Journal of Theoretical Biology* **160**: 97-133.

Faure JD, Vittorioso P, Santoni V, Fraisier V, Prinsen E, Barlier I, Van Onckelen H, Caboche M, Bellini C. 1998. The *PASTICCINO* genes of *Arabidopsis thaliana* are involved in the control of cell division and differentiation. *Development* **125**: 901-918.

Galway ME, Masucci JD, Lloyd AM, Walbot V, Davis RW, Schiefelbein JW. 1994. The *TTG* gene is required to specify epidermal cell fate and cell patterning in the *Arabidopsis* root. *Developmental Biology* **166**: 740-754.

Golberg RB, De Pavia G, Yadegari R. 1994. Plant embryogenesis: zygote to seed. *Science* **266**: 605-614.

Hardtke CS, Berleth T. 1998. The *Arabidopsis* gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development. *EMBO Journal* **17**: 1405-1411.

Hauser MT, Morikami A, Benfey PN. 1995. Conditional root expansion mutants of *Arabidopsis*. *Development* **121**: 1237-1252.

Holding DR, McKenzie RJ, Coomber SA. 1994. Genetic and structural analysis of five *Arabidopsis* mutants with abnormal root morphology generated by the seed transformation method. *Annals of Botany* **74**: 193-204.

Jürgens G, Mayer U, Torrez-Ruiz RA, Berleth T, Miserá S. 1991. Genetic analysis of pattern formation in the *Arabidopsis* embryo. *Development* **Supplement 1**: 27-38.

Kerszberg M, Wolpert L. 1998. Mechanisms for positional signalling by morphogen transport: a theoretical study. *Journal of Theoretical Biology* **191**: 103-114.

Kieber JJ, Rothenberg M, Roman G, Feldmann KA, Ecker JR. 1993. *CTR1*, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the Raf family of protein kineses. *Cell* **72**: 427-441.

Konieczny A, Ausbel FM. 1993. A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant Journal* **4**: 403-410

Krizek BA, Meyerowitz EM. 1996. The *Arabidopsis* homeotic genes *APETALA3* and *PISTILLATA* are sufficient to provide the B class organ identity function. *Development* **122**: 11-22.

Laskowski MJ, Williams ME, Nusbaum HC, Sussex IM. 1995. Formation of lateral root meristem is a two-stage process. *Development* **121**: 3303-3310.

Laux T, Jürgens G. 1997. Embryogenesis: a new start in life. *Plant Cell* **9**: 989-1000.

Ma H. 1994. The unfolding drama of flower development: recent results from genetic and molecular analyses. *Genes and Development* **8**: 745-756.

Malamy JE, Benfey PN. 1997a. Organization and cell differentiation in lateral roots of *Arabidopsis thaliana*. *Development* **124**: 33-44.

Malamy JE, Benfey PN. 1997b. Down and out in *Arabidopsis*: the formation of lateral roots. *Trends in Plant Sciences* **2**: 390-396.

Masucci JD, Rerie WG, Foreman DR, Zhang M, Galway ME, Marks MD, Schiefelbein JW. 1996. The homeobox *GLABRA2* is required for position-dependent cell differentiation in the root epidermis of

Arabidopsis thaliana. *Development* **122**: 1253-1260.

Masucci JD, Schiefelbein JW. 1996. Hormones act dowmstream of TTG and GL2 to promote root hair outgrowth during epidermis development in the *Arabidopsis* root. *Plant Cell* **8**: 1505-1517.

Mayer U, Büttner G, Jürgens G. 1993. Apical-basal pattern formation in the *Arabidopsis* embryo: studies on the role of the *gnom* gene. *Development* **117**: 149-162.

Mendoza L, Alvarez-Buylla ER. 1998. Dynamics of the genetic regulatory network for flowering in *Arabidopsis thaliana*. *Journal of Theoretical Biology* **193**: 307-319.

Meyerowitz EM. 1994. The genetics of flower development. *Scientific American* **271**: 40-47.

Meyerowitz EM. 1997. Plants and the logic of development. *Genetics* **145**: 5-9.

Mizukami Y, Ma H. 1992. Ectopic expression of the floral homeotic gene *AGAMOUS* in transgenic *Arabidopsis* plants alters floral organ identity. *Cell* **71**: 119-131.

Oyama T, Shimura Y, Okada K. 1997. The *Arabidopsis* HY5 gene encodes a bZIP protein that regulates stimulus-induced development of root and hypocotyl. *Genes and Development* **11**: 2983-2995.

Prusinkiewicz P, Lindermayer A. 1990. *The algorithmic beauty of plants*. New York: Springer-Verlag.

Ray A, Robinson-Beers K, Ray S, Baker SC, Lang JD, Preuss D, Milligan SB, Gasser CS. 1994. *Arabidopsis* floral homeotic gene *BELL* (*BEL1*) controls ovule development through negative regulation of *AGAMOUS* gene (*AG*). *Proceeding of the National Academy of Sciences USA* **91**: 5761-5765.

Rerie WG, Feldmann KA, Marks MD. 1994. The *GLABRA2* gene encodes a homeodomain protein required for normal trichome development in *Arabidopsis*. *Genes and Development* **8**: 1388-1399.

Rost TL, Baum SF, Nichol S. 1996. Root apical organization in *Arabidopsis thaliana* ecotype 'WS' and a comment on root cap structure. *Plant and Soil* **187**: 91-95.

Sánchez-Fernández R, Fricker M, Corben LB, White NS, Sheard N, Leave, CJ, Van Montagu M, Inzé D, May MJ. 1997. Cell proliferation and hair tip growth in the *Arabidopsis* root are under mechanistically different forms of redox control. *Proceedings of the National Academy of Sciences USA* **94**: 2745-2750.

Scheres B. 1996. Embryo patterning genes and reinforcement cues determine cell fate in the *Arabidopsis thaliana* root. *Seminars in Cell and Developmental Biology* **7**: 857-865.

Scheres B. 1997. Cell signaling in root development. *Current Opinion in Genetics and Development* **7**: 501-506.

Scheres B, Wolkenfelt H, Willemsen V, Terlouw M, Lawson E, Dean C, Weisbeek P. 1994. Embryonic origin of the *Arabidopsis* primary root and root meristem initials. *Development* **120**: 2475-2487.

Scheres B, Di Laurenzio L, Willemsen V, Hauser MT, Janmaat K, Weisbeek P, Benfey PN. 1995. Mutations affecting the radial organization of the *Arabidopsis* root display specific defects throuhout the embryonic axis. *Development* **121**: 53-62.

Schiefelbein JW, Masucci JD, Wang H. 1997. Building a root: the control of patterning and morphogenesis during root development. *Plant Cell* **9**: 1089-1098.

Schneider K, Wells B, Dolan L, Roberts K. 1997. Structural and genetic analysis of epidermal cell differentiation in *Arabidopsis* primary roots. *Development* **124**: 1789-1798.

Slack J. 1993. Embryonic induction. *Mechanisms of Development* **41**: 91-107.

Sussex IM, Godoy JA, Merk NM, Laskoski MJ, Nusbaum HC, Welsch JA, Williams ME. 1995. Cellular and molecular events in a newly organizing lateral root meristem. *Philosophical Transactions of the Royal Society of London, Series B* **350**: 39-43.

Theiβen G, Saedler H. 1995. MADS-box genes in plant ontogeny and phylogeny: Haekel's "biogenetic law" revisited. *Current*

Opinion in Genetics and Development **5**: 628-639.

Torres-Ruiz RA, Jürgens G. 1994. Mutations in the *FASS* gene uncouple pattern formation and morphogenesis in *Arabidopsis* development. *Development* **120**: 2967-2978.

Traas J, Bellini C, Nacry P, Kronenberger J, Bouchez D, and Caboche M. 1995. Normal differentiation patterns in plants lacking microtubular preprophase bands. *Nature* **375**: 676-677.

Trewavas AJ, Malhó R. 1997. Signal perception and transduction: the origin of the phenotype. *Plant Cell* **9**: 1181-1195.

Van der Berg C, Willemsen V, Hage W, Weisbeek P, Scheres B. 1995. Cell fate in the *Arabidopsis* root meristem determined by directional signalling. *Nature* **378**: 62-65.

Van der Berg C, Willemsen V, Hendriks G, Weisbeck P, Scheres B. 1997. Short-range control of cell differentiation in the *Arabidopsis* root meristem. *Nature* **390**: 287-289.

Weigel D. 1995. The genetics of flower development: from floral induction to ovule morphogenesis. *Annual Review of Genetics* **29**: 19-39.

Willemsen V, Wolkenfelt H, De Vrieze G, Weisbeek P, Scheres B. 1998. The *HOBBIT* gene is required for formation of the root meristem in the *Arabidopsis* embryo. *Development* **125**: 521-531.

Wolpert L. 1969. Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology* **25**: 1-47.

Wolpert L. 1994. The evolutionary origin of development: cycles, patterning, privilege and continuity. *Development* **1994 Supplement:** 79-84.

Wolpert L. 1996. One hundred years of positional information. *Trends in Genetics* **12**: 359-364.

Yanofsky MF. 1995. Floral meristems to floral organs: genes controlling early events in *Arabidopsis* flower development. *Annual Review of Plant Physiology and Plant Molecular Biology* **46**: 167-188.

Zhang H, Forde BG. 1998. An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. *Science* **279**: 407-409.

APENDICE 2

DISCRETE NETWORKS AS A SUITABLE APPROACH FOR THE ANALYSIS OF GENETIC REGULATION

LUIS MENDOZA

Lab. de Genética Molecular y Evolución. Instituto de Ecología, UNAM.
Ciudad Universitaria. Ap. Postal 70-275. Coyoacán, D.F. CP04510. MEXICO
lams@servidor.unam.mx

ABSTRACT

Biological systems are composed of multiple interacting elements; in particular, genetic regulatory networks are formed by genes and their interactions mediated by transcription factors. The establishment of such networks is critical to guarantee the reliability of transcriptional performance in any organism. Whole-genome sequencing efforts have created a growth in those areas related to genetic information, increasing the necessity for integrative tools. The study of genetic regulatory networks as dynamical systems is a helpful methodology to understand the transcriptional behavior of the genome. From a number of theoretical studies, it is known that networks present a complex dynamical behavior that includes stability, redundancy, homeostasis and multistationarity. In this paper, I present some particular biological processes modeled as discrete networks to show that the previously mentioned theoretical properties have a clear biological interpretation. In the discussed examples, the stress relies on the predictions and novel interpretations, rather than on the biological data itself.

INTRODUCTION

Development of multicellular organisms requires the coordinated accomplishment of many molecular and cellular processes, like division and differentiation. Regulation of those processes must be very reliable, capable of resisting fluctuations of the internal and external environment. Without such homeostatic capacity, the viability of the organism would be compromised. For instance, unrestrained division of some cells may lead to the appearance of tumors, with the possible result of the organism death. Cellular processes are finely controlled by a number of regulatory molecules. In particular, transcription factors are proteins that determine the transcription rate of genes, including those involved in development and morphogenesis. Since transcription factors are present inside cells at a low quantity, a random variation in their concentration might alter cellular fate. For such reason, it is reasonable to look for the molecular mechanisms responsible of the homeostatic capacity, and coordinated behavior of the transcriptional machinery.

Modern techniques of molecular genetics have greatly increased the rate at which genes are recognized and their primary sequences determined. Still, classic biochemical and physiological studies are necessary to identify the targets, and to understand the functions, of the coded proteins. For such reasons, the rate at which pathways are described is much slower than the rate of gene sequencing. The large quantity of available sequences creates the challenge for molecular geneticists of linking genes and proteins into functional pathways. Of particular interest for this paper, are those cases when the discovered genes code for transcription factors. These proteins bind to *cis*-regulatory sequences of other genes, and if their targets also code for other transcription factors, then interdependence is created among groups of genes, thus forming

a genetic regulatory network. The observed molecular result is the controlled and coordinated expression of a large group of genes. In this way, genetic regulatory networks control the transcriptional behavior of the cells in an organism. These ideas are commonly accepted by experimental biologists; however, geneticists and molecular biologists generally are not aware of the global dynamical properties of genetic regulatory networks. This problem might be partially solved by presenting models of some particular biological networks, and by showing the descriptive and predictive capacity of them. Network models contain some dynamical properties that are not evident from a global visual inspection of the system connectivity. Thus, it is possible to increase the use of network modeling by stressing some novel results obtained from the interpretation of well-known biological processes as dynamical systems.

One relevant biological problem that may be partially understood by making use of network models, is how different stable genetic patterns arise inside one organism. All cells in an individual organism have the same genes, and therefore the same global genetic regulatory network. However, each cellular type of an organism differs from others in its particular molecular profile, *i.e.* in its pattern of transcriptionally active genes and the presence of other molecular markers. In addition, those genetic activation patterns are stable, in a normal situation cells do not differentiate continually from one type into another. In an adult organism, many stable genetic activation patterns coexist despite the common underlying network in every cell. This characteristic is known as multistationarity. Importantly, it turns out that some particular networks provide the molecular mechanism of multistationarity.

To describe the genetic circuitry, it is a common practice to represent transcriptional regulatory interactions with arrows (A→B) or blunt lines (A⊣B), to indicate activation or repression respectively. Such graphical representation is very helpful to understand the transcriptional machinery of an organism. However, the knowledge of the connectivity is not enough to determine the behavior of a regulatory network. There are properties that cannot be deduced just with the aid of a visual inspection of the circuitry. For example, it is not possible to know how many steady states of genetic activation are allowed by a particular network, neither if those steady states are stable or not. To know these properties, it is necessary to incorporate the transcription rate of each gene as a function of its regulators. In this way, a genetic regulatory network can be translated into a dynamical system. To model a genetic regulatory network, the tendency would be to choose a series of coupled differential equations, with each dependent variable representing the transcriptional activity of one gene. Most of the times, however, there is a lack of quantitative experimental information to fix the set of parameters in the system of equations. In contrast, most published results include qualitative information of the spatio-temporal activation patterns of genes. Thus, there is usually enough information to describe a genetic regulatory network as a system of difference, instead of differential, equations.

It might appear that a discrete approach is somehow inferior to continuos modeling, but it has been shown that continuos and discrete network models share many qualitative dynamic features (Bagley and Glass, 1996; Glass, 1975; Glass and Kauffman, 1973). Also, it is known that the modeling of genetic regulatory network as dynamical systems with discrete variables is adequate, because such systems present global complex behavior including self-organization, stability, redundancy and periodicity (Somogyi and Sniegoski, 1996). Moreover, the presence of feedback loops permits the possibility to obtain homeostasis and multistationarity (Thieffry *et al.*, 1995; Thomas *et al.*, 1995). In contrast, classical hierarchical and static models do not present the rich behavior of networks. In this mini-review, I present some global properties of networks, and in another section some

biological examples of regulatory genetic networks modeled as discrete state dynamical systems. Those models permit to show that genetic regulatory networks have the capacity of forming a set of particular genetic expression patterns.

DISCRETE NETWORKS

Basic concepts

Most biological systems are composed of multiple interacting elements. For example, consider the genes and proteins in a genetic regulatory network, neurons in a nervous system, lymphocytes in the immune system and so on. The molecular response in those systems often involves cooperation, saturation, competition, etc., establishing a non-linear relationship between the input and output. Such non-linear responses of the network elements result in the impossibility of predicting the network behavior just by knowing the response of its isolated elements. Therefore, in a living organism it is necessary to consider the dynamics of each element as part of an entire functional network, rather than isolated entities. To elaborate models that capture the qualitative essence of the system under study it is necessary to consider some simplifications, because of the complexity of the biological regulatory networks.

A common hypothesis in network modeling is that events occur at certain time intervals, rather than continuously. In this case, the state of the system at a particular time depends only on its own state at the preceding time step. Mathematically speaking, the supposition is that the dynamical behavior of the system is governed by an equation of the form $S_{t+1}=f(S_t)$. Here, S represents the state of the system under study, for example the activation pattern of a group of genes. The system behavior is thus expressed by a succession of states, namely $S_1, S_2, S_3 \ldots S_n$. Often, the function f governing the change from one time step to another is unknown. If the function is known, the dynamical behavior of the system can be obtained by an iterative process. In most cases, however, f turns out to be a nonlinear function, making its analysis a hard task.

A network is a system formed by multiple elements, or *nodes*. Each node has a *state of activation*, which depends upon the states of another nodes. To describe the activity of all elements, it is necessary to express each node as a dependent variable in a system of coupled equations. When modeling a genetic regulatory network, the activity of a node represents the transcriptional rate (or *activity*) of a particular gene. A natural tendency would be to represent nodes with continuos variables. However, usually there is only qualitative experimental data regarding the activity of most genes. For such reason, a useful simplification it is to suppose that genes can attain only a finite number of possible states. Thus, it is a common practice to represent the transcriptional rate of a gene with a discrete variable. In the simplest case, a gene might be "turned on" or "turned off" at a given time. In this case, we are dealing with *switching* networks, *boolean* networks, or more properly *binary* networks. Despite the gross simplifying assumptions of discrete time and discrete activity states, the dynamical properties of binary networks are remarkably complex.

To describe the collective behavior of the elements in a network, it is necessary to use the concept of the *state of the network*. Such state is represented by a vector of the form $(x_1, x_2 \ldots x_n)$, which contains the activation state of genes $X_1, X_2 \ldots X_n$. By convention, in binary networks the values '0' and '1' represent transcriptionally inactive and active genes, respectively. Therefore, the state of a binary network is completely characterized by a binary vector, *e.g.* (0,1,1,0,1); although often the shorthand without parentheses or commas are used instead, *e.g.* 01101. Now, since each node in a binary network can attain two states, there are 2^n possible states for a network of size n. Or in more general terms, there are m^n states, where m is the number of discrete states permitted for each node. The collection of all possible activation states forms the *state space* of the network.

An *autonomous* network does not have any input from the outside, and its behavior depends only upon the activation state of its constituent nodes. In such system, each network state at time t determines a unique successor at time $t+1$; it is a *deterministic* system. Over a series of time steps the network passes through a succession of states, thus forming a *trajectory*. The trajectory cannot be infinite because the state space is itself finite. Therefore, the system must eventually repeat a previously visited network state, thus cycling repeatedly around a number of recurrent states. This set of states that are cyclically repeated in a trajectory constitutes an *attractor*. An attractor can be characterized by its period, if the same network state is repeated after m time steps, then it is a period-m attractor. The period of an attractor has a range from 1 to 2^n, where n is the number of nodes in the binary network. Period-1 attractors are also known as *fixed-point* attractors.

Every autonomous finite-state network has at least one attractor, in such case any network trajectory flows into the same attractor. But a network may have more than one attractor; some states may lead to a fixed-point attractor, for example, while others may lead to a period-3 attractor. The collection of all network states leading, or belonging to a particular attractor is called the *basin of attraction*. Thus, the state space on the network is partitioned among its basins of attraction, each draining towards an attractor.

Some properties of discrete networks

It was already mentioned that each network reaches one of the possible attractors draining the state space. Such characteristic confers stability to the networks; which means the possibility to resists some perturbations. To illustrate the property, take the following trajectory of a binary network: 1110 → 1101 → 0000 → 1111. Suppose that the four network states form the whole basin of attraction of the fixed-point attractor 1111. Once the network reaches the attractor, the network state will be 1111 for as long as no perturbations exist. Now, suppose that an external stimulus changes the activation state of the fourth network element from 1 to 0; in that case the new network state would be 1110. It turns out that such state is part of the basin of attraction of 1111; thus, after three time-steps the system returns to the original activation steady state. Notice that a similar situation holds if the third network element is perturbed. Moreover, if all the elements are changed at once, turning the network state from 1111 to 0000, it takes only one time-step for the network to restore its normal state. Therefore, the existence of basins of attractions guarantees that the effect of a number of perturbations will die out after a transient response. Of course, not all perturbations have the same transitory effect. In the previous example, an alteration of the attractor on the first element results in the state 0111, which is outside the depicted basin of attraction. Such new activation state lies necessarily in another basin of attraction; thus, the network will follow a trajectory ending in an attractor different to the original 1111. Therefore, some perturbations may result in a change in the final stable state attained by the network. If a perturbation originates a shift from one attractor to other or not, depends on the particular partition of the state space. In any case, the existence of attractors confers stability and resistance to some perturbations, which are fundamental qualities observed in biological systems.

Genetic regulatory networks present redundancy, the removal of an element may not necessarily affect the global behavior of the system. Redundancy is more easily observed in large systems, which have many elements and abundant interactions. This characteristic is relevant because genetic redundancy is present in biological systems. Experimentally, there are some cases where null mutations do not have phenotypic effects by themselves, but enhance the effects of mutations in other genes (Kempin *et al.*, 1994). Normally, it is assumed that gene duplications, and further point mutations, cause the appearance of similar genes that are partially redundant. An excellent example is the HOM/Hox gene clusters, in which gene duplications and deletions occurred during evolution (for a review see Ruddle *et al.*,

1994), and they present partial redundancy (Horan *et al.*, 1995). However, contrary to what is intuitively expected, network simulations show that the amount of redundancy is not necessarily proportional to the similarity between the coding regions of the original and duplicate gene pair (Wagner, 1996a; 1996b). Even more, genes with redundant or partial overlapping functions but with low sequence similarity are known (Cooke *et al.*, 1997). Redundancy thus seems to be a global property of networks, rather than a particular property of the primary sequence of genes involved in regulatory pathways. In the next section, there are a few examples of networks whose dynamical properties depend only on a subset of their elements.

Regulatory networks may or may not include feedback circuits. Their presence is necessary to ensure multistationarity and homeostasis, which are particularly important properties of biological systems. The logical analysis of feedback loops decomposes any network into a well-defined set of feedback loops; it was first developed by Thomas (1978), and formally demonstrated by others (Plahte *et al.*, 1995; Gouzé, 1998; Snoussi, 1998). According to such methodology, a feedback loop is positive or negative if the involved negative interactions are even or odd, respectively. On the one hand, negative feedback loops generate homeostasis, in the form of damped or sustained oscillations. The importance of homeostasis in maintaining the internal ambient of an organism is well known, and dates back from the work of Cannon (1929). On the other hand, positive feedback loops generate multiple alternative steady states; n independent feedback loops can generate up to 3^n steady states, of which 2^n are stable. The already discussed biological interpretation of multistationarity as cellular differentiation goes back to Delbrück (1949), but has been further developed by the group of Thomas (Thieffry *et al.*, 1995; Thomas *et al.*, 1995).

As noticed, the connection of many simple elements results in the appearance of complex nonlinear dynamics. Such behavior rarely permits the analytical study, and thorough description, of all properties of network models. To solve the problem, a common approach is the statistical study of randomly constructed networks. As a result, there is a large collection of results that are particular of the type of network used (for example, see Kauffman, 1993). Nevertheless, there are some specific properties that are directly relevant for the study of biological systems. In the previous paragraphs I mentioned four: stability, redundancy, homeostasis and multistationarity. Those properties make of the network modeling a suitable tool to describe and predict the characteristics of genetic regulatory networks. To give validity to the previous claim, in the following section I describe some examples of the use of discrete-state networks to model particular biological systems.

DISCRETE NETWORK MODELS OF SOME BIOLOGICAL SYSTEMS

Classical schemes of regulatory cascades make explicit the elements and interactions conforming a pathway. However, they almost never incorporate a description of how the presence of the regulatory elements determines the activity of genes or proteins. Because of this absence, schemes presenting cascades of regulation are not dynamical systems. As a result, it is not possible to evaluate how many steady activation patterns (attractors) are allowed by the regulatory pathway, neither if those patterns are stable or unstable. The intention of the following recount is not to make a thorough description of many models, but to stress some of the novel results and interpretations emanated from the network modeling approach. Since the concept of attractor is of central importance in network modeling, the following examples stress the biological interpretation of different attractors. The area of discrete network modeling of regulatory pathways is in development; therefore, this review covers most of the published examples, but is by no means complete.

The lambda phage is a virus of *E. coli* that can integrate its genetic material into the

host DNA or multiply in the cytoplasm. Many bacterial and viral genes take part in this alternative between the so-called lytic or lysogenic pathways. Thieffry and Thomas (1995) elaborated a regulatory network incorporating the interactions among genes cI, cII, cro and N of the lambda bacteriophage. Those four genes control the choice if the host bacterium will lyse or become lysogenic, a process somewhat similar to cellular differentiation. Despite that the complete genome sequence of the lambda phage is well known, modeling of the dynamical behavior emphasize aspects that are not appreciated intuitively. First, the model shows that the cI and cI-cro circuits are sufficient to obtain the change of stable expression patterns induced by a change of temperature. Specifically, the model has two stable expression patters at low temperature, but only one at a high temperature. Second, the circuit analysis of the model shows that the inclusion of cI-cII and cI-N-cII negative loops, as well as the cI-cro-cII positive loop, results in the increase of cooperativity but do not play a crucial role in the decision for or against immunity. Of course, this model is capable of reproducing the effect of many known mutations, but an important aspect is the establishment of some unexpected results. For example, authors stress the importance of the negative self-regulation of the cro operon in diminishing the effect of gene dosage while the phage is replicating at high rates. In brief, this model reinterprets and predicts the role of positive and negative feedback loops during the infection process of the lambda phage.

Genetic and molecular studies on *Drosophila melanogaster* have shown the existence of gradient mechanisms in the generation of embryonic patterns. Sánchez *et al.* (1997) published a regulatory network constituted by genes *dorsal*, *twist*, *snail*, *decapentaplegic*, *sog*, *toll*, and *rhomboid*, to explain the establishment of the dorso-ventral genetic pattern of *Drosophila*. However, it was found that the consideration of only the first four of the mentioned genes was enough to elaborate a discrete model. The model contains two feedback loops, which generate five steady states. Three of them are stable steady states, which correspond to the genetic expression found in the ectoderm, neuroectoderm and mesoderm. Apart from an adequate description of transient and stable activation pattern, the model also describes the effect of single-gene mutations. Moreover, the model permits to establish that from a formal point of view, the putative autoregulatory function of *snail* is dispensable for the establishment of the dorso-ventral pattern. This result is not evident without the formal analysis of the network model.

Genes that establish the antero-posterior pattern of *Drosophila* also have received attention. Burstein (1995) elaborated a network model incorporating maternal, gap, pair-rule, segment polarity, and homeotic genes. Even though the network contains 16 elements, the author was interested in the dynamics of the *Deformed* gene alone. In brief, the model was developed to establish the requirements of Bicoid, Hunchback and Even-skipped protein concentrations to establish the *Deformed* striped expression pattern. Burstein's model helps to understand how relatively few, broadly expressed gap genes specify organized stripes of downstream genes. The model suggests that the mechanism depends upon a combination of overlapping patterns and gradient concentrations of gap proteins.

Flower development has demonstrated to be a suitable system for network modeling. Mendoza and Alvarez-Buylla (1998) developed a genetic regulatory network with 11 genes that control the flower morphogenesis of *Arabidopsis thaliana*. The network model results in the formation of six attractors. Four attractors correspond to the genetic activity observed in the four floral organs, *i.e.* sepals, petals, stamens and carpels. A fifth attractor, in turn, represents the genetic activity of meristems not competent to initiate a flower; it is a non-flowering state. Finally, the model predicts the existence of a sixth attractor constituted by the activation of two non-flowering genes and two involved in flower development.

Notably, the model posed some predictions regarding the existence of some genetic interactions. From those, the activation of *AG* by *LFY* has been confirmed experimentally (Parcy *et al.*, 1998). Moreover, there is preliminary data confirming the predicted inhibition of *API* by *EMFI* (Aubert, 1998). A subsequent logical analysis of the network model (Mendoza *et al.*, 1999) showed that only two feedback loops, *i.e. API-AG* and *AP3-PI*, are sufficient to obtain the six attractors already mentioned. Also, the analysis predicts the existence of a yet undiscovered activator of the gene *LFY*.

The differentiation process in the root epidermis of *Arabidopsis* also has been subject of modeling. There is a model (Mendoza and Alvarez-Buylla; submitted) that incorporates the genetic regulation and signal transduction pathways leading to the development of root hairs. The dynamic system incorporates eight elements, including transcription factors and signaling transduction proteins. Interestingly, the initial state of two variables represents the ethylene availability of cells, and another uncharacterized signal coming from the root cortex. The particular combination of the two signals determines the attractor reached by the network. Each attractor represents a genetic expression pattern leading to the appearance of different number of root hairs. Normally, alternate files of hair and non-hair cells form the root epidermis of *Arabidopsis*, but there are multiple mutants altering the number and distribution of hairs. An important difference of this model with other previous non-network models for the root epidermis, is its capacity to describe and predict the morphological effects of single and multiple mutations, as well as the response to some pharmacological treatments.

Network elements not necessarily represent genes, proteins or other single molecules. Muraille *et al.* (1996) elaborated a network model to study the neuroendocrine regulation of the immune response. The model contains only four elements representing a pathogen, the immune response, the hypothalamo-pituitary-adrenocortical axis, and the host organism itself. With the use of a logical analysis, the authors showed that the model contains six feedback loops, functional in some regions of the variable space. Such analysis also revealed the existence of ten steady states. Seven of the steady states are characteristic of the feedback loops, one is the zero state, one represents the state where the pathogen is dead, and finally one represents the organism's death. In addition, the discrete model was used to construct a more refined differential model. Briefly, the study of the discrete and continuos models accounts for some classical and peculiar immune responses, particularly immunogenicity, toxicity, neuro-hormonal feedback, toxic shock syndrome, the relation of pathogen with infection, and the relation of stress with the immune response. In Muraille's work, the elaboration of a discrete network helped to locate and identify the nature of steady states in a more elaborated continuos model. Moreover, the logical approach also was used to predict some dynamical changes as the loss of multistationarity or homeostasis.

Without doubt, the most elaborated models of genetic regulatory networks are made for *Drosophila* early embryogenesis. In particular, Bodnar (1997) presented an extraordinary spatio-temporal discrete model that integrates the genetic and nuclear events from the egg to the syncytial blastoderm, comprising 13 nuclear divisions. In the model all genes and protein products have four possible states, while the rules controlling gene's switching depend upon the protein and chromatin states. Briefly, maternal effects are simulated as the initial state of the system, and the subsequent genetic expression establishes protein step gradients that activate o repress other genes in neighbor nuclei. A network of 11 genes was sufficient to form the 16 compartments corresponding to the parasegments formed along the antero-posterior axis of the *Drosophila* embryo. In contrast, a network of 14 elements was necessary to simulate the expression pattern of homeotic genes. Finally, the dorso-ventral expression was modeled with a network

formed by seven elements. This model permits not only to describe the early development of the *Drosophila* embryo, but also can be used to study the evolution of developmental pathways. Just with the elimination of two genes, and the addition of one gene and three connections, it is possible to obtain the homeotic gene patterning of the beetle or grasshopper.

CONCLUDING REMARKS

Biological components, most often macromolecules, altogether with their interactions controlling physiological processes are considered regulatory networks. In the particular case of genetic regulatory networks, nodes represent genes, while the connections usually depict the regulatory interactions established among genes by means of transcription factors. With the advance of molecular techniques, the availability of gene sequences is increasing at a high rate. Thus, it is necessary to integrate such molecular information into functional networks. For a long time there have been theoretical studies of the properties of discrete state networks as models for genetic regulation (Kauffman, 1969; Thomas, 1973; Glass, 1975). However, only until recently there are some models of particular genetic regulatory networks of some biological systems. *Drosophila melanogaster* is by far the most studied system from the point of view of regulatory networks. The complexity of the models and data acquisition technology elaborated for *Drosophila* have reached high sophistication (see for example Kosman and Reinitz, 1998). Apart from such case, other models are arising slowly. With the advent of whole genome sequencing projects, plus some technologies like DNA chips, the identification of new genes and their regulatory interactions are imminent. For such reason, it is important to show the integrative capacity of the regulatory network modeling approach.

Traditionally, when genetic and biochemical studies suggest a sequence of regulatory events, people use arrows to represent the interactions among macromolecules. Such graph representation is necessary but not sufficient to understand the dynamical behavior of the pathway. This mini-review was aimed to show that finite-state networks offer a simple methodology that permits to study the collective behavior of a large number of elements. Moreover, this kind of modeling permits a thorough dynamical and analytical study, which gives important information to elaborate more realistic continuous models. Despite the clear usefulness of regulatory networks, it is important to keep in mind the difficulties behind network modeling. A significant problem in elaborating genetic regulatory networks of any kind is the correct inference of the genetic interactions. Such process involves a thorough analysis of a large quantity of relevant experimental literature to infer the connectivity. This step includes at least two possible sources of error. The first is involved with the discrimination process, what information is relevant and what is superfluous? It is not evident at all if a reported experimental result is related to a particular regulatory process, or rather to a molecular response not expected by the modeler. An ideal solution would be the formation of interdisciplinary groups with experts in the modeling process and experimentalists familiar with biological systems. However, this is not a common trend; usually the modeler has to incorporate information from different experimental groups. This brings about the second problem. Different experimental laboratories often work with different methodologies, animal subspecies, plant ecotypes, etc. Here again, how to distinguish between changes due to regulatory process or due to differences in the data acquisition methodology? There is no easy response; one has to balance between knowledge of the experimental system and modeling intuition.

To circumvent the problems inherent to the inference of network connectivity, there is an interest in the development of algorithms for the deduction of genetic networks, departing from temporal expression patterns of a group of genes. Many algorithms are based on the supposition that the regulatory networks behave as boolean networks. For

example, Liang *et al.* (1998) proposed an algorithm named REVEAL for network inference departing from state transition tables, which might correspond to time series. The algorithm uses a mutual information analysis to reduce the quantity of data points necessary for the inference. Notably, a small number of state transition pairs (100 from a total of 10^{15} possible) were sufficient to infer networks made of 50 nodes. Using another algorithm (Akutsu *et al.*, 1999), it has been proved mathematically that the number of transition pairs necessary to identify the network is proportional to $log\, n$, where n is the number of nodes. Despite the efficiency of the algorithms, their implementation in experimental laboratories requires a high precision in data acquisition of genetic expression, which is very difficult. However, recent advances, like the use of RT-PCR assays, permit the simultaneous measurement of expression of a large number of genes at different time points. A notable example of this kind of technology is the gene expression matrix of Wen *et al.* (1998), which contains the expression levels of 112 genes.

The convergence of theoretical and experimental advances has an important impact on the basic understanding of global biological processes. For example, it is known that all cells in a multicellular organism have the same genetic regulatory network. Nevertheless, the expression patterns and genetic responses are different in distinct cellular types. Thus the response to environmental changes and cellular signals are particular to each type of tissue. This differential response despite the common underlying genetic material can now be understood with the use of the concept of genetic regulatory networks as dynamical systems. As explained before, the network may have different equilibrium states in its expression patterns. Each of those stable expression patterns might correspond to the genetic profile characteristic of a particular cell type. Of course, there are many conceptual issues that need to be clarified bot experimentally and theoretically. For example, what kind of mechanism permits the correct divergence in expression of the involved genetic networks during ontogenesis? To answer this and other questions, it is necessary to elaborate more integrative models that take into account not only the molecular level, but also the cellular constitution of an organism.

There is also a practical side in the use of genetic regulatory network models. The recent availability of large financial resources to establish multi-species genetic analyses, specially in biological systems with potential economic impact like crops, attracts the interest of many scientists (Bennetzen *et al.*, 1999). The technical achievements to assess whole-genome genetic expression are paving the ground for the establishment of large research groups to generate massive data sets of gene expression patterns. Such information availability of information does not warrant the understanding of the basic processes, controlling gene expression, and thus differentiation and morphogenesis. Now, the understanding of basic processes is necessary to the control of organism with potential economical importance. Think for example on the flowering process, its control might lead to higher yield of cotton, fruits, etc., with the consequent economical impacts.

The use of integrative methodologies is necessary to reproduce the complex dynamical spatio-temporal patterns of biological systems. In particular, the analysis of genetic regulatory networks as dynamical systems provide with a suitable tool for the integrative analysis of the large quantity of genetic data that is becoming available. Moreover, network models not only synthesize data but also permit the elaboration of predictions that are not evident by using the classical conception of hierarchical static models. For example, the prediction of missing regulatory interactions, of missing genes, or stable genetic expression patterns, are useful guides for the experimental biologists to continue with the molecular analysis of certain experimental organisms. Certainly, there are very few biological examples to show the utility of the network modeling. In this paper, I presented some characteristics of seven network models

to give a glimpse of the usefulness of the translation of regulatory pathways into dynamical systems. The understanding and use of genetic regulatory models is becoming indispensable as the output of whole genome projects accelerates.

BIBLIOGRAPHY

Akutsu, T.; Miyano, S.; and Kuhara, S. (1999). Identification of genetic networks from a small number of gene expression patterns under the boolean network model. *Pacific Symposium on Biocomputing* **4**: 17-28.

Aubert, D. (1998). *EMF* genes regulate shoot development in *Arabidopsis*. *Proceedings of the 9th international conference on Arabidopsis research*. University of Wisconsin, Madison. p156.

Bagley, R.J.; and Glass, L. (1996). Counting and classifying attractors in high dimensional dynamical systems. *J. theor. Biol.* **183**: 269-284.

Bennetzen, J.L. (1999). Plant genomics takes root, branches out. *Trends Genet.* **15**: 85-87.

Bodnar, J.W. (1997). Programming the Drosophila embryo. *J. theor. Biol.* **188**: 391-445.

Burstein, Z. (1995). A network model of developmental gene hierarchy. *J. theor. Biol.* **174**: 1-11.

Cannon, W.B. (1929). Organization for physiological homeostasis. *Physiol. Rev.* **9**: 399-431.

Cooke, J.; Nowak M.A.; Boerlijst, M.; Maynard-Smith, J. (1997). Evolutionary origins and maintenance of redundant gene expression during metazoan development. *Trends Genet* **13**: 360-364.

Delbrück, M. (1949). Discussion. In: *Unités Biologiques Douées de Continuité Génétique*. CNRS (Lyon) p33-34.

Glass, L. (1975). Classification of biological networks by their qualitative dynamics. *J. theor. Biol.* **54**: 85-107.

Glass, L.; and Kauffman, S.A. (1973). The logical analysis of continuous, non-linear biochemical control networks. *J. theor. Biol.* **39**: 103-129.

Gouzé, J.L. (1998). Positive and negative circuits in dynamical systems. *J. Biol. Systems* **6**: 11-15.

Horan, G.; Kovacs, E.; Behringer, R.; and Featherstone, M. (1995). Mutations in paralogous *Hox* genes result in overlapping homeotic transformations of the axial skeleton: evidence for unique and redundant function. *Dev. Biol.* **169**: 359-372.

Kauffman, S.A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *J. theor. Biol.* **22**: 487-467.

Kauffman, S.A. (1993). *The origins of order: self-organization and selection in evolution*. Oxford University Press.

Kempin, S.; Savidge, B.; and Yanofsky, M.F. (1994). Molecular basis of the *cauliflower* phenotype in *Arabidopsis*. *Science* **267**: 522-525.

Kosman, D.; and Reinitz, J. (1998). Automated assay of gene expression at cellular resolution. *Pacific Symposium on Biocomputing* **3**: 6-17.

Liang, S.; Fuhrman, S.; and Somogyi, R. (1998). REVEAL, a general reverse engineering algorithm for inference of genetic network architectures. *Pacific Symposium on Biocomputing* **3**: 18-29.

Mendoza, L.; and Alvarez-Buylla, E.R. (1998). Dynamics of the genetic regulatory network for *Arabidopsis thaliana* flower morphogenesis. *J. theor. Biol.* **193**: 307-319.

Mendoza, L.; and Alvarez-Buylla, E.R. (ms). A regulatory network model for the appearance of root hairs in *Arabidopsis thaliana*. **Submitted**.

Mendoza, L.; Thieffry, D.; Alvarez-Buylla, E.R. (1999). Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis. *Bioinformatics*. **15**: 593-606

Muraille, E.M.; Thieffry, D.; Leo, O.; and Kaufman, M. (1996). Toxicity and neuroendocrine regulation of the immune response: a model analysis. *J. theor. Biol.* **183**: 285-305.

Parcy, F.; Nilsson, O.; Busch, M.A.; Lee, I.; and Weigel, D. (1998). A genetic framework for floral patterning. *Nature* **395**: 561-566.

Plahte, E.; Mestl, T.; and Omholt, S.W. (1995). Feedback loops, stability and multistationarity in dynamical systems. *J. Biol. Syst.* **3**: 409-413.

Ruddle, F.H.; Bartels, J.L.; Bentley, K.L.; Kappen, C.; Murtha, M.T.; and Pendleton, J.W. (1994). Evolution of *HOX* genes. *Annu. Rev. Genet.* **28**: 423-442.

Sánchez, L.; Van Helden, J.; and Thieffry, D. (1997). Establishment of the dorso-ventral pattern during embryonic development of *Drosophila melanogaster*: a logical analysis. *J. theor. Biol.* **189**: 377-389.

Snoussi, E.H. (1998). Necessary conditions for multistationarity and stable periodicity. *J. Biol. Systems* **6**: 3-9.

Somogyi, R.; and Sniegoski, C.A. (1996). Modeling the complexity of genetic networks: understanding multigenic and pleiotropic regulation. *Complexity* **1**: 45-63.

Thieffry, D.; Snoussi, E.H.; Richelle, J.; and Thomas, R. (1995). Positive loops and differentiation. *J. Biol. Syst.* **3**: 457-466.

Thieffry, D.; and Thomas, R. (1995). Dynamical behaviour of biological regulatory networks –II. Immunity control in bacteriophage lambda. *Bull. Math. Biol.* **57**: 277-297.

Thomas, R. (1973). Boolean formalization of genetic control circuits. *J. theor. Biol.* **42**: 563-585.

Thomas, R. (1978). Logical analysis of systems comprising feedback loops. *J. Theor. Biol.* **73**: 631-656.

Thomas, R.; Thieffry, D.; and 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.

Wagner, A. (1996a). Genetic redundancy caused by gene duplications and its evolution in networks of transcriptional regulators. *Biol. Cybern.* **74**: 557-567.

Wagner, A. (1996b). Can nonlinear epigenetic interactions obscure causal relations between genotype and phenotype?. *Nonlinearity* **9**: 607-629.

Wen, X.; Fuhrman, S.; Michaelis, G.S.; Carr, D.B.; Smith, S.; Barker, J.L.; and Somogyi, R. (1998). Large-scale temporal gene expression mapping of central nervous system development. *Proc. Natl. Acad. Sci. USA*. **95**: 334-339.

APENDICE 3

J. theor. Biol. (1998) **193**, 307–319
Article No. jt980701

Dynamics of the Genetic Regulatory Network for *Arabidopsis thaliana* Flower Morphogenesis

LUIS MENDOZA AND ELENA R. ALVAREZ-BUYLLA*

Lab. de Genética Molecular y Evolución, Instituto de Ecologia, UNAM, Ap. Postal 70-275, *Coyoacán, D.F. CP*04510, *Mexico*

(*Received on* 26 *September* 1997, *Accepted in revised form on* 3 *March* 1998)

We present a network model and its dynamic analysis for the regulatory relationships among 11 genes that participate in *Arabidopsis thaliana* flower morphogenesis. The topology of the network and the relative strengths of interactions among these genes were based from published genetic and molecular data, mainly relying on mRNA expression patterns under wild type and mutant backgrounds. The network model is made of binary elements and we used a particular dynamic implementation for the network that we call semi-synchronic. Using this method the network reaches six attractors; four of them correspond to observed patterns of gene expression found in the floral organs of *Arabidopsis* (sepals, petals, stamens and carpels) as predicted by the ABC model of flower morphogenesis. The fifth state corresponds to cells that are not competent to flowering, and the sixth attractor predicted by the model is never found in wild-type plants, but it could be induced experimentally. We discuss the biological implications and the potential use of this network modeling approach to integrate functional data of regulatory genes of plant development.

1. Introduction

As experimental data on gene function and regulatory interactions accumulate in biological model systems, the need of formal and modeling paradigms for functional inference and integration of large data sets is becoming imminent. In this paper we put forward a first trial to apply dynamic analyses of a genetic regulatory network model to integrate molecular and genetic data of *Arabidopsis thaliana* genes involved in flowering morphogenesis. In contrast to cascade or hierarchical models of regulatory genes, that are widely used in molecular biology (see for example Kornfeld, 1997), models of genetic regulatory networks consider direct and indirect feedback regulatory relationships (Garzon, 1990). Such feedback regulatory interactions are analogous to those present in metabolic routes, in which the product of a reaction regulates its own synthesis rate. The product of a certain gene might also regulate its own transcription rate directly or indirectly. Feedback loops make genetic regulatory network models dynamic systems, which may have fixed or periodic activity patterns. Moreover, feedback loops, characteristic of genetic regulatory networks, constitute the necessary mechanism to explain multiple equilibria and homeostasis of a given network (Thieffry *et al.*, 1995; Thomas, 1991; Thomas & D'Ari, 1990; Thomas *et al.*, 1995).

Transcriptional and post-transcriptional regulation are central issues for understanding the origin of cellular differentiation. Each cell type can be identified by its molecular profile (i.e. by the pattern of all molecular markers present in the cell), and in theory the cellular identity might be determined by describing all the active genes in the cell (Kauffman, 1969, 1991, 1993). Gene activity, in turn, depends on

* Present address: Department of Biology and Centre for Molecular Genetics, UCSD, La Jolla, CA 92037, U.S.A.

0022–5193/98/140307 + 13 $30.00/0

the presence of one or several transcription factors, which are themselves gene products, creating in this way a functional interdependence among a large group of genes. These groups of genes regulating the activity of each other are known as genetic regulatory networks. Formal dynamic analyses of these networks can be useful to explore the possible long-term genetic activation patterns (called attractors) of a given genotype. Such attractors have been identified with the different cell types of an organism (Kauffman, 1993; Zuckerkandl, 1994), thus in order to achieve cellular differentiation it would be sufficient to give a perturbation to the genetic regulatory network to change from one attractor to another. The perturbation might be in the form of a morphogen or an environmental factor, and the identification of such signaling factors or molecules is a matter of intensive experimental research (see Wolpert, 1996).

Many studies have explored the behavior of partial hypothetical regulatory networks (Kauffman, 1993; Clarket *et al.*, 1993), but there are few published applications of this dynamic approach to specific biological systems. Most of the published applications are for *Drosophila melanogaster* (Burstein, 1995; Reinitz & Sharp, 1995; Spirov, 1996) for which abundant genetic and molecular data has accumulated over the past few years. From this kind of model it has been possible to acquire insights that could have not been reached otherwise. For example, these models have suggested that a small group of homeotic proteins are sufficient to coordinate morphogenesis (Burstein, 1995), that the connectivity of the network determines a metabolic pathway (McAdams & Shapiro, 1995), or that it is possible to find a correlation among biochemical regulators and morphogenesis even when the mechanism is not completely known (Mjolsness *et al.*, 1991). Network models have been useful to make inferences on the evolutionary pathway of a group of regulatory genes (Spirov, 1996), or even to suggest missing components of a regulatory system (Loomis & Sternberg, 1995).

During recent years much has been learned about the genetic and molecular basis of flower morphogenesis in *Arabidopsis thaliana* (Coen, 1991; Weigel, 1995; Yanofsky, 1995). We used published genetic and molecular data for 11 genes to construct a genetic regulatory network for *Arabidopsis thaliana* flower morphogenesis and we provide analyses of the dynamic behavior of this network. To our knowledge, this is the first genetic regulatory network model for a plant, or part of a plant. We are particularly interested in the dynamics of the genetic regulatory network and we explore the hypothesis that the attractors of a network correspond to the activation states of specific cell groups. In this paper, we specifically addressed if the four gene activation states predicted by the ABC model of *Arabidopsis* flower morphogenesis (Meyerowitz, 1994a; see later) can be recovered as stable activation states of the regulatory network model that we put forward. We found that the dynamics of the network predicts six stable states: the four gene activation states of the ABC model, an activation state of cells in vegetative tissue, and anactivation state not found in wild-type plants.

1.1. *ARABIDOPSIS THALIANA* FLOWERS AND THE ABC MODEL

Flowers of *Arabidopsis thaliana* (hereafter *Arabidopsis*) are formed by four concentric whorls of flower organs made of, from outside to inside: four sepals (whorl 1), four petals (whorl 2), six stamens (whorl 3) and two fused carpels (whorl 4). This particular organization can be disrupted by mutations in different genes, and the analysis of such mutations have led to the proposition of a combinatorial model that has been used extensively to describe the morphology of *Arabidopsis* flowers in wild-type and mutant plants (Fig. 1). The so-called ABC model (Coen & Meyerowitz, 1991; Meyerowitz 1994a) postulates the existence of three different activities (A, B and C) which are each active in two adjacent whorls, and their combination determines the identity of the organs that develop in the flowers. According to the model, the presence of activity A will determine the differentiation of sepals, a combination of activities A and B will result in petals, while the presence of both B and C will give rise to stamens, and finally activity C alone results in the formation of carpels. Additionally, the ABC model postulates a mutual inhibition between activities A and C, such that when function A is absent function C substitutes it and vice versa.

There are several specific genes related to the three above mentioned activities. *APETALA*1 (*AP*1) and *APETALA*2 (*AP*2) have been considered A function genes (Bowman *et al.*, 1991, 1993. In the *Arabidopsis* literature, proteins are abbreviated using uppercase letters, wild-type genes with uppercase italics, and mutated genes with lowercase italics), because plants mutated in either gene yield flowers lacking sepals and petals. However, while molecular data confirmed the expected spatial distribution of *AP*1 mRNA according to the ABC model (Mandel *et al.*, 1992), *AP*2 mRNA is present throughout the flower and is also present in non-floral organs (Jofuku *et al.*, 1994). Even though *AP*2 might be regulated at a

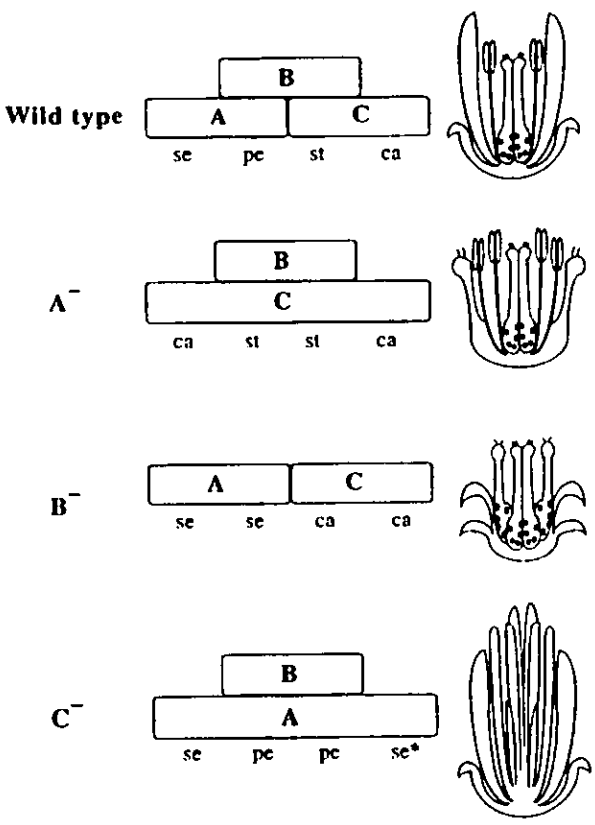


FIG. 1. The ABC combinatorial model proposes that the four different floral organs are determined by the specific combination of three different functions or activities. Activity A specifies sepals, activities A and B specify petals, activities B and C specify stamens, and activity C alone specifies carpels. Additionally, the ABC model postulates a mutual inhibition between activities A and C, such that when the function A is absent function C takes its place and vice versa. The rightmost figures are schematic representations of *Arabidopsis thaliana* flowers, which in wild-type plants (top) are composed from the outside to the inside of four sepals, four petals, six stamens and two carpels. Mutants in genes that confer each of the three floral activities and their effect on floral morphology are shown. se = sepals, pe = petals, st = stamens, ca = carpels, se* = an iteration of sepals, petals and petals.

post-transcriptional level, for simplicity we will consider *AP1* as the only A function gene. On the other hand, it is known that *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) constitute the B activity (Krizek & Meyerowitz, 1996). Finally, *AGAMOUS* (*AG*) is the only reported C function gene (Bowman *et al.*, 1991; Sieburth *et al.*, 1995; Yanofsky, 1995; Yanofsky *et al.*, 1990). Even though some authors have suggested minor modifications to the ABC model (Ma, 1994; Ray *et al.*, 1994) based on new molecular data, the model has been very robust for describing overall flower morphology of mutant (Coen & Meyerowitz, 1991; Ma, 1994; Meyerowitz, 1994a) and transgenic plants (Krizek & Meyerowitz, 1996; Mizukami & Ma, 1992). Moreover, the molecular and genetic mechanisms implied in this model seem to be conserved in virtually all angiosperms (Bowman, 1997; but see Vergara & Alvarez-Buylla, 1997).

2. The Network Model

We reviewed the literature looking for molecular and/or morphological data that could reveal the regulatory interactions among 11 genes involved in flowering morphogenesis of *Arabidopsis*. Four of these genes are the ABC genes described above. Based on this information we constructed the genetic regulatory network that we present as the NET model in Fig. 2. Most of the postulated gene interactions represent regulatory interactions at the transcriptional level. Each element in the network represents one gene with one exception: the element referred to as *BFU* in the network implementation denotes a protein heterodimer formed by AP3 and PI, this complex forms an active transcription factor (see later), therefore this interaction is represented by merging arrows acting over *AP3* and *PI*. Further

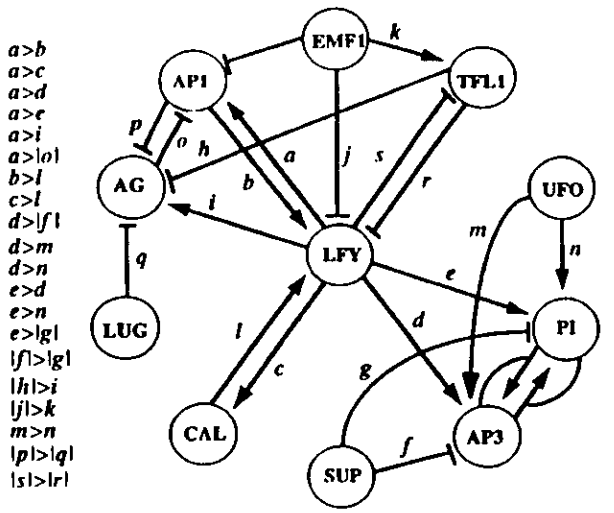


FIG. 2. The NET model. Arrows represent activations and flat-end lines represent inhibitions. Genes included are *EMBRYONIC FLOWER 1* (*EMF1*), *TERMINAL FLOWER 1* (*TFL1*), *LEAFY* (*LFY*), *APETALA1* (*AP1*), *CAULIFLOWER* (*CAL*), *LEUNIG* (*LUG*), *UNUSUAL FLORAL ORGANS* (*UFO*), *APETALA3* (*AP3*), *PISTILLATA* (*PI*), and *SUPERMAN* (*SUP*). Lowercase letters indicate the weight of interactions while inequalities at the left are the relative values inferred from experimental data. Absolute values represent the effect of de-repression, for example $|h| > i$ means that the absence of the repression of *TFL1* has a stronger impact over *AG* expression than its activation by *LFY*.

more, molecular data enabled us to obtain the sign of each interaction (activation or inhibition, arrow-head or dash-head, respectively), and the *relative strength* of many interactions. We explain case by case the logic that led us to each of these relative values.

2.1. IMPLEMENTATION OF THE SET MODEL

We wanted to know if the dynamics of the genetic regulatory network leads to attractors that correspond to the genetic activities proposed in the ABC model for each of the four floral organs. Despite the great advances in the studies of gene expression of *Arabidopsis*, the published molecular data is still qualitative. Nevertheless, we have used this data to put forward the first network architecture and we explore the final stable patterns of gene activity (attractors of the network).

We do not have data regarding rates of transcription but rather know only if a gene is either active or inactive at certain time intervals. Furthermore, the spatial resolution of such data does not provide information on the expression dynamics of genes inside individual cells. Therefore, we cannot construct a state transition diagram to infer the network connectivity as has been achieved in some studies of neural networks (Glass & Young, 1979), or biochemical control networks (Glass & Kauffman, 1973). However, the molecular data at hand provide enough information to establish gene connectivity and a first proposition of the relative effect of some interactions (see later). Consequently, we decided to use a network realization that takes into account both gene activation states and relative interaction weights, the implementation takes the form of a difference equation:

$$x_i(t+1) = \mathbf{H}\left(\sum_j w_{ij} x_j(t) - \theta_i\right) \qquad (1)$$

where $\mathbf{H}$ is the Heaviside step function:

$$\mathbf{H}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

The vector state indicating the activity of the n elements (genes) of the network is represented by $\mathbf{X} = (x_1, x_2, \ldots x_n)$ where $x_i \in \{0, 1\}$: meaning that a gene can only be in either of two states, active (1, maximum transcription rate) or inactive (0). In this network implementation a gene x_i becomes active if the weighted input of all the genes regulating it exceeds a certain activation threshold θ_i. Biologically, thresholds represent the amount of activators/inhibitors needed for turning on/off the transcription of that gene. Since we do not have any experimental evidence regarding the magnitude of such thresholds, we used integer values to keep computation simple, i.e. $\theta \in \mathbf{Z}$. A similar situation applies for the weights (w_{ij}) of every interaction, thus $w_{ij} \in \mathbf{Z}$. Regarding the weights or relative strengths of the interactions, however, we based their sign (positive for activations and negative for inhibitions) and *relative* magnitudes on experimental data unless otherwise indicated (see next section and inequalities in Fig. 2). The relative values that we propose here imply differences in the transcription factors' efficiency, concentration, or both. Although we propose here a simple network model with on/off elements, previous theoretical work has yielded maps to compare some shared dynamic features between continuous and discrete networks (Glass, 1975; Glass & Kauffman, 1973). Therefore, the model we present here could be used, in principle, to construct more realistic continuous models, that should still keep the same long-term dynamic behaviors.

For our purposes, the thresholding behavior implied in eqn (1) was adequate because we were able to obtain the gene expression patterns observed in the flower, namely, four stationary states corresponding to the four regions of the ABC model that give rise to the floral organs (see ahead). Nevertheless, such implementation does not allow for interactions like those shown in Fig. 2 between the products of *AP3* and *PI*. To circumvent this problem we introduced a network element (*BFC*, for B function) receiving inputs only from *AP3* and *PI*, in such a way that it becomes active if and only if *AP3* and *PI* are both active (this is a logical AND function, see experimental data later). Finally, we assigned numeric values to all thresholds and interactions between elements of the network. In networks, however, different values might result in different dynamics, therefore we tried to find the lowest integer values that result in the long-term activation patterns observed in the *Arabidopsis* flower. Remember that our objective is to evaluate if the *architecture* of the network, rather than the specific values, is compatible with the observed experimental data. For that reason, we used the lowest numerical values in our model to avoid as much as possible hypotheses that cannot be supported by experimental data at the moment.

All interactions among network elements depicted in Fig. 2 are presented in matrix $\mathbf{W}$, where each element w_{ij} represents the weight of the interaction from gene x_j to gene x_i. Likewise, in vector $\boldsymbol{\theta}$ the θ_is represent the thresholds of activation for each network element x_i. The order of the elements are

EMF1, *TFL1*, *LFY*, *AP1*, *CAL*, *LUG*, *UFO*, *BFU*,*AG*, *AP3*, *PI* and *SUP*.

$$\mathbf{W} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & -1 & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & -2 & 1 & -2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & -2 \\ 0 & 0 & 4 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \boldsymbol{\theta} = \begin{bmatrix} 0 \\ 0 \\ 3 \\ -1 \\ 1 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

The numerical values of matrix **W** and vector $\boldsymbol{\theta}$ were obtained using a genetic algorithm. We implemented a program that used only mutations, where a particular value of w_{ij} or θ_i chosen at random was changed for another one that fulfilled the following restrictions: (a) inequalities in Fig. 2; (b) network topology (i.e. no interactions were created or disappeared); and (c) the sign of interactions in Fig. 2. Such procedure was done with a population of 20 networks. After each mutation, every network was studied to examine if it had as stable states the four gene activation patterns corresponding to the ABC model and to non-flowering. After this step, each network was assigned with a fitness value. The fitness value was a linear function of the number of desired fixed points. Finally, to avoid local maxima during the selection step, besides the network with the highest fitness value, a randomly chosen network was kept. These two networks were used to replenish the network population, and the cycle of mutation, assignment of fitness values and selection were repeated until the population fitness reached a maximum.

The final issue regarding the implementation of the NET model concerns how to solve the transition from the vector state $\mathbf{X}(t)$ to $\mathbf{X}(t+1)$. The easiest way is the synchronous approach (as in Kauffman, 1969, 1991) in which eqn (1) is applied to all network elements at the same time. This dynamic description is problematic from the biological point of view, because it implies the unlikely situation in which all the genes respond exactly at the same time. Conversely, the asynchronous approach (as in Thomas, 1991) consists in solving eqn (1) for one network element at each time step, once the order in which the elements to be solved has been specified. This asynchronous approach is also problematic for our purpose, because the order for solving eqn (1) might change the gene activation pattern and the long term stable states, and there is still no experimental data on the precise order of activation of individual genes. We therefore decided to introduce a combined, biologically inspired approach, henceforth named semi-synchronic. In this method, the elements of the network are divided into groups. Hence, eqn (1) is solved synchronically for elements within the same group, and asynchronically for elements in different groups (see Fig. 3). We used experimental data to decide the order of activation of the different groups. Therefore, we think that this method is more appealing from the biological point of view, than previous ones.

Experimental biologists have grouped the genes related to flowering into a hierarchy of four sets of genes depending on their time of activation as the transition to flowering and flower morphogenesis proceeds. For the genes included in our network, *EMF1* and *TFL1* belong to the first group of genes to become active, namely the group of early and late floral genes (Coupland, 1995). The next set of genes to become active is the group of meristem identity genes (Weigel *et al.*, 1992), in which *LFY*, *AP1* and *CAL* are included. Then, the so-called cadastral genes become active (Weigel & Meyerowitz, 1993a), from

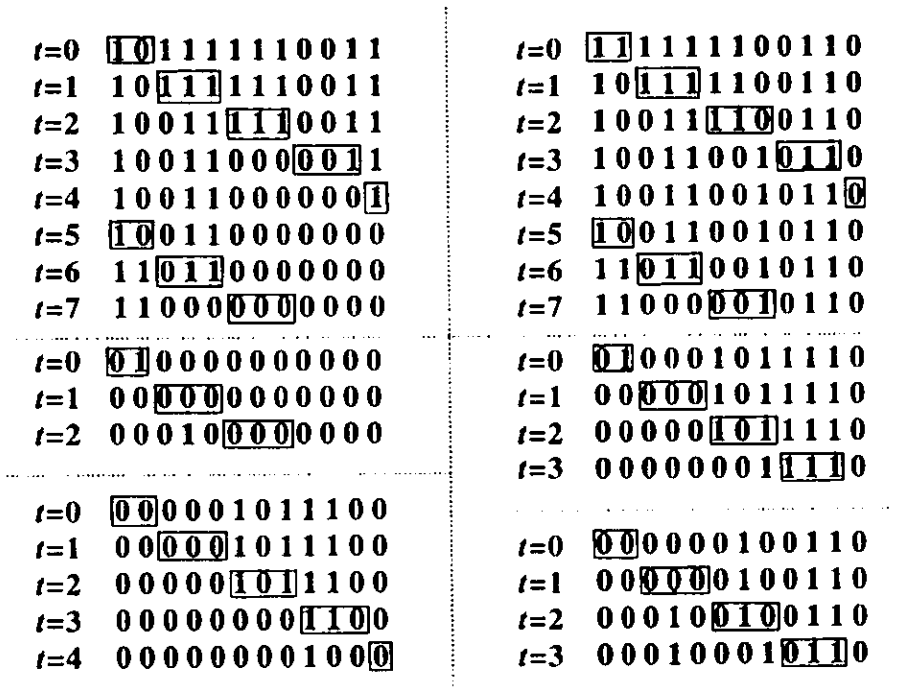


FIG. 3. A sample of the network activation patterns. These are some of the initial states (marked by $t = 0$) that eventually lead to one of the six fixed points. Rectangles indicate the group of element to which eqn (1) is applied to obtain the following network activation state. The order of the network elements here and throughout the text is from left to right *EMF1*, *TFL1*, *LFY*, *AP1*, *CAL*, *LUG*, *UFO*, *BFU* (B function, see text for details), *AG*, *AP3*, *PI* and *SUP*.

which *LUG* and *UFO* are representatives. It is important to mention that *LUG* and *UFO* are included in the same category due to their effect on flower morphology rather than to their spatial and temporal expression patterns, since *LUG* is not yet cloned and *UFO* is active at all developmental stages (Lee *et al.*, 1997). Finally, the ensemble of organ identity genes (Ma, 1994), represented in our model by *AG*, *AP3* and *PI*, become active. It is important to say that *SUP* was first considered as a cadastral gene, but it becomes active after the organ identity genes (Sakai *et al.*, 1995). Based on these considerations, we have divided the 12 network elements into five groups: the first containing *EMF1* and *TFL1*; the second *LFY*, *AP1* and *CAL*; the third *LUG*, *UFO*, and *BFU*; the fourth *AG*, *AP3* and *PI*; and the fifth with only *SUP*.

Another important biological consideration while performing the dynamic analysis, concerns the activation state of *EMF1*. This gene is part of the proposed floral repressor (Haughn *et al.*, 1995; Weigel, 1995), which supposedly controls the transition from vegetative to reproductive growth, and it is proposed to be under the influence of many upstream genes (Coupland, 1995). To incorporate such an external influence, we decided to fix the activation state of *EMF1* (either 0 or 1) throughout each dynamic cycle. In this way, we simulated the effect of an external factor not incorporated in the network. We achieved this by changing w_{11} from 0 to 1.

2.2. MOLECULAR BASIS OF THE NET MODEL

LEAFY (*LFY*) is known to be positively regulated by *AP1* and *CAULIFLOWER* (*CAL*), because its mRNA is reduced in *ap1 cal* plants (Bowman *et al.*, 1993; Kempin *et al.*, 1995; Weigel & Nilsson, 1995). However, *AP1* and *CAL* products are not needed simultaneously to activate *LFY* because single and double mutants for these two genes have different floral morphologies and *LFY* mRNA levels (Bowman *et al.*, 1993; Gustafson-Brown *et al.*, 1994). Therefore, *AP1* and *CAL* regulations over *LFY* are part of two independent pathways. Conversely, *AP1* mRNA (Weigel & Nilsson, 1995) onset is delayed in *lfy* mutants. The relative reduction of *AP1* mRNA in *lfy* plants is more pronounced than the reduction of *LFY* mRNA in *ap1* plants. Therefore, we propose that the up-regulation of *AP1* by *LFY* is stronger than the up-regulation of *LFY* by *AP1* ($a > b$; see Fig. 2). Furthermore, the phenotype of *35S::LFY ap1* plants

(Weigel & Nilsson, 1995; *35S::LFY* is a molecular construction that over-expresses *LFY*) is between that of wild-type and *35S::LFY* plants, suggesting that a large function of *LFY* is to activate *AP1*. We incorporated this in the NET model by making the relative value of such activation greater than any other activation from *LFY* to any other gene (i.e. $a > c, a > d, a > e$ and $a > i$). Unpublished data suggest that a similar regulatory interaction to that documented between *LFY* and *AP1* exists between *LFY* and *CAL* (Savidge & Yanofsky, unpublished data; see also Fig. 6 in Weigel & Nilsson, 1995). In order to maintain the symmetry of the relationship between *LFY* and *AP1*/*CAL*, we made the activation of *LFY* over *CAL* stronger than that of *CAL* over *LFY* ($e > 1$ in the NET model). Finally, because *ap1* plants have a more pronounced mutant phenotype than *cal* plants we made $b > 1$.

TERMINAL FLOWER 1 (*TFL1*) is a repressor of *LFY* (Okamuro *et al.*, 1993; Weigel *et al.*, 1992), because in *tfl1* mutants, both *AP1* and *LFY* mRNAs are ectopically expressed (an expression found outside the wild-type pattern, Bowman *et al.*, 1993; Gustafson-Brown *et al.*, 1994). But the effect of *TFL1* over *AP1* seems to be mediated by *LFY*, because *tfl1 ap1* mutants have an additive phenotype (Bowman *et al.*, 1993; Shannon & Meeks-Wagner, 1993). If the morphology of double mutants is the summation of the phenotypes of the individual single mutants, we consider that one of the two genes is not directly downstream of the other. Besides, *tfl1 lfy* double mutants (*tfl1*-10 *lfy*-16) have determinate growth and a flowering time similar to that of the *tfl1* single mutant (Shannon & Meeks-Wagner, 1993), suggesting that *LFY* and *TFL1* are in the same pathway, one downstream of the other. Conversely, *LFY* seems to inhibit *TFL1* transcription, because *35S::LFY* plants are very similar to *tfl1* mutants (Shannon & Meeks-Wagner, 1991; Weigel & Nilsson, 1995). The *ap1*-1 mutation largely attenuates the *35S::LFY* phenotype, nevertheless those plants still form terminal flowers (Weigel & Nilsson, 1995) as in *tfl1* mutants, indicating that *LFY* inhibits *TFL1* using an *AP1*-independent pathway. However, *35S::LFY* plants produce a terminal flower before and at a shorter inflorescence stem size than *tfl1* mutants, therefore, we postulate that the inhibition of *TFL1* over *LFY* is weaker than the inverse inhibition ($|s| > |r|$). In other words, the effect caused by the disinhibition of *LFY* in the *tfl1* mutants is not sufficient to cause the dramatic effects seen with the constitutive expression of *LFY* in *35S::LFY* plants. Such contrast might be due to a difference in the levels of the LFY protein, to the activation or inhibition of other downstream genes, or a combination of both. This mutual inhibition between *LFY* and *TFL1* is supported by the fact that their spatial domains of mRNA expression are contiguous but they do not overlap (Bradley *et al.*, 1997).

There is much evidence indicating that *LFY* activates both B activity genes, *AP3* and *PI* (Goto & Meyerowitz, 1994; Jack *et al.*, 1994; Weigel *et al.*, 1992; Weigel & Meyerowitz, 1993a,b). However, in *lfy* mutants *PI* mRNA reduction is more pronounced than *AP3* mRNA reduction, therefore suggesting that the activation of *PI* by *LFY* is stronger than that of *AP3*, ($e > d$). Likewise, the gene *UNUSUAL FLORAL ORGANS* (*UFO*) activates *AP3* and *PI* (Lee *et al.*, 1997; Levin & Meyerowitz, 1995; Wilkinson & Haughn, 1995), but in this case *AP3* mRNA levels are more dramatically reduced in *ufo* mutants than levels of *PI* mRNA ($m > n$). Finally, morphological analyses of mutant flowers suggest that *LFY* is a stronger activator of B function genes than *UFO* ($e > d, e > m, e > n, d > m, d > n$, and $m > n$). It is necessary to mention that *UFO* has a role in cell proliferation or organ initiation (Lee *et al.*, 1997), therefore it is quite probable that the effect of this gene over *AP3* and *PI* is not a direct regulation of expression. Nevertheless, the dynamic study presented here might be valid even when some of the proposed regulatory interactions turn out to be indirect mediated by other products or processes (see discussion). It has been proposed that *UFO* and *LFY* might act as coregulators (Lee *et al.*, 1997). However, since the single and double mutants have different phenotypes, the putative interaction between LFY and UFO are not yet clear and therefore we decided to wait for more experimental results before including such interaction in our model.

In Fig. 1, merging arrows between AP3 and PI indicate that a dimer of the proteins encoded by these two genes is formed and that it maintains the activities of both *AP3* and *PI* (Goto & Meyerowitz, 1994; Jack *et al.*, 1994). Finally, *SUP* inhibits both *AP3* and *PI* (Bowman *et al.*, 1992; Goto & Meyerowitz, 1994; Shultz *et al.*, 1991); and once more, *AP3* mRNA expression is the most affected one in *sup* mutants (Sakai *et al.*, 1995) suggesting that $|f| > |g|$. In *sup* mutants, in contrast to wild-type, *AP3* mRNA is partially expressed in the fourth whorl during late stages of flower development. *PI* mRNA, on the other hand, is present in the fourth whorl from the onset of its expression, as normally does, but is maintained at detectable levels throughout development in contrast to the pattern of mRNA expression in wild-type (Sakai *et al.*, 1995). These changes in mRNA expression of *AP3* and *PI* are less drastic than

those observed in *lfy* mutants, where *AP3* and *PI* mRNAs are abolished completely, therefore suggesting that $e > |g|$ and $d > |f|$ in Fig. 2. Here again, *SUP* seems to have a role in cell division (Sakai *et al.*, 1995) rather than a direct regulation of *AP3* and *PI* gene expression. Its inclusion in the NET model, however, reveals the possible existence of another, yet undiscovered, inhibitor of the B function genes (see Discussion).

The mechanism by which *AP1* and *AP2* (activity A) inhibit *AG* (activity C) is still uncertain, with the extra problem that most of the genetic analyses are reported for weak *ap1* and *ap2* alleles. It is known that *AP2* is needed for the inhibition of *AG* in the whorls that will give rise to sepals and petals, but not in those where stamens and carpels will arise even though *AP2* messenger is expressed throughout the flower (Jofuku *et al.*, 1994). However, since *AP1* mRNA is present in the two outer whorls (Mandel *et al.*, 1992) many authors have suggested that a combination of both AP1 and AP2 are needed for the inhibition of *AG* (Bowman *et al.*, 1993; Liu & Meyerowitz, 1995; Ma, 1994). Nevertheless, these regulatory interactions are still controversial and the mechanism is far from clear. However, our purpose is to investigate the dynamic properties of the genetic regulatory network, and *AP2* seems to be a constitutively expressed gene in the *Arabidopsis* flower. Therefore, for simplicity we will assume that *AP1* inhibits *AG*, keeping in mind that *AP2* and perhaps other factors are also needed.

AG is activated by *LFY* (Weigel & Meyerowitz, 1993c), and inhibited by *LEUNIG* (*LUG*; Liu & Meyerowitz, 1995; Meyerowitz, 1994b). However, single *ap1* mutants have more severe floral transformations than single *lug* mutants ($|p| > |q|$). *AG* itself provides an important feedback regulation by means of its inhibitory effect over *AP1* (Gustafson-Brown *et al.*, 1994). But the strongest regulation over *AP1* seems to come from LFY (rather than from *AG*) because flowers of *35S::AG* plants (Mizukami & Ma, 1992) start diverging morphologically from wild-type flowers when *LFY* expression diminishes in the center of the flower primordium (Weigel *et al.*, 1992). Hence the up-regulation of *AP1* by *LFY* seems to be stronger than the down-regulation of *AP1* by *AG* ($a > |o|$). An experimental way to test this hypothesis is to observe flower development in *35S::AG lfy* plants, our model predicts that such flowers would diverge morphologically from wild-type earlier than their *35S::AG* counterparts, since the former plants would not have the activation of *LFY* over *AP1*, while the inhibition of *AG* over *AP1* would always be present. There are other possible experiments of this kind that might help to unravel genetic interactions, for example if *35S::AG ap1* plants start diverging from wild-type while *LFY* diminishes its expression in the center of the flower, that would indicate that there is another gene acting downstream of both *AG* and *LFY* but mainly regulated by *LFY*.

Finally, concerning *AG*, it seems that *TFL1* inhibits *AG*. Flowers of *ap1 ap2* double mutants have an inflorescence-like morphology and also lack a proper central pistil. For example, the fourth whorl of *ap1-1 ap2-2* (strong) mutants is formed by carpels that often fail to fuse and have sectors of stamen tissue [Bowman *et al.*, 1993, see Fig. 8(c)], suggesting an alteration of the C activity. However, *tfl1 ap1 ap2* mutant flowers do have pistil (Shannon & Meeks-Wagner, 1993), suggesting that the *tfl1* mutation re-establishes the C activity. Given that *AG* is the only reported C function gene, we infer that the *tfl1* mutation causes an increase in *AG* mRNA, and hence in wild-type plants *TFL1* would be normally inhibiting *AG*. This inhibitory relationship is not mediated by *LFY*, because *tfl1 lfy* plants flower at the same time as *tfl1* plants do (Shannon & Meeks-Wagner, 1993), indicating that the *LFY* product is not needed. Furthermore, *lfy* plants flower later and *tfl1* earlier than wild-type plants; then if *tfl1 lfy* plants flower at the same time as *tfl1* mutants, we can say that *AG* expression is more severely affected by the lack of *TFL1* product than by the lack of *LFY* product, thus establishing that $|h| > i$.

EMBRYONIC FLOWER 1 (*EMF1*) is the most upstream gene we have considered. Actually, *EMF1* has not been cloned and we rely only on morphological data of the mutant plants to make inferences about possible regulatory interactions involving this gene. *EMF1* is considered as part of the so-called floral repressor (Haughn *et al.*, 1995; Weigel, 1995), and it may fulfil such a function by two, non-excluding possibilities: either *EMF1* inhibits the floral promoting genes (*LFY* and *AP1*) or it activates the other floral repressing gene (*TFL1*). The morphology of mutants suggests that both pathways may be operating. First, *lfy* mutants grown under short photoperiod have a stronger mutant phenotype than the same plants grown under continuous light (Schultz & Haughn, 1993). Actually, *lfy* plants under short photoperiods have hardly any flowers, just like *lfy ap1* plants under long photoperiods (Haughn *et al.*, 1995; Weigel *et al.*, 1992). This suggests that under short photoperiods *AP1* expression is absent or very reduced in *lfy* plants, thus acquiring a *lfy ap1*-like phenotype. This putative inhibition might be carried out by either *EMF1* or *TFL1*, the late-flowering genes. However, *TFL1* mRNA is found just

below the inflorescence dome and absent in the regions that will give rise to flowers (Bradley *et al.*, 1997), thus *TFL1* mRNA is never present in the regions where *AP1* mRNA eventually appears, making very unlikely that *TFL1* would repress *AP1*. Such spatial considerations leaves only *EMF1* the possible repressor of *AP1*.

Indirect evidence suggests that *EMF1* activates *TFL1*. Long day grown *emf1 tfl1* plants flower just like *tfl1* single mutants (Yang *et al.*, 1995), suggesting that *TFL1* is directly downstream of *EMF1*. Additionally, pistils are the most prominent organs in *emf1* flowers (Yang *et al.*, 1995), suggesting that those mutants over-express *AG*, presumably due to the absence of the inhibitory effect of *TFL1*. Finally, *EMF1* could also be directly inhibiting *LFY*, because *emf1 lfy* plants develop *lfy*-like flowers with leaf-like organs, but do not develop *emf1*-like flowers (Yang *et al.*, 1995), again suggesting that *LFY* is directly downstream of *EMF1*. Furthermore, it has been shown that the gene *CONSTANS* (*CO*) activates *LFY* (Simon *et al.*, 1996) mediated by the inhibition of *EMF1* (Coupland, 1995). This is, *CO* inhibits *EMF1* which in turn inhibits *LFY*. Finally, the *emf1* mutation causes an early flowering which is more similar to that presented by *35S::LFY* than *tfl1* plants. This morphological evidence suggests that the morphology of *emf1* plants is due to an over-activation of *LFY* rather than to a lack of *TFL1* expression ($|j| > k$).

2.3. DYNAMICS OF THE NET MODEL

We made an exhaustive examination of the dynamics of the model. We tested all possible initial states ($2^{12} = 4096$ states) and we used the semi-synchronic method for solving eqn (1) for a sufficiently long time period ($t = 200$) to find the final activation states of the network. In such a way we obtained the attractors and basins of attraction of the NET model (Fig. 4). Six stable fixed points were obtained, five of them with a clear biological meaning. Instead of writing the network activity state as a vector, like $X(t) = (0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0)$, we will simply write it down as a 12-digit binary number, i.e. 000100010110.

The first attractor (000100000000) corresponds to the exclusive activation of *AP1*, the A function gene. As discussed above, this pattern of activation is found in the first whorl of flowers, where sepals develop. The second fixed point (000100010110) is a stable activation of genes *AP1*, *AP3* and *PI*, and the presence of the AP3/PI protein complex. This attractor corresponds to the activation of A and B functions that determine petal formation in the

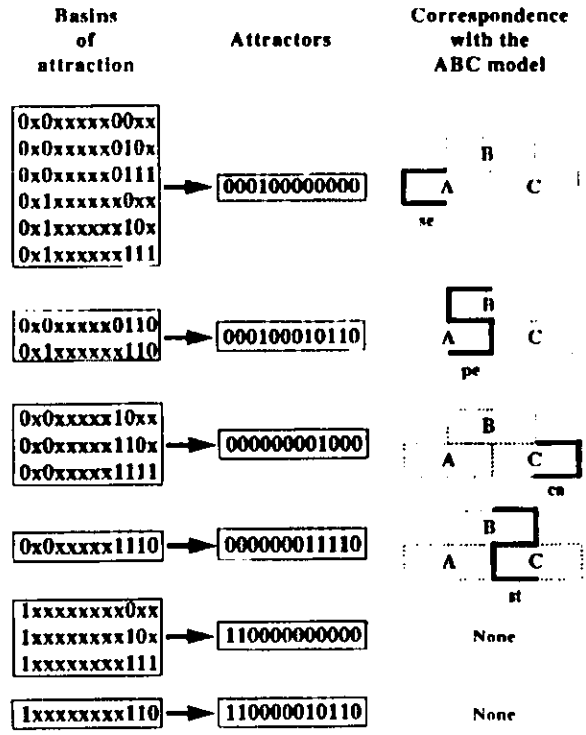


FIG. 4. Attractors, basins of attraction of the NET model and their correspondence with the ABC model. A basin (left column) is composed of all those initial states that eventually lead to a particular attractor (central column). 0 and 1 represent inactive and active genes respectively, while x means that an element might be either active or inactive. Four attractors have a direct equivalence with the four regions of the ABC model (right column), therefore it is possible to infer the floral organs that will develop from those fixed-point of the NET model

second whorl. We also obtained a fixed point (000000011110) representing the stable activation of *AG*, *AP3*, *PI* and the presence of the AP3/PI complex, (i.e. B and C functions active). This activation state corresponds to that proposed by the ABC model for the third whorl cells that differentiate into stamens. The fourth attractor (000000001000) corresponds to an exclusive activation of *AG*, the only C function gene. This activation pattern is found in cells of the fourth whorl that give rise to carpels. The last biologically meaningful attractor (110000000000) is a stable activation of the floral inhibitor genes *EMF1* and *TFL1*, characteristic of cells that will not become part of flowers. Heretofore, the dynamics of the NET model agrees with the well supported ABC model (Meyerowitz, 1994a).

The NET model has a sixth attractor (110000010110) with stable expression of *EMF1*, *TFL1*, *AP3*, *PI* and the complex of AP3/PI. In this

case we obtained the stable expression of the flowering inhibition genes at the same time as the presence of floral B function genes and products. This attractor does not agree with any observed gene expression pattern in wild-type plants, but we think that it can be experimentally induced. *AP3* and *PI* can be turned on despite the activation of the *EMF1* and *TFL1*, because of the mutual positive feedback present between *AP3* and *PI* (see Fig. 2). If for any reason the AP3 PI dimer is present at any stage of development, and in the absence of an inhibitory signal (notice that *SUP* is inactive in this basin of attraction), the B function genes will be expressed. The model predicts, therefore, that the presence of the AP3 PI dimer will be enough to originate and sustain their own expression as long as the proteins do not diffuse or degrade. Nevertheless, flowers would not appear in this putative region of induced *AP3* and *PI* expression because meristem identity genes (mainly *LFY* and *AP1*) will remain inactive. This type of experiment has not been performed but plants over-expressing both B function genes (Krizek & Meyerowitz, 1996) showed AP3 protein expression in some vegetative tissues.

3. Discussion

Many genes involved in *Arabidopsis thaliana* flowering and flower morphogenesis have been cloned in recent years (Weigel, 1995; Yanofsky, 1995). Messenger RNA expression patterns, altogether with their related phenotypes are accumulating fast and providing valuable functional information regarding such genes. Some interesting pictures of the complex regulatory interactions established by these genes are slowly arising (see for example Theißen & Saedler, 1995), but these schemes are mostly hierarchical and static. In this paper we provide a first example of how nonlinear dynamic models can help us understand the molecular mechanisms underlying flowering and flower morphogenesis.

The network presented here provides a first provisional architecture that yields as stable states the A, B and C activities and their combinations as proposed in the widely used ABC model (Fig. 1). This network model is, therefore, a first proposition of at least some of the molecular mechanisms underlying the ABC model. We do not claim that the NET model incorporates all the important genes involved in determining the cellular fate in flower morphogenesis. The inclusion of further genes would alter the network topology but it could be possible to maintain the four attractors that represent the activation states characterizing the regions where sepals, petals, stamens or carpels differentiate.

In this work, our main goal was to know if the molecular data available for genes related to *Arabidopsis* flower morphogenesis could be synthesized in a dynamic model compatible with the ABC model. Our goal in the long run is, however, to obtain a more general dynamic model that includes the ABC model as a particular case. In the meantime, the NET model provides with both a tool for further theoretical analysis, and a guide for future experimental work. For example, we are currently looking for the minimal network representation that still has the four activation regions represented by the ABC model. This exercise might prove useful for the experimental biologist working with *Arabidopsis* in order to distinguish direct from indirect regulatory relationships among these genes. Further simulations on this NET model may be used also to identify interactions for which the quantitative value is critical to recover the expected stable states. Special effort might be put to obtain quantitative data for such interactions.

The attractors of the NET model (Fig. 4), show that two types of elements can be distinguished: those which initial states affect the attractor reached and those which initial states do no affect the attractor reached. The former are *EMF1*, *LFY*, *AG*, *AP3*, *PI* and *SUP*, and those that do not affect the final network state are *TFL1*, *AP1*, *CAL*, *LUG*, *UFO* and the AP3 PI complex (*BFU*). This could imply that a perturbation of one of the elements of the first group might be sufficient to change the fate of the cells in the developing organ by altering the identity of the attractor. On the other hand, a perturbation in any of the elements of the second group would not alter the cell differentiation commitment. Moreover, if our model turns out to be experimentally robust, it is possible to predict that floral genes still to be discovered most probably will either belong to the second group or connect to elements in that group. In any case, the inclusion of such new genes in the NET model would alter the topology of the network but would not modify its final attractors. Hence the capacity of a network of absorbing perturbations due to the inclusion or deletion of elements depends on the connectivity of the whole network. Previous theoretical statistical analyses (Wagner, 1996) have also shown that the final activation pattern in a network might remain unchanged, or very slightly modified, despite large deletions.

Genes belonging to each of these two groups might be prone to different evolutionary forces. We may speculate that network elements that are less critical

for the establishment of adequate activation patterns, might be more prone to neutral mutations, than those that are critical for maintaining the correct pattern. For example, if a mutation permanently "turns off" a gene member of those genes that do not alter the attractors of the network, then the transient activation pattern of the network might be altered, but the final activation state would remain unchanged. The reverse hypothesis can also be postulated. Genes belonging to the group that affect the attractors might be more sensible to mutations, and therefore the ratio of non-neutral to neutral mutations would be higher in these genes than in those not involved in establishing the network attractors. This bold hypothesis could be tested by making comparative analyses of sequence data of genes belonging to each of these two groups from different species.

There is another way in which the addition of a newly discovered gene would not alter our results. Suppose that the inhibition of *AG* by *LUG* is not a direct one as depicted in Fig. 2 but is indirectly mediated by a still undiscovered gene *X*. Adding this new gene to the NET model would certainly modify the topology of the network and might also alter the transient dynamic behavior. However, since *LUG* belongs to the group of genes that are not critical to reach the attained attractors, the activation state of any gene that is exclusively controlled by *LUG* cannot affect the attractors, because this imaginary gene X would be only a relay between *LUG* and *AG*. Therefore, we do not postulate that all the genes directly connected in our model have direct interactions *in vivo*. This has to be tested experimentally for most cases. We state, however, that many intermediary genes might act only as relays and would not affect the stable states reached. Nevertheless, not all possible alterations are of this kind, some gene inclusions might actually add new attractors to the dynamics of the network. These new attractors would not invalidate our model, however. In this first model we are dealing only with four major regions of gene activation as stated by the ABC model, but each floral whorl contains different cell types. Therefore, future experimental results may be incorporated into larger networks, ideally yielding as many attractors as cell types are in the entire flower.

We want to make clear that the dynamic approach used here has clear advantages over a static analysis. This point can be illustrated by comparing, for example, the activation state of *AP3*, *PI* and *SUP* in the basins of attraction (last three values in every vector, Fig. 4) to their activation in the corresponding attractors. Notice that those attractors in which the B activity is present (AP3 and PI active) come from initial states in which *AP3* and *PI* are active but *SUP* is inactive. This means that the activity of *SUP* determines if the B activity will be present or not: if during the initial state *SUP* is active, the B activity will not be established. This result indicates that a gene inhibiting the B activity genes is absolutely necessary to obtain a result compatible with the ABC model. However, *SUP* is probably acting at the cellular level regulating cell division rate rather than directly regulating the transcription of B genes. If this proves to be the case, our dynamic model predicts that there would be a still undiscovered gene that inhibits *AP3* and *PI* at the transcriptional level, because a negative transcriptional regulator of the B genes is required to yield some regions with and other regions without the B activity.

One of the objectives of this study was to ask if the network model could reach as stable states the four patterns of gene activation predicted for *AP1*, *AP3*, *PI* and *AG* in the ABC model. Although other genes may be involved in determining the A and C functions (Bowman, 1997), the ABC model is the best global description of the activities of flower organ identity genes. Therefore, the numerical values of network parameters that could not be estimated based on current experimental results were obtained with a genetic algorithm using a fitness function defined according to the gene activation patterns proposed in the ABC model. However, the topology of the NET model included more genes and more interactions than those included in the ABC model. Hence, there is no *a priori* reason to expect that the dynamic model proposed here should reach the same gene activation patterns predicted by the combinatorial static ABC model. Furthermore, the NET model constitutes a new hypothesis regarding the molecular mechanisms underlying the establishment of the ABC functions.

Although the dynamics of the NET model heavily depends on the non-zero values of **W** and $\boldsymbol{\theta}$, the topology of the network imposes restrictions. For example, it was impossible to obtain values of **W** and $\boldsymbol{\theta}$ that yielded only the four floral stable states plus a unique non-floral state. The sixth stable pattern of Fig. 4, that does not have a biological interpretation, is therefore a restriction imposed by the topology of the network on its dynamic behavior. On the other hand, certain network topologies never attain the ABC activation patterns as stable states, no matter which numerical values of **W** and $\boldsymbol{\theta}$ are used (including those used here, results not shown). Therefore, the coincidence between the NET and ABC models validates some of the assumptions incorporated in both models. Simulations in progress

(Mendoza & Alvarez-Buylla, in prep.) and additional theoretical (Thieffry *et al.*, 1995; Thomas, 1991; Thomas & D'Ari, 1990, Thomas *et al.*, 1995) and experimental work should help us elucidate which of these assumptions are valid and which are not.

How can we reconcile the use of the network approach with the temporal pattern of flower development in *Arabidopsis*? In early flower primordia (stages 1 and 2, Smyth *et al.*, 1990) before floral organ primordia are apparent, the regulatory genetic network of each cell forming the bud would be in the same state, namely 110000000000 according to the NET model. As cell division proceeds and mediated by still unknown cell cell signaling cues and mechanisms, the stable activation pattern of the network is changed only by turning off *EMF1*, for example. This putative signal acting only on one gene would be enough to ensure commitment of cells to differentiate into sepals, the organs of the outermost whorl of floral organs. Such a perturbation would take the network to the state 010000000000, that is part of the attractor that eventually leads to the state in which only *AP1* is on (000100000000) and determines sepal differentiation according to a simplification of the ABC model. As flower development proceeds, the same or other putative signals might prompt the genetic network to each of the other attractors for petals, stamens and carpels, thus completing flower development. This account is largely speculation, but it might be useful for thinking about the role that genetic regulatory networks play in the process of cellular differentiation and development of flowers.

Francisco Vergara contributed with some important ideas during the first stages of this work and his comments to previous versions of the manuscript were very useful. We thank Marty Yanofsky for detailed revisions and very illuminating discussions, as well as for sharing various unpublished results. Comments made by John Bowman, Caroline Burgeff, Alicia Gamboa, Liz Izquierdo, Medard Ng, Rosalinda Tapia, Detlef Weigel, Elliot Meyerowitz and Francois Parcy were also very useful. Pedro Miramontes and Denis Thieffry provided also very useful suggestions regarding the mathematical aspects of this work. This work was funded by an UNAM scholarship to L.M. and CONACYT and PSPA grants to E.A.B. as well as by financial support from the Institute of Ecology, UNAM, and a PEW fellowship to E.A.B.

REFERENCES

Bowman, J. L. (1997). Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. *J. Biosc. Rev.* (in press).

Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. (1991). Genetic interactions among floral homeotic genes of *Arabidopsis*. *Development* **112**, 1–20.

Bowman, J. L., Sakai, H., Jack, T., Weigel, D., Mayer, U. & Meyerowitz, E. M. (1992). SUPERMAN, a regulator of floral homeotic genes in *Arabidopsis*. *Development* **114**, 599–615.

Bowman, J. L., Alvarez, J., Weigel, D., Meyerowitz, E. M. & Smyth, D. R. (1993). Control of flower development in *Arabidopsis thaliana* by APETALA1 and interacting genes. *Development* **119**, 721–743.

Bradley, D., Ratcliffe, O., Vincent, C., Carpenter, R. & Coen, E. (1997). Inflorescence commitment and architecture in *Arabidopsis*. *Science* **275**, 80–83.

Burstein, Z. (1995). A network model of developmental gene hierarchy. *J. theor. Biol.* **174**, 1–11.

Clarket, B., Mittenthal, J. E. & Senn, M. (1993). A model for the evolution of networks of genes. *J. theor. Biol.* **165**, 269–289.

Coen, E. (1991). The role of homeotic genes in flower development and evolution. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **42**, 241–279.

Coen, E. S. & Meyerowitz, E. M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* **353**, 31–37.

Coupland, G. (1995). Genetic and enviromental control of flowering time in *Arabidopsis*. *Trends Genet.* **11**, 393–397.

Garzon, M. (1990). Cellular automata and discrete neural networks. *Physica D* **45**, 431–440.

Glass, L. (1975). Classification of biological networks by their qualitative dynamics. *J. theor. Biol.* **54**, 85–107.

Glass, L. & Kauffman, S. A. (1973). The logical analysis of continuous, non-linear biochemical control networks. *J. theor. Biol.* **39**, 103–129.

Glass, L. & Young, R. E. (1979). Structure and dynamics of neural network oscillators. *Brain Res.* **179**, 207–218.

Goto, K. & Meyerowitz, E. M. (1994). Function and regulation of the *Arabidopsis* floral homeotic gene PISTILLATA. *Genes & Dev.* **8**, 1548–1560.

Gustafson-Brown, C., Savidge, B. & Yanofsky, M. F. (1994). Regulation of the *Arabidopsis* floral homeotic gene APETALA1. *Cell* **76**, 131–143.

Haughn, G. W., Schultz, E. A. & Martinez-Zapater, J. M. (1995). The regulation of flowering in *Arabidopsis thaliana*: meristems, morphogenesis and mutants. *Can. J. Bot.* **73**, 959–981.

Jack, T., Fox, G. L. & Meyerowitz, E. M. (1994). *Arabidopsis* homeotic gene APETALA3 ectopic expression: transcriptional and post-transcriptional regulation determine floral organ identity. *Cell* **76**, 703–716.

Jofuku, K. D., den Boer, B. G. W., Van Montagu, M. & Okamuro, J. K. (1994). Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. *Plant Cell* **6**, 1211–1225.

Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *J. theor. Biol.* **22**, 437–467.

Kauffman, S. A. (1991). Antichaos and adaptation. *Sci. Am.* **265**, 64–71.

Kauffman, S. A. (1993). *The Origins of Order: Self-organization and Selection in Evolution*. Oxford: Oxford University Press.

Kempin, S. A., Savidge, B. & Yanofsky, M. F. (1995). Molecular basis of the *cauliflower* phenotype in *Arabidopsis*. *Science* **267**, 522–525.

Kornfeld, K. (1997). Vulval development in *Caenorhabditis elegans*. *Trends Genet.* **13**, 55–61.

Krizek, B. A. & Meyerowitz, E. M. (1996). The *Arabidopsis* homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. *Development* **122**, 11–22.

Lee, I., Wolfe, D. S., Nilsson, O. & Weigel, D. (1997). A *LEAFY* co-regulator encoded by *UNUSUAL FLORAL ORGANS*. *Curr. Biol.* **7**, 95–104.

Levin, J. Z. & Meyerowitz, E. M. (1995). UFO: An *Arabidopsis* gene involved in both floral meristem and floral organ development. *Plant Cell* **7**, 529–548.

Liu, Z. & Meyerowitz, E. M. (1995). LEUNIG regulates AGAMOUS expression in *Arabidopsis* flowers. *Development* **121**, 975–991.

LOOMIS, W. F. & STERNBERG, P. W. (1995). Genetic networks. *Science* **269**, 649.

MA, H. (1994). The unfolding drama of flower development: recent results from genetic and molecular analyses. *Genes & Dev.* **8**, 745–756.

MANDEL, M. A., GUSTAFSON-BROWN, C., SAVIDGE, B. & YANOFSKY, M. F. (1992). Molecular characterization of the *Arabidopsis* floral homeotic gene APETALA1. *Nature* **360**, 273–277.

MCADAMS, H. H. & SHAPIRO, L. (1995). Circuit simulation of genetic networks. *Science* **269**, 650–656.

MEYEROWITZ, E. M. (1994a). The genetics of flower development. *Sci. Am.* **271**, 40–47.

MEYEROWITZ, E. M. (1994b). Flower development and evolution: New answers and new questions. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5735–5737.

MIZUKAMI, Y. & MA, H. (1992). Ectopic expression of the floral homeotic gene AGAMOUS in transgenic *Arabidopsis* plants alters floral organ identity. *Cell* **71**, 119–131.

MJOLSNESS, E., SHARP, D. H. & REINITZ, J. (1991). A connectionist model of development. *J. theor. Biol.* **152**, 429–453.

OKAMURO, J. K., DER BOER, B. G. W. & JOFUKU, K. D. (1993). Regulation of *Arabidopsis* flower development. *Plant Cell* **5**, 1183–1193.

RAY, A., ROBINSON-BEERS, K., RAY, S., BAKER, S. C., LANG, J. D., PRUSS, D., MILLIGAN, S. B. & GASSER, C. S. (1994). *Arabidopsis* floral homeotic gene BELL (BEL1), controls ovule development through negative regulation of AGAMOUS gene (AG). *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5761–5765.

REINITZ, J. & SHARP, D. H. (1995). Mechanism of eve stripe formation. *Mech. Dev.* **49**, 133–158.

SAKAI, H., MEDRANO, L. J. & MEYEROWITZ, E. M. (1995). Role of SUPERMAN in maintaining *Arabidopsis* floral whorl boundaries. *Nature* **378**, 199–203.

SCHULTZ, E. A. & HAUGHN, W. (1993). Genetic analysis of the floral initiation process (FLIP). in *Arabidopsis*. *Development* **110**, 745–765.

SCHULTZ, E. A., PICKETT, F. B. & HAUGHN, G. W. (1991). The FLO10 gene product regulates the expression domain of homeotic genes AP3 and PI in *Arabidopsis* flowers. *Plant Cell* **3**, 1221–1237.

SHANNON, S. & MEEKS-WAGNER, D. R. (1991). A mutation in the *Arabidopsis* TFL1 gene affects inflorescence meristem development. *Plant Cell* **3**, 877–892.

SHANNON, S. & MEEKS-WAGNER, D. R. (1993). Genetic interactions that regulate inflorescence development in *Arabidopsis*. *Plant Cell* **5**, 639–655.

SIEBURTH, L. E., RUNNING, M. P. & MEYEROWITZ, E. M. (1995). Genetic separation of third and fourth whorl function of AGAMOUS. *Plant Cell* **7**, 1249–1258.

SIMON, R., IGEÑO, M. I. & COUPLAND, G. (1996). Activation of floral meristem identity genes in *Arabidopsis*. *Nature* **384**, 59–62.

SMYTH, D. R., BOWMAN, J. L. & MEYEROWITZ, E. M. (1990). Early flower development in *Arabidopsis*. *Plant Cell* **2**, 755–767.

SPIROV, A. V. (1996). Self-assemblage of gene nets in evolution via recruiting of new netters. *CSTB bulletin* Spring 1996, (http://132.248.34.3/denis_html/cstbspring96/CSTB-spring96-2.0.html).

THEIßEN, G. & SAEDLER, H. (1995). MADS-box genes in plant ontogeny and phylogeny: Haekel's "biogenetic law" revisited. *Current Op. Gen. Dev.* **5**, 628–639.

THIEFFRY, D., SNOUSSI, E. H., RICHELLE, J. & THOMAS, R. (1995). Positive loops and differentiation. *J. Biol. Syst.* **3**, 457–466.

THOMAS, R. (1991). Regulatory networks seen as asynchronous automata: a logical description. *J. theor. Biol.* **153**, 1–23.

THOMAS, R. & D'ARI, R. (1990). *Biological Feedback*. Boca Raton, FL: CRC Press.

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**(2), 247–276.

VERGARA, F. & ALVAREZ-BUYLLA, E. (1997). Developmental genetic basis of the origin and evolution of the homeotic phenotype of *Lacandonia schismatica*, an endemic angiosperm from Chiapas (México), with inverted reproductive whorls. *Keystone Symposium: Evolution of Plant Development*, Taos, NM, p. 28.

WAGNER, A. (1996). Genetic redundancy caused by gene duplications and its evolution in networks of transcriptional regulators. *Biol. Cybern.* **74**, 557–567.

WEIGEL, D. (1995). The genetics of flower development: from floral induction to ovule morphogenesis. *Annu. Rev. Genetics* **29**, 19–39.

WEIGEL, D., ALVAREZ, J., SMYTH, D., YANOFSKY, M. F. & MEYEROWITZ, E. M. (1992). LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* **69**, 843–859.

WEIGEL, D. & MEYEROWITZ, E. M. (1993a). Genetic hierarchy controlling flower development. In: *Molecular Basis of Morphogenesis*, pp. 93–107. New York: Wiley-Liss.

WEIGEL, D. & MEYEROWITZ, E. M. (1993b). LEAFY controls meristem identity in *Arabidopsis*. In: *Cellular Communication in Plants* (Amasino, R. M., ed.), pp. 115–122. New York: Plenum Press.

WEIGEL, D. & MEYEROWITZ, E. M. (1993c). Activation of floral homeotic genes in *Arabidopsis*. *Science* **261**, 1723–1726.

WEIGEL, D. & NILSSON, O. (1995). A developmental switch sufficient for flower initiation in diverse plants. *Nature* **377**, 495–500.

WILKINSON, M. D. & HAUGHN, G. W. (1995). UNUSUAL FLORAL ORGANS controls meristem identity and organ primordia fate in *Arabidopsis*. *Plant Cell* **7**, 1485–1499.

WOLPERT, L. (1996). One hundred years of positional information. *Trends Genet.* **12**, 359–364.

YANG, C. H., CHEN, L. J. & SUNG, Z. R. (1995). Genetic regulation of shoot development in *Arabidopsis*, role of the EMF genes. *Dev. Biology* **169**, 421–435.

YANOFSKY, M. F. (1995). Floral meristems to floral organs: genes controlling early events in *Arabidopsis* flower development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**, 167–188.

YANOFSKY, M. F., MA, H., BOWMAN, J. L., DREWS, G. N., FELDMANN, K. A. & MEYEROWITZ, E. M. (1990). The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature* **346**, 35–40.

ZUCKERKANDL, E. (1994). Molecular pathways to parallel evolution: I. Gene nexuses and their morphological correlates. *J. Mol. Evol.* **39**, 661–678.

APENDICE 4

BIOINFORMATICS

Vol. 15 nos 7/8 1999
Pages 593-606

Genetic control of flower morphogenesis in Arabidopsis thaliana: a logical analysis

Luis Mendoza[1], Denis Thieffry[2] and Elena R. Alvarez-Buylla[1]

[1]Instituto de Ecología, UNAM, Ap. Postal 70-275, Coyoacán, DF, CP04510, Mexico and [2]Département de Biologie Moléculaire, Université Libre de Bruxelles, Rue des Chevaux, 67, B1640 Rhode-Saint-Genèse, Belgium

Received on October 30, 1998, revised on February 22, 1999, accepted on February 23, 1999

Abstract

Motivation: *A large number of molecular mechanisms at the basis of gene regulation have been described during the last few decades. It is now becoming possible to address questions dealing with both the structure and the dynamics of genetic regulatory networks, at least in the case of some of the best-characterized organisms. Most recent attempts to address these questions deal with microbial or animal model systems. In contrast, we analyze here a gene network involved in the control of the morphogenesis of flowers in a model plant, Arabidopsis thaliana.*

Results: *The genetic control of flower morphogenesis in Arabidopsis involves a large number of genes, of which 10 are considered here. The network topology has been derived from published genetic and molecular data, mainly relying on mRNA expression patterns under wild-type and mutant backgrounds. Using a 'generalized logical formalism', we provide a qualitative model and derive the parameter constraints accounting for the different patterns of gene expression found in the four floral organs of Arabidopsis (sepals, petals, stamens and carpels), plus a 'non-floral' state. This model also allows the simulation or the prediction of various mutant phenotypes. On the basis of our model analysis, we predict the existence of a sixth stable pattern of gene expression, yet to be characterized experimentally. Moreover, our dynamical analysis leads to the prediction of at least one more regulator of the gene LFY, likely to be involved in the transition from the non-flowering state to the flowering pathways. Finally, this work, together with other theoretical and experimental considerations, leads us to propose some general conclusions about the structure of gene networks controlling development.*

Contact: *lams@servidor.unam.mx; denis@dbm.ulb.ac.be; abuylla@servidor.unam.mx*

Introduction

The study of the molecular basis underlying development and morphogenesis is an active area of biological research. Most studies use animals as model systems, mainly *Drosophila melanogaster*, *Caenorhabditis elegans*, *Xenopus laevis* and *Mus musculus*. More recently, plants have also become the focus of molecular geneticists, and new insights into the genetic and molecular mechanisms underlying plant development have been gained through the study of vegetal model systems, especially *Arabidopsis thaliana*. Flower development has received special attention, and many genes involved in this process have now been characterized at the genetic and molecular levels (for a review, see Weigel, 1995). Though still fragmentary, published data already support the existence of a complex regulatory network controlling flower morphogenesis in *Arabidopsis*.

A qualitative model for flower morphogenesis in *A.thaliana* was recently proposed (the 'NET model'; Mendoza and Alvarez-Buylla, 1998), together with a preliminary analysis of its dynamical behavior. Here, we develop an extensive analysis of the relationship between the architecture of this gene network (in terms of feedback circuits) and its dynamical behavior. Basic formal prerequisites are introduced hereafter, whereas the experimental background supporting the model is summarized in the last part of this introductory section. A second section is devoted to the presentation of the NET model and its feedback circuit analysis, whereas a third section contains a thorough discussion of the biological relevance of the results, followed by some conclusions and further prospects.

Feedback circuits and logical formalization

Experimental and theoretical biologists have long been aware of the important biological and dynamical roles of genetic 'feedback circuits' or 'feedback loops' (see, for example, Monod and Jacob, 1961; Rosen, 1968). Feedback circuits are defined as circular chains of interactions, such that each element of a circuit influences its own future level. Whenever specific signs can be associated to each interaction, a given circuit can be unambiguously classified in one of the two following classes: either each element of the circuit exerts a positive direct or indirect effect on itself and the circuit is said to be 'positive', or each element of the circuit exerts a negative effect on itself and the circuit is said to be 'negative'. The sign of a circuit is easily determined by

checking the parity of the number of negative interactions involved. If this number is even or zero, the circuit is positive, but if the number is odd the circuit is negative. Positive- and negative-feedback circuits have different dynamical and biological properties (for a review, see Thomas *et al.*, 1995). From a dynamical point of view, positive circuits yield multistationarity, whereas negative circuits generate (damped or sustained) oscillatory behavior. Biologically, positive circuits are necessary to accomplish differentiation (i.e. alternative patterns of gene expression), whereas negative circuits are necessary to generate homeostasis. Formulated by Thomas about 20 years ago, these conjectures have recently been demonstrated in a general formal context (Plahte *et al.*, 1995; Gouzé, 1998; Snoussi, 1998).

Formally, a gene network can be represented by a set of differential equations. Most of the time, however, little is known about the precise shapes of regulatory interactions, or about the values of the various parameters. This situation led several authors to propose a qualitative or 'logical' formalization of sets of interacting genes (e.g. Kauffman, 1969; Glass and Kauffman, 1973). In this paper, we use the formalism introduced by Thomas, which has the advantage of being particularly suited to evaluating the roles of the various feedback circuits present in a given network.

Following Thomas, we associate a logical variable (noted by a lowercase letter, e.g. 'x') and a logical function (noted by an uppercase letter, e.g. 'X') to each gene of a network. The logical variable represents the actual expression state of a gene, while its logical function indicates the gene's future state as determined by the action of the genes (activators and inhibitors) that regulate it. The expression states are characterized by a limited number of integer values (0, 1, 2, ...), but 'threshold' values $[s^{(1)}, s^{(2)}, \ldots]$ that separate the expression states are also considered. Finally, logical parameters (noted by Ks and appropriate indices) are introduced to qualify the weight of each interaction, or combination of interactions, on the expression of a regulated gene (Snoussi, 1989). For example, K_a represents the lowest expression state of a gene a, $K_{a,b}$ indicates the effect of an activation of gene b over a, $K_{a,bc}$ the combined activatory effect of genes b and c over the state of a, and so forth. In the simplest cases, our logical variables, functions and parameters take one of the three values {0, s, 1}, where s stands for the threshold between '0' and '1' (for a review, see Thomas, 1991). A simple case of two interacting genes and the dynamical consequences of different values of Ks is shown in Figure 1.

In the context of this logical formalism, an n-element gene network can be described by two matrices: the first ($n \times n$) matrix contains the signs (and eventually the thresholds) of all interactions, the second ($n \times 2^n$) matrix contains the values of the logical parameters. The state of the system is represented by a vector of dimension n. Whenever this state vector and its image (i.e. the vector formed by the values of the corresponding functions) are equal, there is a steady state in the system. Because we explicitly take into account threshold values, we will distinguish between 'regular' and 'singular' logical states. Regular states are those including only zeros and ones, while singular states may include one or more threshold values. Finally, the discrete approach used here guarantees the explicit calculation of numerical values of all parameters and functions. Therefore, it is possible to identify all the steady states of the system.

A much more interesting approach, however, consists of using the notion of feedback circuit. Clearly, the effect of a circuit will not depend only on the mere existence of the relevant interactions, but also on their relative strengths. Within appropriate parametric ranges, positive circuits generate multistationarity and negative circuits generate homeostasis. Therefore, if a circuit is found to generate its associated dynamical behavior, we will say that the circuit is 'functional'. That is, a positive circuit is functional if it actually produces different stable steady states separated by unstable steady states. Conversely, a negative circuit is functional when it generates damped or sustained oscillatory behavior. Figure 1 presents a simple example of how circuit functionality depends on the parametric values (Ks). In this example, the circuit is functional only for the first set of values.

In the context of Thomas' logical formalism, it is possible to compute the parameter constraints making any single feedback circuit functional. Moreover, it was found that whenever such a circuit is functional, it generates a singular ('characteristic') steady state located on the thresholds of the interactions forming the circuit. More precisely, when the circuit is positive, its characteristic state is always unstable (typically a saddle point in two dimensions) and stands on a separatrix. When the circuit is negative, its characteristic state can be either stable or unstable (typically a focus in two dimensions). Snoussi and Thomas (1993) have demonstrated that only those singular states which are circuit characteristic can be steady, thus making superfluous the scanning of all 'non-characteristic' singular steady states.

The logical formalism outlined above enables us to dissociate a complex network into a well-defined set of feedback circuits and check their dynamical roles individually, yet keeping complete control of the ways in which these circuits are interconnected. The scope of this paper consists precisely in applying this approach to the analysis of the genetic network controlling flower morphogenesis in *Arabidopsis*. Before that, however, we still need to summarize the basic genetic and molecular data supporting the 'NET model'.

Flower development in Arabidopsis

Mature flowers of *Arabidopsis* display a stereotyped architecture that consists of four concentric whorls of organs. These are from the external to the internal whorls: four sepals

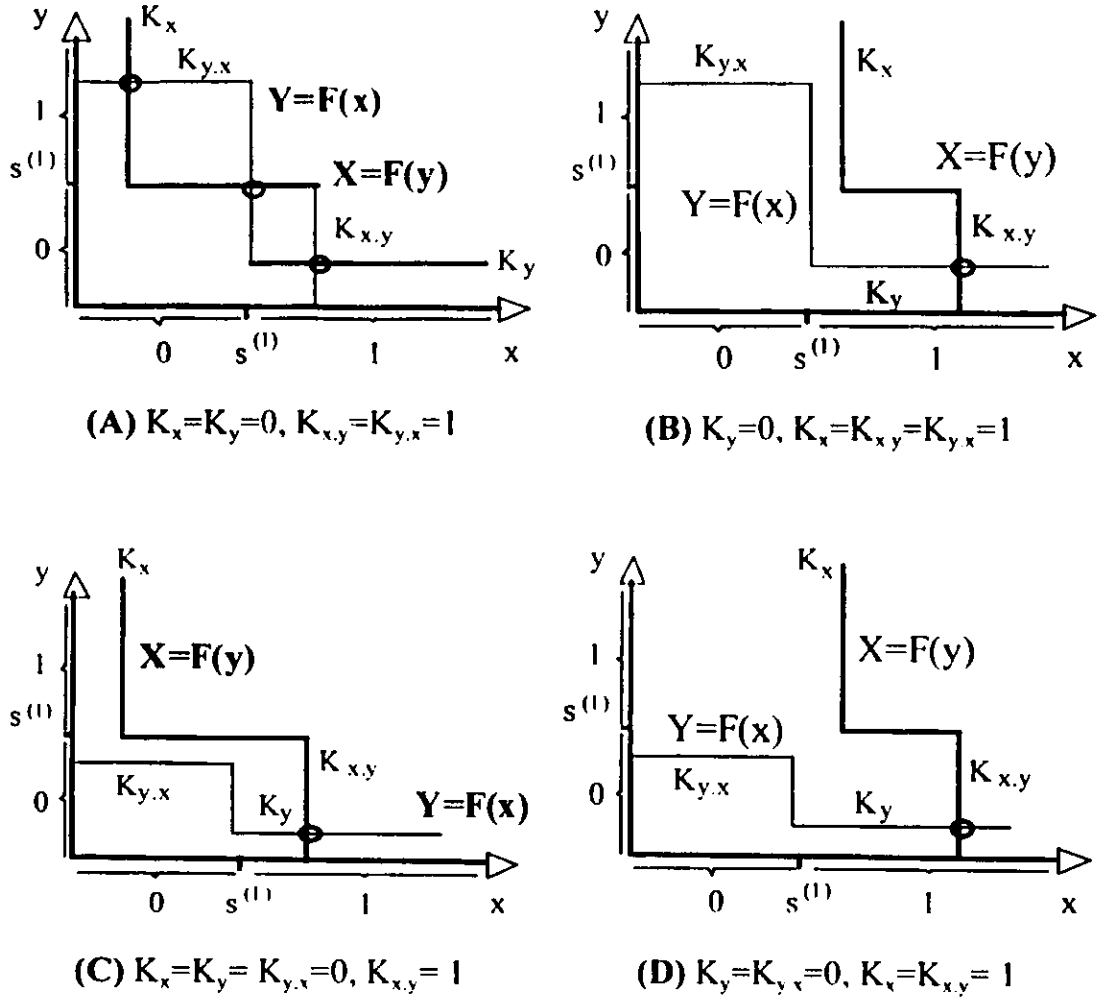


Fig. 1. Schematic representation of the parameter constraints on circuit functionality for a two-element cross-inhibitory positive circuit. The step curves give the value of the logical function X (respectively Y) as a function of the value of the logical variable y (respectively x). In biological terms, the step curve $X = F(y)$ represents the level of expression of gene X in response to increasing concentrations of the regulatory product y [similarly for $Y = F(x)$]. These step curves correspond to the 'nullclines' in the differential formalism (ODEs) and, consequently, intersections (circled) between the step curves give the steady states of the system. The number of steady states depends on the parameter values. The four diagrams correspond to four different combinations of parameter values. As we account only for two states corresponding to maximal and minimal transcriptional activities, the variables, functions and parameters can take only one of the two regular values 0 and 1, or, in special cases, the threshold value s. The logical parameters (Ks) determine the amplitude of each function F, accordingly with the value of the relevant variable. For example, K_x stands for the level of expression of gene X (function X) in the presence of high concentrations of its repressor, the regulatory product y (variable y); its value is set to 0 in (A) and (C), whereas it is set to 1 in (B) and (D) [leftmost vertical segment of $X = F(y)$]. $K_{x.y}$ stands for the level of expression of gene X in the absence of the repressor y; it has the value 1 in all four illustrated cases [rightmost vertical segment of $X = F(y)$]. K_y stands for the level of expression of gene Y in the presence of its repressor x; its value is set to 0 in all four examples [rightmost horizontal segment of $Y = F(x)$]. Finally, $K_{y.x}$ stands for the expression of gene Y in the absence of its repressor x; it has the value 1 in cases (A) and (B), and the value 0 in case (C) and (D) [leftmost horizontal segment of $X = F(y)$]. Note that multistationarity is found only in case (A), where $K_x = K_y = 0$ and $K_{x.y} = K_{y.x} = 1$.

(whorl 1), four petals (whorl 2), six stamens (whorl 3) and two fused carpels (whorl 4). Many mutations that disrupt normal flower development, and therefore alter the final floral architecture, have been already reported. The analysis of homeotic floral mutants was the basis for the so-called 'ABC' combinatorial model of flower morphogenesis that is

widely used to describe the morphology of wild-type and mutant *Arabidopsis* flowers. This model (Coen and Meyerowitz, 1991) postulates the existence of three different genetic functions or activities (A, B and C), each of them present in two adjacent whorls. According to this model, each whorl requires a specific combination of genetic activities: activity A in the first whorl determines sepal identity; activities A and B combined in the second whorl determine petal identity; activities B and C combined in the third whorl determine stamen identity; finally, activity C alone in the fourth whorl determines carpel identity. Furthermore, the ABC model postulates a mutual inhibition between activities A and C.

The ABC model also accounts for the phenotype of homeotic floral mutants (see Figure 2). Nevertheless, we still lack a thorough understanding of the molecular mechanisms underlying the ABC model. Recent genetic and molecular studies have already uncovered over a dozen genes involved in the control of flower morphogenesis. Some of these are clearly associated with each of the ABC functions described above. The gene *APETALA1* (*AP1*) is associated with function A (Mandel *et al.*, 1992; Bowman *et al.*, 1993), *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) are B genes (Krizek and Meyerowitz, 1996), and *AGAMOUS* (*AG*) is the only gene associated with the C function (Yanofsky *et al.*, 1990). Moreover, these genes have been found to cross-regulate with many others and thus form a complex regulatory network summarized in Figure 3. The molecular and morphological data underlying this net model are shown in Table 1.

Feedback circuit analysis of the NET model

A quick look at the interaction matrix in Figure 3 indicates that *EMF1*, *LUG*, *UFO* and *SUP* do not receive any regulatory input, although each of them regulates at least one gene of the network. Closer inspection leads to the identification of eight feedback circuits involving from one to four of the remaining six genes. Moreover, considering the variables involved, we can split these circuits into two groups: *TFL1-LFY*, *LFY-API*, *API-AG*, *LFY-AG-API* and *TFL1-AG-API-LFY*, on the one hand; *AP3*, *PI* and *AP3-PI*, on the other hand. Our 10-variable network can thus be considered as the combination of two smaller subnetworks, one formed by *AP3* and *PI*, and the other one formed by *TFL1*, *LFY*, *API* and *AG*. The other four genes (*EMF1*, *LUG*, *UFO* and *SUP*) can be represented as input variables. Note, however, that the two subnetworks are not isolated, because the smallest receives two inputs from *LFY*, which is part of the larger subnetwork. In the next two sections, we consider the dynamics of each of these two subnetworks and then reconstitute the dynamics of the whole network.

Table 1. Experimental data that support the gene interactions included in the NET model. For a detailed account, see Mendoza and Alvarez-Buylla (1998). The interactions supported by Mendoza and Alvarez-Buylla (1998) are not based on new experimental data, but rather are inferred from previous analysis of the patterns of gene expression

Interactions	Main evidence	Main references
AG —| *API*	*API* mRNA accumulates uniformly in *ag*-1 mutant flowers	Gustafson-Brown *et al.* (1994)
API —| *AG*	Sepals are replaced by carpels, and petals by stamens in *ap1* mutants. *AG* mRNA found in all the flower primordium of *ap1-1* plants	Bowman *et al.* (1993) Weigel and Meyerowitz (1993)
API —> *LFY*	Reduction of *LFY* mRNA in *ap1 cal* double mutants (independent pathways)	Weigel and Nilsson (1995) Bowman *et al.* (1993) Kempin *et al.* (1995)
AP3/PI —> *AP3/PI*	*AP3* and *PI* mRNA levels are not maintained in *ap3*-3, *pi*-1 or double mutants Co-immunoprecipitation of AP3 and PI proteins	Goto and Meyerowitz (1994) Jack *et al.* (1992)
EMF1 —| *API LFY*	Inferred by morphological evidence that *EMF1* inhibits the flowering promoting genes	Mendoza and Alvarez-Buylla (1998)
EMF1 —> *TFL1*	Inferred by morphological evidence that *EMF1* activates the late late-flowering genes	Mendoza and Alvarez-Buylla (1998)
LFY —> *AG*	Early expression of *AG* is abnormally low in *lfy*-6 flowers	Weigel and Meyerowitz (1993)
LFY —> *API*	*API* mRNA delayed in *lfy* mutants. Earlier *API* promoter induction in plants overexpressing *LFY*	Weigel and Nilsson (1995) Parcy *et al.* (1998)
LFY —> *AP3*	Amount and domain of *AP3* expression reduced in *lfy*-6 mutants	Weigel and Meyerowitz (1993)
LFY —> *PI*	Amount and domain of *PI* expression reduced in *lfy*-6 mutants	Weigel and Meyerowitz (1993)
LFY —| *TFL1*	Plants overexpressing *LFY* are very similar to *tfl1* mutants	Weigel and Nilsson (1995)
LUG —| *AG*	Ectopic expression of *AG* in *lug*-1 mutants	Liu and Meyerowitz (1995)
SUP —| *AP3*	Ectopic expression of *AP3* in *sup*-1 mutants	Sakai *et al.* (1995)
SUP —| *PI*	Contrary to wild type, *PI* expression is not reduced in the center of *sup*-1 flowers	Goto and Meyerowitz (1994)
TFL1 —| *AG*	Inferred from morphological evidence. Double mutants *ap1-1 ap2-2* have a disrupted C activity, which is rescued with the addition of *tfl1* mutation	Mendoza and Alvarez-Buylla (1998)
TFL1 —| *LFY*	Precocious appearance of floral buds expressing *LFY* in *tfl1*-2 plants	Weigel *et al.* (1992)
UFO —> *AP3*	*AP3* protein and messenger levels reduced in *ufo*-2 plants	Levin and Meyerowitz (1995)
UFO —> *PI*	*PI* mRNA reduced in early stages of flower development in *ufo*-2 plants	Levin and Meyerowitz (1995)

Table 2. General state table for the subnetwork *AP3-PI*. Variables p and i stand for the levels of activity of the regulatory products of genes *AP3* and *PI*, respectively, whereas functions P and I stand for the levels of expression of the corresponding genes. Each row correspond to a specific combination of the possible values (0 or 1) for the variables p and i. In biological terms, the rows thus correspond to different combinations between the two levels of activity for the two regulatory products. For example, '10' in the penultimate row means that the first regulatory product (*AP3*) is present at a high level, whereas the second product (*PI*) is absent (or present at a negligible level). Under the columns P and I are given the corresponding levels of gene expression in terms of the logical parameters (Ks). K_p stands for the basal expression of gene *AP3*, $K_{p,p}$ for its expression in the presence of its own product (autoregulation), $K_{p,i}$ for its expression in the presence of the sole product of gene *PI* and $K_{p,pi}$ for its expression in the presence of both regulatory products. The meaning of parameters K_i, $K_{i,p}$, $K_{i,i}$ and $K_{i,pi}$ is analogous to that associated with *AP3*. Clearly, different values of the logical parameters (Ks) may give rise to different dynamics

p	i	P	I
0	0	K_p	K_i
0	1	$K_{p,i}$	$K_{i,i}$
1	0	$K_{p,p}$	$K_{i,p}$
1	1	$K_{p,pi}$	$K_{i,pi}$

Table 3. Feedback circuits of the *AP3-PI* subnetwork and the corresponding functionality constraints. This subnetwork formed by *AP3* and *PI* (variables p and i) contains three circuits that are indicated in the leftmost column. The K columns give the parameter constraints to be fulfilled for each feedback circuit to be functional in a particular region of the phase space ('Domain' column). A '–' means that there are no constraints on the corresponding parameter. The values selected on the basis of experimental data (see the text) are indicated in the lowermost row. Note that for these values, the two autoregulatory circuits are functional only in a limited region of the phase space, whereas the two-element circuit is functional in the entire state space. Notation of the parameters is as in Table 2

Loop	Domain	K_p	$K_{p,p}$	$K_{p,i}$	$K_{p,pi}$	K_i	$K_{i,p}$	$K_{i,i}$	$K_{i,pi}$
p (+)	S[1]	0	–	0	1	–	–	–	–
	S[0]	0	1	–	1	–	–		–
i (+)	[1]S	–	–	–	–	0	0	–	1
	[0]S	–	–	–	–	0	–	1	1
pi (+)	SS	0	–	–	1	0	–		1
Selected values		0	0	0	1	0	0	0	1

AP3-PI subnetwork

The analysis of this subsystem is straightforward as it is made of only two elements. The general state table is given in Table 2.

Depending on the value of the logical parameters, this table covers various dynamics. Note that the system involves three positive feedback circuits. The parameter constraints that make each of these circuits functional are given in Table 3.

Remember that *AP3* and *PI* genes are both associated with the B function. Both genes are expressed in petals and stamens, but not in sepals and carpels. Moreover, it is known that *AP3* and *PI* regulate their own, as well as each other's, transcription through the formation of a heterodimer of their protein products (Goto and Meyerowitz, 1994). In logical terms, these regulatory interactions can be represented by $K_p = K_{p,p} = K_{p,i} = 0$, $K_{p,pi} = 1$ for the parameters associated with the regulation of the expression of gene *AP3* (indices p), and by $K_i = K_{i,i} = K_{i,p} = 0$, $K_{i,pi} = 1$, for the parameters associated with *PI* (indices i). Once included in Table 2, these parameter values give two stable states, [00] and [11], corresponding to the absence and presence of B function, respectively. For the parameter values selected, inspection of Table 3 reveals that the two-variable positive circuit is fully functional, whereas the one-variable positive circuits are functional only in some limited regions of the variable space. Additionally, the system comprises a third unstable steady state, namely [ss], which is the characteristic state of the two-element positive circuit.

TFL1-LFY-AP1-AG subnetwork

Analysis of the four-variable subnetwork is somewhat more complex, but still straightforward. By hand or using a computer program (Thieffry *et al.*, 1993), we obtain Table 4 (general state table) and 5 (functionality constraints for the five feedback circuits of the system).

For the four-variable subnetwork, we have also derived some of the parameter values from gene expression data in the outer and inner whorls of the flower. First, we need to account for a default 'non-flowering state'. Considered as a flower suppressor, *TFL1* is expressed in plant apices before flower development (Ratcliffe *et al.*, 1998). In contrast, *LFY* expression has been shown to be very low before flowering, reaching its highest level upon flower induction (Nilsson *et al.*, 1998). Finally, *AP1* and *AG* genes are known to be specifically expressed during flowering. This leads us to associate the state 1000 (vector notation for *TFL1* = 1, *LFY* = 0, *AP1* = 0, and *AG* = 0) with the non-flowering state. Looking at the corresponding (tenth) row in Table 4, we find that state 1000 will be steady if $K_{t,l} = 1$, and $K_l = K_{a,g} = K_{g,a} = 0$, where indices t, l, a and g stand for *TFL1*, *LFY*, *AP1* and *AG*, respectively.

In contrast with the default state mentioned above, the flower suppressor gene *TFL1* should be 'off' during flower morphogenesis and in each of the developing flower organs. In addition, as *LFY* is known to be only transiently expressed in most of the developing flower (Weigel *et al.*, 1992), we consider that its final steady value is '0' in the flower organs. Finally, taking into account the association of *AP1* with A function, that of *AG* with C function, as well as their probable mutual inhibition, we are led to associate the states 0010 and 0001 with A and C functions, respectively. On the basis of Table 4, the first of these states will be steady if $K_{t,l} = K_{l,ta} = K_{g,t} = 0$ and $K_{a,g} = 1$, whereas the second will be steady for $K_{t,l} = K_{l,l} = K_a = 0$ and $K_{g,a} = 1$.

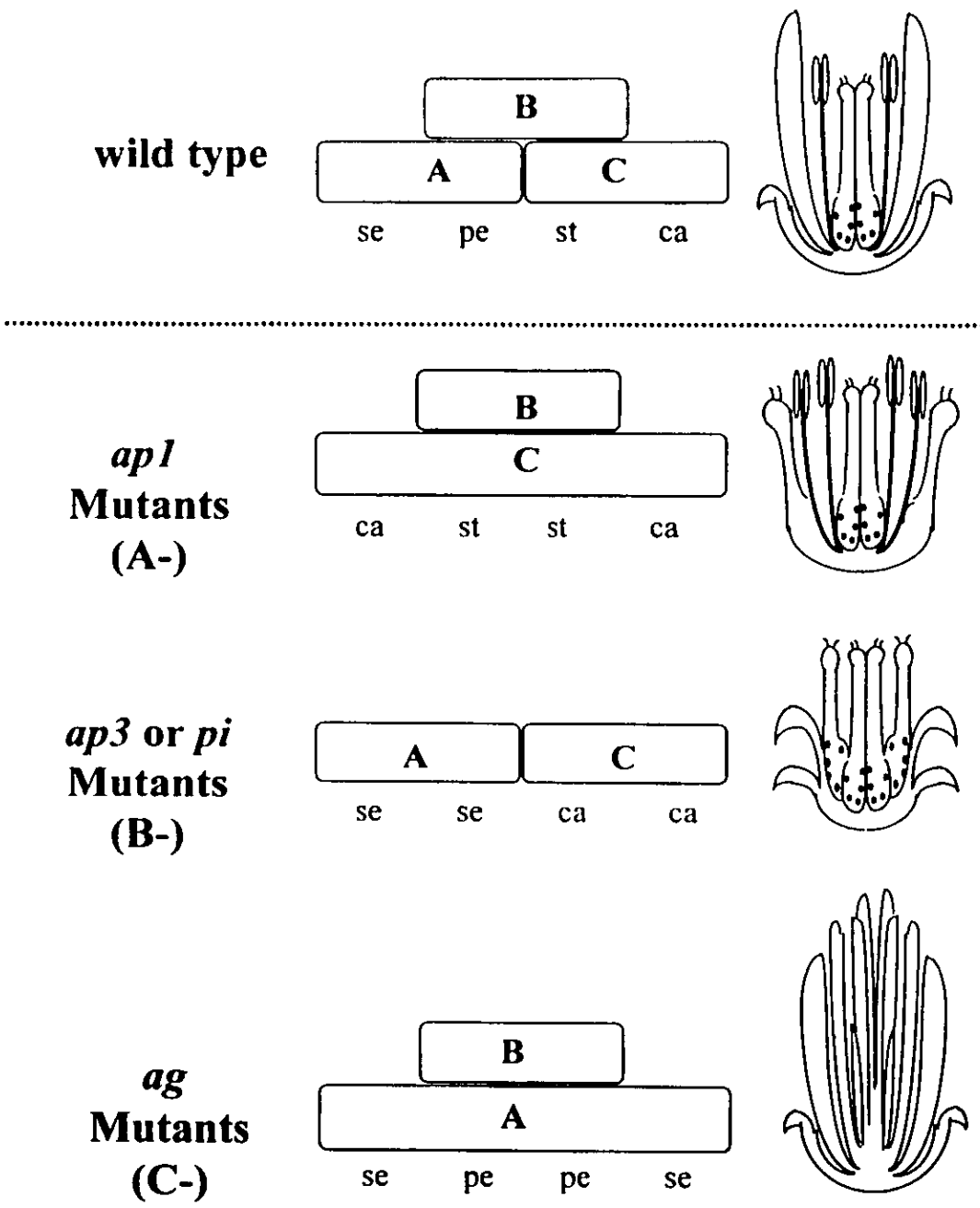
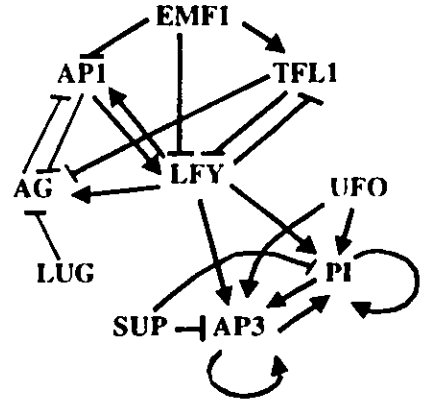


Fig. 2. Schematic representation of the wild-type *Arabidopsis* flower and the ABC model. The mutant types are in the leftmost column. The domains of expression of the ABC genetic activities in the four floral whorls are shown in the central column of the diagram. Longitudinal sections of wild-type (upper part) and mutant (lower part) *Arabidopsis* flowers are schematized in the rightmost column. For example, mutations in the *AP1* gene eliminate activity A, and activity C expands to the first two whorls; this causes the *ap1* mutant flowers to have carpels, stamens, stamens and carpels. se, sepals; pe, petals; st, stamens; ca, carpels.



	EMF1	TFL1	LFY	AP1	AG	LUG	UFO	AP3	PI	SUP
EMF1	0	0	0	0	0	0	0	0	0	0
TFL1	+	**0**	-	**0**	**0**	0	0	0	0	0
LFY	-	-	**0**	+	**0**	0	0	0	0	0
API	-	**0**	+	**0**	-	0	0	0	0	0
AG	0	-	+	-	**0**	-	0	0	0	0
LUG	0	0	0	0	0	0	0	0	0	0
UFO	0	0	0	0	0	0	0	0	0	0
AP3	0	0	+	0	0	0	+	+	+	-
PI	0	0	+	0	0	0	+	+	+	-
SUP	0	0	0	0	0	0	0	0	0	0

Fig. 3. Diagram (upper part) and matrix (lower part) describing the interactions established among 10 genes that regulate flower morphogenesis in *Arabidopsis*. Activatory and repressory relationships are indicated by arrows and blunt lines, respectively. In the matrix, the rows give the sign of the transcriptional response of each gene (italics) to the gene products indicated in the column headings. Positive and negative transcriptional interactions are represented by plus (+) and minus (−) signs, respectively. Zeros indicate the absence of a transcriptional interaction. Focusing on the genes involved in feedback circuits, two subnetworks can be readily distinguished: one involving the genes *AP3* and *PI*, and the other involving the genes *TFL1*, *LFY*, *API* and *AG*. These subnetworks are highlighted in the matrix.

Note that the consideration of these different steady states leads to conflicting values for $K_{t,l}$ and $K_{a,g}$. In fact, this apparent contradiction disappears when we include the input variables in our description. Specifically, *EMF1* is a key flowering repressing element (Yang *et al.*, 1995), that activates *TFL1* and represses *LFY* and *API* (see Table 1). Rather than directly introducing an additional index in the parameter notation, we hereafter consider two sets of parameter values, corresponding to the presence and the absence of *EMF1*, respectively.

In the first case, where the flower inhibitor is present (*EMF1* = 1), we have $K_{t,l} = 1$ and $K_{a,g} = 0$, giving rise to a unique, regular steady state, namely [1000] (all other Ks = 0, except $K_{a,lg} = K_{g,tl} = K_{g,ta} = K_{g,la} = K_{g,tla} = 1$). This state corresponds to the default state. Looking back to Table 5, we find that all three two-element circuits are functional, but only in a limited region of the variable space. Therefore, they do not generate any complete separatrix, while the system includes a single attractor representing the non-flowering stage.

In the second case, with the flower inhibitor absent (*EMF1* = 0), we have $K_{t,l} = 0$ and $K_{a,g} = 1$ (all other parameters identical to the preceding case). This leads to the two regular stable states [0010] and [0001], representing functions A and C, respectively. In this case, only one circuit is functional (see Table 5), corresponding to the mutual exclusion of *API* and *AG*. This circuit generates a separatrix across the variable space, defining two basins of attraction, each draining towards one of the stable states [0010] or [0001]. An additional, singular, steady state [00ss] is found on this separatrix. This singular state is unstable because it is generated by the positive circuit *API-AG*.

Table 4. General state table for the *TFL1, LFY, AP1* and *AG* subnetwork. Variables t, l, a and g stand for the levels of activity of the regulatory products of genes *TFL1*, *LFY*, *AP1* and *AG*, respectively, whereas functions T, L, A and G stand for the levels of expression of the corresponding genes. Notation of the parameters is as in Table 2

t	l	a	g	T	L	A	G
0	0	0	0	$K_{t\,l}$	$K_{l\,t}$	$K_{a\,g}$	$K_{g\,ta}$
0	0	0	1	$K_{t\,l}$	$K_{l\,t}$	K_{a}	$K_{g\,ta}$
0	0	1	0	$K_{t\,l}$	$K_{l\,ta}$	$K_{a\,g}$	$K_{g\,t}$
0	0	1	1	$K_{t\,l}$	$K_{l\,ta}$	K_{a}	$K_{g\,t}$
0	1	0	0	K_{t}	$K_{l\,t}$	$K_{a\,lg}$	$K_{g\,tla}$
0	1	0	1	K_{t}	$K_{l\,t}$	$K_{a\,l}$	$K_{g\,tla}$
0	1	1	0	K_{t}	$K_{l\,ta}$	$K_{a\,lg}$	$K_{g\,tl}$
0	1	1	1	K_{t}	$K_{l\,ta}$	$K_{a\,l}$	$K_{g\,tl}$
1	0	0	0	$K_{t\,l}$	K_{l}	$K_{a\,g}$	$K_{g\,a}$
1	0	0	1	$K_{t\,l}$	K_{l}	K_{a}	$K_{g\,a}$
1	0	1	0	$K_{t\,l}$	$K_{l\,a}$	$K_{a\,g}$	K_{g}
1	0	1	1	$K_{t\,l}$	$K_{l\,a}$	K_{a}	K_{g}
1	1	0	0	K_{t}	K_{l}	$K_{a\,lg}$	$K_{g\,la}$
1	1	0	1	K_{t}	K_{l}	$K_{a\,l}$	$K_{g\,la}$
1	1	1	0	K_{t}	$K_{l\,a}$	$K_{a\,lg}$	$K_{g\,l}$
1	1	1	1	K_{t}	$K_{l\,a}$	$K_{a\,l}$	$K_{g\,l}$

Table 5. Feedback circuits of *TFL1, LFY, AP1* and *AG* subnetwork and the corresponding functionality constraints. This subnetwork contains five circuits that are indicated in the leftmost column. The K columns give the parameter constraints to be fulfilled for each feedback circuit to be functional in a particular region of the phase space ('Domain' column). For the parameter notation, see Table 2

Loop	Domain	K_t	$K_{t\,l}$	K_l	$K_{l\,t}$	$K_{l\,a}$	$K_{l\,ta}$	K_a	$K_{a\,l}$	$K_{a\,g}$	$K_{a\,lg}$	K_g	$K_{g\,t}$	$K_{g\,l}$	$K_{g\,a}$	$K_{g\,tl}$	$K_{g\,ta}$	$K_{g\,la}$	$K_{g\,tla}$
tl (+)	SS[1][01]	0	1	0	–	0	1	–	–	–	–	–	–	–	–	–	–	–	–
	SS[0][01]	0	1	0	1	–	1	–	–	–	–	–	–	–	–	–	–	–	–
la (+)	[1]SS[1]	–	–	0	–	1	1	0	1	–	1	–	–	–	–	–	–	–	–
	[1]SS[0]	–	–	0	–	1	1	0	–	0	1	–	–	–	–	–	–	–	–
	[0]SS[1]	–	–	0	0	–	1	0	1	–	1	–	–	–	–	–	–	–	–
	[0]SS[0]	–	–	0	0		1	0	–	0	1			–					
ag (+)	[1][1]SS	–	–	–	–	–	–	0	0	–	1	0	–	0	–	–	–	1	1
	[1][0]SS	–	–	–	–	–	–	0		1	1	0	–	–	1	–	1	1	1
	[0][1]SS	–	–	–	–	–	–	0	0	–	1	0	0	0	–	0	–	–	1
	[0][0]SS	–	–	–	–	–	–	0		1	1	0	0	–			1	–	1
lag (–)	[1]SSS	–	–	0	–	1	1	0	–	–	1	0	–	1	0	1	–	1	1
	[0]SSS	–	–	0	0	–	1	0	–	–	1	0	0	–	0	1	0	–	1
tlag (–)	SSSS	0	1	0	–	–	1	0	0	1	1	0	1	0	0	1	1	0	1

Dynamics of the whole network

On the basis of the results obtained for the two subnetworks, it is now possible to derive some key features of the dynamics of the whole system. The steady states of the whole network can be constructed by combining the steady states of the two subnetworks and relevant values of the four input variables. Since *EMF1* is a floral repressor gene, it is considered active only in non-floral attractors. In addition, it is known that *UFO*, *LUG* and *SUP* regulate the fine spatial expression patterns of the organ identity genes (Levin and Meyerowitz, 1995; Liu and Meyerowitz, 1995; Sakai *et al.*, 1995). Consequently, these three genes play a role only in the transient activation patterns of the organ identity genes *AG*,

AP3 and *PI*, but not in the establishment of the floral organs themselves. Therefore, the combination of the stable states found for the two subnetworks leads to the following six attractors:

[00] + [0010] → [0001000000] (A = sepals)
[11] + [0010] → [0001000110] (AB = petals)
[11] + [0001] → [0000100110] (BC = stamen)
[00] + [0001] → [0000100000] (C = carpels)
[00] + [1000] → [1100000000] (flower inhibition)
[11] + [1000] → [1100000110] (6th attractor)

using the vector notation [*pi*] + [*tlag*] → [*etlagufpis*], where *e*, *t*, *l*, *a*, *g*, *u*, *f*, *p*, *i* and *s* stand for *EMF1*, *TFL1*, *LFY*, *API*, *AG*, *LUG*, *UFO*, *AP3*, *PI* and *SUP*, respectively.

This set of attractors can be obtained for several combinations of parameter values, e.g. all values equal to zero, except $K_{t,el}$, $K_{a,eg}$, $K_{g,tau}$, $K_{i,pis}$, $K_{i,pis}$, and the parameters necessarily greater or equal to these. Note that these parameters take into account the input variables e, u and s. As the two modules behave largely independently, we obtain all possible combinations of their respective steady states, including the sixth global attractor that does not correspond to any experimentally characterized cell type.

The methodology used here allows for the location of all the steady states of gene activation implied by the network connectivity, but in contrast to dynamic systems, this methodology does not identify the transitory activation pathways that lead to each attractor. The sequence of activation and repression of the input variables *EMF1*, *UFO*, *SUP* and *LUG*, as well as the timing of response of the genes in the circuit, determine which attractor is attained. In order to incorporate this information in the model, we need biological data such as the time needed for synthesis and degradation of regulatory factors. This information would enable us to obtain a flux diagram similar to that in Figure 4, which shows possible routes of gene activation starting at the vegetative stage and reaching the four genetic activation states characteristic of the floral organs.

Five of the six attractors have a clear biological interpretation. The first one, represented by [0001000000], corresponds to the exclusive activation of *AP1*, i.e. function A in the ABC model. This activation pattern is found in the first whorl of flowers, formed by sepals. The second attractor, i.e. [0001000110], represents the stable activation of genes *AP1*, *AP3* and *PI*. These genes are responsible for A and B functions, whose combination determines petal formation in the second whorl. The third attractor [0000100110] denotes the activation of *AG*, *AP3* and *PI*, representing the presence of both B and C functions. This state corresponds to the differentiation of stamens in the third whorl. The fourth attractor [0000100000] corresponds to the exclusive activation of *AG*, the only reported C function gene. This pattern is found in cells of the fourth whorl which give rise to carpels. The last biologically meaningful attractor [1100000000] consists of the stable expression of the two floral inhibitor genes *EMF1* and *TFL1*, characteristic of cells that will not become part of flowers. Hitherto, these four long-term patterns of gene expression are consistent with the ABC model.

The analysis performed here revealed a sixth attractor, i.e. [1100000110], representing the stable expression of *EMF1*, *TFL1*, *AP3* and *PI*. The combined expression of the flower inhibition and B function genes has not yet been reported experimentally. This result can thus be interpreted as a prediction of our analysis. From realistic initial conditions, however, it is supposed that this sixth attractor is never reached.

In addition to the six attractors described above, the 10-variable system also comprises six singular steady states, all unstable because they are each characteristic of at least one positive feedback circuit:

[000SS00110] ↔ circuit *API-AG*
[000SS00000] ↔ circuit *API-AG*
[1100000SS0] ↔ circuit *AP3-PI*
[0001000SS0] ↔ circuit *AP3-PI*
[0000100SS0] ↔ circuit *AP3-PI*
[000SS00SS0] ↔ circuits *API-AG* + *AP3-PI*

Each of these unstable steady states is found on a separatrix generated by the corresponding circuit(s). Altogether, these separatrices are dividing the variable space into six basins of attractions, each draining towards one stable state.

Simulation of mutations

Our model also allows the simulation or the prediction of mutant phenotypes. As an example, let us analyze the effect of *LFY* null mutation. Recall that *LFY* specifies the floral fate and activates the four organ identity genes represented in this model, i.e. *API*, *AG*, *AP3* and *PI*. Null *lfy* mutants are practically devoid of flowers, and those that eventually appear present organs with characteristics of wild-type sepals and carpels. To simulate a null *lfy* mutant in our model, it suffices to eliminate the variable *l*. As a result, the *LFY*-dependent *AG* activation pathway is lost, whereas the *LFY*-independent pathway remains unaltered.

Similarly, we can simulate a mutation in *AG*, the only reported gene for the C activity. As *AG* is involved in a functional positive circuit, its inactivation altogether eliminates the corresponding multistable behavior. Indeed, eliminating variable g in the four-variable subnetwork leads to the single stable state [001(0)], draining the entire variable space. Combining this steady state with those of the small subnetwork, we find that only the states corresponding to the A function alone and the combination AB are reached. This result accounts for the phenotype of strong *ag* mutants, whose flowers contain only sepals and petals (Gustafson-Brown *et al.*, 1994).

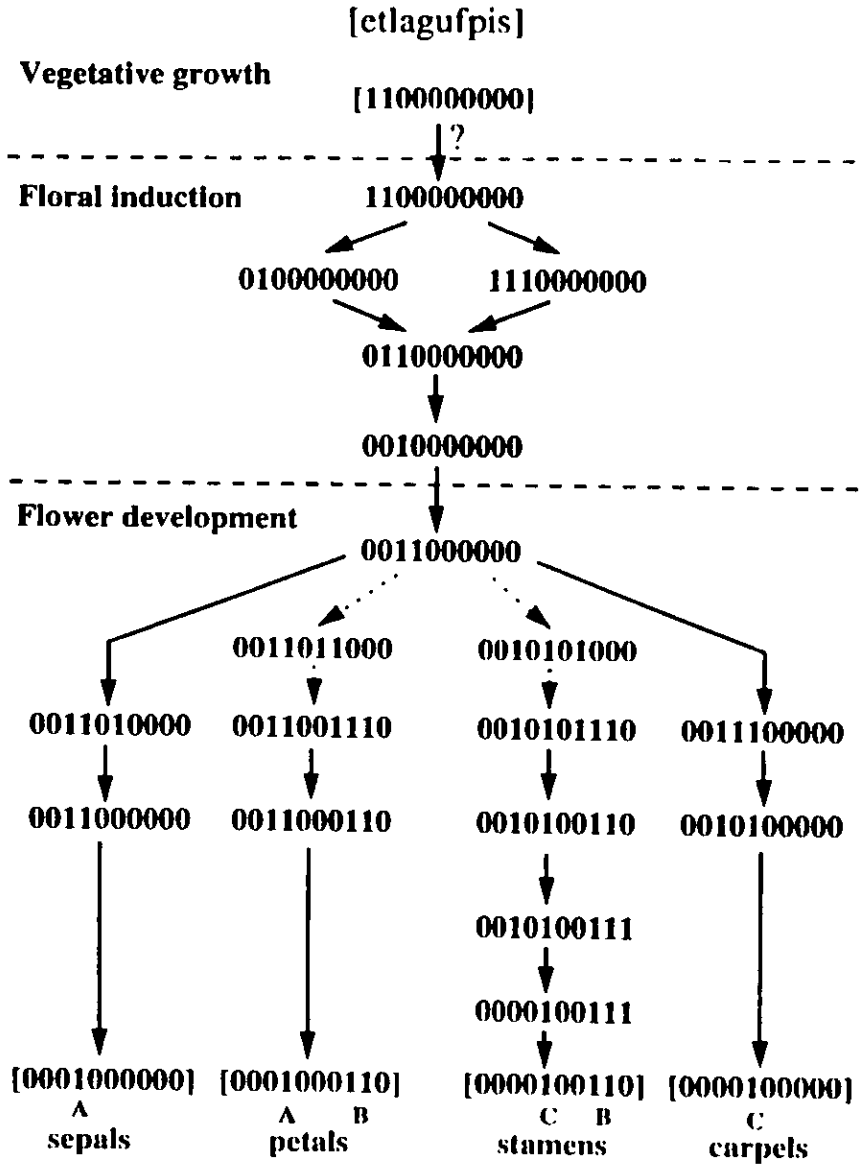


Fig. 4. Possible pathways of gene expression during the transition to flowering and during flower morphogenesis in which sepals, petals, stamens and carpels are formed. The initial state is always the same, but as flower development proceeds, different cells follow alternative pathways. This is only a rough scheme of what might be the dynamics of the whole network. Until now, there are no published reports on the spatio-temporal patterns of expression of significant combinations of these genes in a single plant. We use here a vector notation of the form (*etlagufpis*), where variables *e, t, l, a, g, u, f, p, i* and *s* stand for the levels of activity of the regulatory products of genes *EMF1, TFL1, LFY, AP1, AG, LUG, UFO, AP3, PI* and *SUP*, respectively. Brackets '[]' indicate steady states. An induction signal is needed to switch from the stable non-floral pattern of gene expression to any of the pathways leading to the floral organs. Since we do not know the precise nature of this signal, we indicate it with the use of the question mark. Note that several genes (*LFY, UFO, LUG* and *SUP*) are only transiently expressed. Solid arrows stand for single variable commutations. Dotted arrows stand for multiple variable commutations.

Our model can also be used to predict floral phenotypes of mutations in the non-ABC genes included in it. For example, a simulation of a null mutation in *TFL1* results in the elimination of the non-floral steady state. However, the stable ac-

tivation of the A, B and C functions remains unaltered. The biological interpretation of such a result is that a mutation in the *TFL1* gene gives rise to plants with severely reduced vegetative growth and flowers with wild-type phenotype. Plants with the *tfl1* mutation have very short primary shoots terminated prematurely by a group of flowers with normal or slightly altered organs (Shannon and Meeks-Wagner, 1991). Therefore, the biological effect of inactivating *TFL1* is described accurately in the model by the loss of stability of the non-floral state, converting it into a transient state. In contrast with the regulatory network presented here, the ABC model does not take into account a non-floral state, and hence cannot explain mutations that accelerate or retard the onset of flowering. Within flowers, we can also simulate the dynamical consequences of *cis*-regulatory mutations by adjusting the values of the relevant logical parameters. Moreover, it is possible to assign a phenotype to each simulation. Our model cannot, however, be used to describe mutant phenotypes of *TFL1*, or other genes, outside the flower. A comprehensive analysis of various types of loss and gain-of-function mutations will be presented elsewhere (Mendoza *et al.*, in preparation).

Discussion

We previously presented the topology of the genetic regulatory network shown here and a system of difference equations to model the stable gene activation patterns observed in the flowers of *Arabidopsis* (Mendoza and Alvarez-Buylla, 1998). Here, we focused on the analysis of the dynamical roles of the feedback circuits contained in the same regulatory network. Despite the difference in the formalism used, both models predict the same stable states, but how much of these results depend on the precise genes involved in the network? New developmental genes are rapidly being discovered and characterized in *Arabidopsis*. Will the incorporation of these new genes soon invalidate our results? We think that this will not be the case, because the dynamic properties of feedback circuits depend on their sign (positive or negative) and not on their size (Thomas *et al.*, 1995). Therefore, changes in the interaction matrix accounting for new intermediate elements should not invalidate our results regarding the number and nature of stable gene activation states. For example, if an intermediate gene between *TFL1* and *LFY* were to be discovered, the interaction matrix of Figure 3 would be modified, but the positive circuit involving *TFL1* and *LFY* would still operate and fulfill the same dynamical role. The only difference is that it would become a three-element instead of a two-element circuit. In short, the qualitative analysis presented here does not depend on the precise molecular nature of the regulatory pathways.

We found eight circuits in the regulatory graph and matrix (Figure 2), of which only two are functional in the whole space of gene expression states. These are the positive circuits *AP1-AG* and *AP3-PI*. These circuits, together with the input variable *EMF1*, account for the generation of several separatrices across the phase space, that define six basins of attraction, each containing one stable state. These attractors correspond to the patterns of gene expression found in the four floral organs of *Arabidopsis* (sepals, petals, stamens and carpels), plus a non-floral state and a new state awaiting experimental characterization. Although the regulatory matrix contains several other circuits, none of these are found to play any crucial dynamical role. Once again, we may be concerned by the discovery of new genes and interactions that define additional feedback circuits in the network. We predict that if such a hypothetical loop includes any of the ABC genes, this circuit should not be functional. Indeed, the experimental evidence available up to now suggests that there are only two functional feedback loops among the ABC genes, which correspond to the two functional circuits identified in this analysis. Messenger RNA *in situ* hybridization experiments in mutant backgrounds have shown that the A and C genes inhibit each other (Bowman *et al.*, 1993). This creates the stable exclusive choice between A and C functions. There is also experimental data that support the view that the B function genes only have two possible stable states of activation. They are either both active or both inactive (Krizek and Meyerowitz, 1996). We exclude the possibility of additional functional circuits involving ABC genes because this would change the well-established combinations of activities described by the ABC model. However, we do not exclude the possibility of additional functional feedback circuits involving some of the non-ABC genes included in our model.

Our model accounts only for extreme levels of gene activation, either active (or more properly, with an activity above its threshold of functionality) or inactive (below this threshold). We used this simple binary representation because there is a scarcity of quantitative data. As a result, we reached a consistent but somewhat schematic picture of the differentiation pathways leading to the alternative states of gene expression in the developing flower of *Arabidopsis*. It would be more realistic to think in terms of a gradient of activation between these extreme values. As more data on different patterns of gene expression become available, we will be able to take advantage of the more sophisticated aspects of the generalized logical formalism (e.g. multilevel variables, functions and parameters, analysis of state transitions, etc.) to refine our model. These refined logical models could lead to more precise and detailed accounts of the kinetics of gene expression during flower development. This logical model and its potential further refinements could also help the development of a quantitative differential description. In this respect, note that previous theoretical analyses have shown that analogous continuous and discrete models yield com-

parable results in terms of steady states (Thomas *et al.*, 1995).

The last issue regarding the accuracy of our model concerns the particular values of the parameters (*K*s) that we used. Most of these were based on existing data about the patterns of gene expression observed in *Arabidopsis*. However, not all of the parameters could be inferred from experimental data. Indeed, the values for $K_{a,l}$, $K_{g,l}$, $K_{g,ll}$ and $K_{c,ll}$ were adjusted to obtain plausible transient gene activation patterns that did not affect the configuration of steady states. Nevertheless, an important prediction arises from the corresponding constraints on the values of some of the logical parameters. More precisely, our analysis predicts that *LFY* has at least one additional regulator yet to be discovered. Indeed, in order to obtain the three stable states experimentally found in wild-type *Arabidopsis* (non-flowering, activation only of *AG* or of *AP1*), $K_{l,ll}$ (representing the maximum expression state of *LFY*, i.e. when the products of *TFL1* and *EMF1* are below their respective inhibitory thresholds, and when the product of *AP1* is over its activatory threshold) needs to be set equal to zero, but the experimental evidence available shows that *LFY* is expressed at high levels at the onset of flowering (Blázquez *et al.*, 1997), and at low levels during later stages of flower development. This discrepancy strongly suggests that there is at least one additional regulator of *LFY* still to be discovered. We further predict that this regulator should play a key role in the transition from the non-floral attractors to the flowering pathways, implying both temporal and functional specific features.

The analysis presented here includes all the possible activation states of the network. However, cells in an organism likely only access a limited subset of these theoretical states during development. Each cell probably follows a limited set of states until it attains its final differentiated fate. In the particular case of *Arabidopsis*, some plausible pathways of gene activation states are shown in Figure 4. At the onset of flower development, flower inhibitors (*EMF1* and *TFL1*) are active and cells are trapped in the stable non-flowering state [1100000000]. An external signal, which alters the activation of the flower inhibitors *EMF1* and *TFL1*, is thus needed to move the regulatory network towards one of the floral steady states. From this point onwards, a choice between four main pathways that reflect alternative gene activation routes in different parts of the flower arises. *AP1* mRNA is found throughout floral primordia during the very first stages of development, but by the time the sepals arise, its expression is restricted to the first two whorls. *LFY* follows a similar expression pattern, except that its mRNA becomes undetectable in the outer whorls at late stages of flower development. This temporal expression leads to the long-term activation pattern of the first and second whorls. On the other hand, *LFY* and *AP1* expression are found in the inner whorls of the flower until sepals arise, when *AG* expression sets in (Weigel and Meyerowitz, 1993).

The dynamics explained above account for the differentiation of those regions of the flower characterized by either A or C activity. To complete this picture, it is necessary to take into account B activity, as well. In fact, the transient expression of *LFY* mentioned above might be important for the activation of *AP3* and *PI*, ultimately leading to the stable states [0001000110] (AB combination) and [0000100110] (BC combination). This role of *LFY* in the establishment of B activity is supported by the observation that *LFY* mutant plants bear flowers that often lack petals and stamens, and that the expression of *AP3* and *PI* messengers is reduced. Recent data on transgenic plants bearing a super activator form of the *LFY* product (LFY protein fused to the activation domain VP16) also suggest that LFY is an important activator of B genes, particularly *AP3*. In addition, *UFO* might act as a co-regulator in the activation of *AP3* and *PI* (Lee *et al.*, 1997; Parcy *et al.*, 1998), but the mechanism is not yet known. Once *AP3* and *PI* are fully active, the activation of *LFY* becomes dispensable, because *AP3* and *PI* cross-activation ensures their mutual maintenance. Additional data on the kinetics of expression of cadastral genes and on their upstream regulators are needed to refine this scheme.

In addition to the vegetative state and the four states corresponding to floral organs already accounted for by the ABC model, our analysis predicts a sixth attractor [1100000110], which implies a stable activation of *EMF1*, *TFL1*, *AP3* and *PI*. This state automatically results from the combination of the steady states [11] and [1100] generated by the two regulatory modules. Such relative independence of B and A/C functions is supported by the fact that B activity can be disrupted without affecting other aspects of flower morphogenesis (Day *et al.*, 1995).

Conclusions and prospects

We have derived a network of 10 genes involved in the genetic control of flower morphogenesis in *Arabidopsis* from published molecular and genetic data. Using a logical formalism, we provided a qualitative dynamical analysis of this regulatory network and derived the parameter constraints accounting for the different patterns of gene expression found in the four floral organs of *Arabidopsis* (sepals, petals, stamens and carpels), as well as for a non-floral state. In addition, our model leads to the prediction of a new stable state characterized by a simultaneous expression of the floral repressor and B function genes. Finally, we predict the discovery of at least a new regulator of the gene *LFY*, likely to be involved in the transition from the non-flowering state to the flowering pathways.

Our analysis also explains how specific genes cooperate to generate the alternative patterns of gene expression through two functional positive circuits. From a general perspective, this work further supports the idea that positive regulatory cir-

cuits play a central role in development. Such a role was postulated by several theoretical biologists in the late 1970s (see, for example Lewis *et al.*, 1977; Meinhardt, 1978; Thomas, 1978). More recent studies have confirmed that key developmental genes are often involved in such positive circuits, thus positively affecting their own expression. Often, but not always, this is accomplished through a direct autoregulation, like in the case of the gene MyoD that codes for a muscle-cell specific transcription factor and is involved in both cell determination and differentiation [see Murre *et al.*, 1989; more examples are found in Thieffry *et al.*, 1995].

The present model also emphasizes another important feature of developmental gene networks, namely their modularity. Indeed, developmental biologists are now realizing that developmental genes can be grouped in relatively independent 'regulatory modules' or 'syntagms', acting in parallel or in temporal cascades (the best available example is the segmentation network of *D.melanogaster*; see, for example, Reinitz and Sharp, 1995; Sánchez *et al.*, 1997). The present analysis points to a straightforward formal criterion to isolate regulatory modules in complex intertwined network, namely that a module can be simply defined as a set of feedback circuits sharing some element(s).

Our analysis accounts for various stable states and, to a lesser extent, for the kinetics of gene expression of the ABC and other interacting genes included in the network. It is clear, however, that pattern formation and cellular differentiation most often involve interplay between subcellular and supracellular levels. In addition, this interplay should be considered in the context of cellular growth and multiplication. Our model should thus be considered as a first step toward a more ambitious model of flower development, covering intercellular communication, cell growth and division, as well as three-dimensional morphogenesis. For example, our model could not account for flower phenotypes with the same four floral organs, but with a different spatial arrangement (e.g. sepals, petals, carpels and stamens), because it does not incorporate explicit spatial information. Future models will have to consider genes underlying cellular processes that are at the basis of aspects of floral morphology, such the shape, number or color of floral organs.

In spite of the limitations of our model, it can already be used to explore the role of regulatory genes in determining important aspects of floral morphological variation across angiosperms (plants with flowers). The stereotypical structure of flowers of >250 000 angiosperm species, and molecular and genetic data available for a few other model angiosperm systems, strongly suggest that the ABC model is widely conserved across flowering plants (Bowman, 1997). All flowering species, except one (*Lacandonia schismatica*; Martínez and Ramos, 1989; Alvarez-Buylla *et al.*, 1999), have concentric whorls of floral organs which, from the outside to the inside of the flower, are: sepals (or modified sterile organs), petals (or modified sterile organs), stamens and carpels. Furthermore, comparative molecular studies have shown that the ABC genes and their expression patterns are highly conserved across distantly related angiosperm species (Bradley *et al.*, 1996; Mena *et al.*, 1996; Purugganan, 1997). Therefore, the ABC model may be used to describe important aspects of the floral architecture of most angiosperms. In this context, the NET model is useful to elaborate hypotheses regarding variations in the network that underlie important aspects of flower morphological diversification during evolution. We refer to the morphological response that results from the variation of the patterns of gene expression involved in the control of the morphogenesis of floral organs (e.g. sepals, petals, stamens and carpels). This model also constitutes a formal framework susceptible to guiding future experiments aimed at identifying new regulatory genes or interactions. For example, our previous analysis of the NET model (Mendoza and Alvarez-Buylla, 1998) led us to propose that the gene *AG* was activated by *LFY*, a regulatory interaction which was soon confirmed experimentally by an independent group (Parcy *et al.*, 1998).

Acknowledgements

This work was funded by an UNAM scholarship to L.M., and CONACYT and PSPA grants to E.A.B., as well as by financial support from the Institute of Ecology, UNAM and a PEW fellowship to E.A.B. Thanks to R.Thomas, L.Sánchez and two anonymous referees for valuable comments on previous versions of this paper.

References

Alvarez-Buylla,E.R., Chaos,A., Burgeff,C., Gamboa,A., Izquierdo,L., Mendoza,L., Tapia,R. and Vergara,F. (1999) Genes homeóticos: genética molecular, redes de regulación, mecanismos de desarrollo y evolución. In Merchant,H. and Jimenez-García,L.F. (eds), *Biología Celular y Molecular en México*. In press.

Blázquez,M.A., Soowal,L.N., Lee,I. and Weigel,D. (1997) LEAFY expression and flower initiation in *Arabidopsis*. *Development*, **124**, 3835–3844.

Bowman,J.L (1997) Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. *J. Biosci.*, **22**, 515–527.

Bowman,J.L., Alvarez,J., Weigel,D., Meyerowitz,E.M. and Smyth,D.R. (1993) Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. *Development*, **119**, 721–743.

Bradley,D., Carpenter,R., Copsey,L., Vincent,C., Rothstein,J. and Coen,E. (1996) Control of inflorescence architecture in *Antirrhinum*. *Nature*, **379**, 791–797.

Coen,E.S. and Meyerowitz,E.M. (1991) The war of the whorls: genetic interactions controlling flower development. *Nature*, **353**, 31–37.

Day,C.D., Galgoci,B.F.C. and Irish,V.F. (1995) Genetic ablation of petal and stamen primordia to elucidate cell interactions during floral development. *Development*, **121**, 2887–2895.

Glass,L. and Kauffman,S.A. (1973) The logical analysis of continuous, non-linear biochemical control networks. *J. Theor. Biol.*, **39**, 103–129.

Goto,K. and Meyerowitz,E.M. (1994) Function and regulation of the *Arabidopsis* floral homeotic gene PISTILLATA. *Genes Dev.*, **8**, 1548–1560.

Gouzé,J.L. (1998) Positive and negative circuits in dynamical systems. *J. Biol. Syst.*, **6**, 11–15.

Gustafson-Brown,C., Savidge,B. and Yanofsky,M.F. (1994) Regulation of the Arabidopsis floral homeotic gene *APETALA1*. *Cell*, **76**, 131–143.

Jack,T., Brockman,L.L. and Meyerowitz,E.M. (1992) The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell*, **68**, 683–697.

Kauffman,S.A. (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.*, **22**, 437–467.

Kempin,S.A., Savidge,B. and Yanofsky,M.F. (1995) Molecular basis of the *cauliflower* phenotype in *Arabidopsis*. *Science*, **267**, 522–525.

Krizek,B.A. and Meyerowitz,E.M. (1996) The *Arabidopsis* homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. *Development*, **122**, 11–22.

Lee,I., Wolfe,D.S., Nilsson,O. and Weigel,D. (1997) A *LEAFY* co-regulator encoded by *UNUSUAL FLORAL ORGANS*. *Curr. Biol.*, **7**, 95–104.

Levin,J.Z. and Meyerowitz,E.M. (1995) *UFO*: an *Arabidopsis* gene involved in both floral meristem and floral organ development. *Plant Cell*, **7**, 529–548.

Lewis,J., Slack,J. and Wolpert,L. (1977) Thresholds in development. *J. Theor. Biol.*, **65**, 579–590.

Liu,Z. and Meyerowitz,E.M. (1995) *LEUNIG* regulates *AGAMOUS* expression in *Arabidopsis* flowers. *Development*, **121**, 975–991.

Mandel,M.A., Gustafson-Brown,C., Savidge,B. and Yanofsky,M.F. (1992) Molecular characterization of the *Arabidopsis* floral homeotic gene APETALA1. *Nature*, **360**, 273–277.

Martinez,E. and Ramos,C.H. (1989) Lacandoniaceae (Triuridales): Una nueva familia de Mexico. *Ann. M. Bot. Gard.*, **76**, 128–135.

Meinhardt,H. (1978) Space-dependent cell determination under the control of a morphogen gradient. *J. Theor. Biol.*, **74**, 307–321.

Mena,M., Ambrose,B.A., Meeley,R.B., Briggs,S.P., Yanofsky,M.F. and Schmidt,R.J. (1996) Diversification of C-function activity in maize flower development. *Science*, **274**, 1537–1540.

Mendoza,L. and Alvarez-Buylla,E.R. (1998) Dynamics of the genetic regulatory network for *Arabidopsis thaliana* flower morphogenesis. *J. Theor. Biol.*, **193**, 307–319.

Monod,J. and Jacob,F. (1961) General conclusions: telenomic mechanisms in cellular metabolism, growth and differentiation. *Cold Spring Harbor Symp. Quant. Biol.*, **26**, 389–401.

Murre,C., McCaw,P.S. and Baltimore,D. (1989) A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD and myc protein. *Cell*, **56**, 777–783.

Nilsson,O., Lee,I., Blázquez,M.A. and Weigel,D. (1998) Flowering-time genes modulate the response to *LFY* activity. *Genetics*, **150**, 403–410.

Parcy,F., Nilsson,O., Busch,M.A., Lee,I. and Weigel,D. (1998) A genetic framework for floral patterning. *Nature*, **395**, 561–566.

Plahte,E., Mestl,T. and Omholt,S.W. (1995) Feedback loops, stability and multistationarity in dynamical systems. *J. Biol. Syst.*, **3**, 409–413.

Purugganan,M.D (1997) The MADS-box floral homeotic gene lineages predate the origin of seeds plants: Phylogenetic and molecular clock estimates. *J. Mol. Evol.*, **45**, 392–396.

Ratcliffe,O.J., Amaya,I., Vincent,C.A., Rothstein,S., Carpenter,R., Coen,E.C. and Bradley,D.J. (1998) A common mechanism controls the life cycle and architecture of plants. *Development*, **125**, 1609–1615.

Reinitz,J. and Sharp,D.H. (1995) Mechanism of formation of eve stripes. *Mech. Dev.*, **49**, 133–158.

Rosen,R. (1968) Recent developments in the theory of control and regulation of cellular processes. *Int. Rev. Cytol.*, **23**, 25–88.

Sakai,H., Medrano,L.J. and Meyerowitz,E.M. (1995) Role of *SUPERMAN* in maintaining *Arabidopsis* floral whorl boundaries. *Nature*, **378**, 199–203.

Sánchez,L., Van Helden,J. and Thieffry,D. (1997) Establishment of the dorso-ventral pattern during embryonic development of *Drosophila melanogaster*: a logical analysis. *J. Theor. Biol.*, **189**, 377–389.

Shannon,S. and Meeks-Wagner,D.R. (1991) A mutation in the *Arabidopsis* TFL1 gene affects inflorescence meristem development. *Plant Cell*, **3**, 877–892.

Snoussi,E.H. (1989) Qualitative dynamics of piece-linear differential equations: a discrete mapping approach. *Dyn. Stability Syst.*, **4**, 189–207.

Snoussi,E.H. (1998) Necessary conditions for multistationarity and stable periodicity. *J. Biol. Syst.*, **6**, 3–9.

Snoussi,E.H. and Thomas,R. (1993) Logical identification of all steady states: the concept of feedback loop characteristic states. *Bull. Math. Biol.*, **55**, 973–991.

Thieffry,D., Colet,M. and Thomas,R. (1993) Formalisation of regulatory networks: a logical method and its automatization. *Math. Model. Sci. Comput.*, **2**, 144–151.

Thieffry,D., Snoussi,E.H., Richelle,J. and Thomas,R. (1995) Positive loops and differentiation. *J. Biol. Syst.*, **3**, 457–466.

Thomas,R. (1978) Logical analysis of systems comprising feedback loops. *J. Theor. Biol.*, **73**, 631–656.

Thomas,R. (1991) Regulatory networks seen as asynchronous automata: a logical description. *J. Theor. Biol.*, **153**, 1–23.

Thomas,R., Thieffry,D. and 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.

Weigel,D. (1995) The genetics of flower development: from floral induction to ovule morphogenesis. *Annu. Rev. Genet.*, **29**, 19–39.

Weigel,D. and Meyerowitz,E.M. (1993) Activation of floral homeotic genes in *Arabidopsis*. *Science*, **261**, 1723–1726.

Weigel,D. and Nilsson,O. (1995) A developmental switch sufficient for flower initiation in diverse plants. *Nature*, **377**, 495–500.

Weigel,D., Alvarez,J., Smyth,D., Yanofsky,M.F. and Meyerowitz,E.M. (1992) LEAFY controls floral meristem identity in *Arabidopsis*. *Cell*, **69**, 843–859.

Yang,C.H., Chen,L.J. and Sung,Z.R. (1995) Genetic regulation of shoot development in *Arabidopsis*: role of the EMF genes. *Dev. Biol.*, **169**, 421–435.

Yanofsky,M.F., Ma,H., Bowman,J.L., Drews,G.N., Feldmann,K.A. and Meyerowitz,E.M. (1990) The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature*, **346**, 35–40.

APENDICE 5

GENETIC REGULATION OF ROOT HAIR DEVELOPMENT IN *Arabidopsis thaliana*: A NETWORK MODEL

Luis Mendoza and Elena R. Alvarez-Buylla

Laboratorio de Genética Molecular y Evolución
Instituto de Ecología, UNAM. Ap. Postal 70-275
Coyoacán, D.F. CP04510. MEXICO

ABSTRACT

The root epidermis of *Arabidopsis thaliana* is formed by alternate files of hair and non-hair cells. Epidermal cells overlying two cortex cells eventually develop a hair, while those overlying only one cortex cell do not. Here we propose a network model that integrates most of the available genetic and molecular data on the regulatory and signaling pathways underlying root epidermal differentiation. The network architecture includes two distinctive pathways; one formed by the transcription factors *TTG*, *R* homolog, *GL2* and *CPC*, and the other one by the signal transduction proteins ETR1 and CTR1. Both parallel pathways regulate the activity of AXR2 and RHD6, which in turn control the development of root hairs. The regulatory network was simulated as a dynamical system of eight discrete state variables. The distinction between epidermal cells contacting one or two cortical cells was accounted for by fixing the initial states of CPC and ETR1. The model allows for predictions of mutants and pharmacological effects because it includes the ethylene receptor. The dynamical system reaches one of six stable states depending upon the initial state of the gene *CPC* and the ethylene receptor. Two of the stable states describe the activation patterns observed in mature trichoblasts (hair cells) and atrichoblasts (non-hair cells), in the wild type phenotype and under normal ethylene availability. The other four states correspond to changes in the number of hair cells due to experimentally induced changes in ethylene availability. This model is a new interpretation of the interactions among genes that encode transcription factors that regulate root hair development and the proteins involved in the ethylene transduction pathway. This is the first effort to use a dynamical system to understand the complex genetic regulatory interactions that rule *Arabidopsis* primary root development. The advantages of this type of models over static schematic representations are discussed.

INTRODUCTION

The increasing amount of data on the molecular mechanisms underlying cellular differentiation and morphogenesis are putting forward complex networks of genetic interactions. These are dynamical systems that operate in specific cellular and spatial scales. Network models of genetic regulatory interactions are dynamical representations of molecular interactions, and can be used to predict the stable activation states attained by a given genetic architecture within mature cells (Kauffman, 1993). These models provide a formal dynamical framework to analyze cellular differentiation, in contrast to schematic representations of gene and protein interactions often published in experimental papers. However, network models for specific biological systems are still scarce (for notable examples see Bodnar, 1997; Reinitz & Sharp, 1995). In this paper we propose one network model for cell differentiation during root epidermis development of the experimental plant system *Arabidopsis thaliana*.

In contrast to most animals, plant cells cannot move, while positional cues are important in ruling cellular fate. Also, plant organ development is remarkably plastic in response to environmental conditions (Meyerowitz, 1997; Pigliucci, 1996; Scheres & Wolkenfelt, 1998). Therefore, cell signaling should play important roles in plant cellular specification and pattern formation, which take place in shoot and root meristems, where organs continuously develop (Scheres, 1997). Within floral meristems, transcription factors are known to be involved in regional specification (Coen & Meyerowitz, 1991; Weigel & Meyerowitz, 1993), but the gene products required for the signaling processes during pattern formation are only starting to be studied. The complexity of the shoot meristem, however, hinders the experimental approaches that are required to understand the role of cell signaling during development. In contrast, the regular cellular architecture, small size and transparency of the *Arabidopsis* root meristem, enables the combination of genetic, cell ablation and microinjection approaches to study the signaling mechanisms involved in cell specification (see Scheres, 1996; 1997). This makes the root a particularly attractive system to develop mechanistic models, which help understand how molecular signals and cellular interactions are integrated across cells during development.

The root epidermis of *Arabidopsis* is an exceptionally simple system to address questions regarding the cellular and genetic mechanisms that determine a stereotyped pattern of cellular differentiation. The root epidermis is made of alternate columns of hair (trichoblasts) and non-hair cells (atrichoblasts); and inside it, cellular fate is determined by the relative location of epidermal cells with respect to the cortex cell walls. However, it seems that the atricoblast cell fate is a default state (Berger *et al.*, 1998a). Recent studies in the root of *Arabidopsis* have identified at least four transcription factors that regulate root hair development, and the ethylene signaling pathway as one of the signaling mechanisms involved in it (see Scheres & Wolkenfelt, 1998). It is unclear if ethylene is a molecule available to all epidermal cells, or if it forms a gradient conveying positional information during sub-specification of epidermal cell-types. Additional experimental results are necessary to solve the details of the interplay between transcriptional regulation and ethylene transduction pathway; but to integrate all the data available to date we put forward the first genetic regulatory network model for root hair development.

No previous model has proposed a genetic regulatory mechanism responsible for the cellular patterning observed in the *Arabidopsis* root epidermis. In this paper we propose a regulatory network model of discrete states that may be used as a framework to predict the effect of mutations that alter root hair formation. The network is implemented as a dynamical system with discrete state variables. Departing from any initial activation state, the model reaches one of six stable activation patterns. Two of these correspond to stable activation states that characterize mature trichoblasts and atrichoblasts under wild type conditions. The other four correspond to genetic patterns that would diminish or augment the trichoblasts to atrichoblasts ratio according to induced changes in ethylene availability. This model contrasts in two important ways with our previous models of flower development (Mendoza & Alvarez-Buylla, 1998; Mendoza *et al.*, 1999). First, in our previous network implementations we used the ABC model of flower morphogenesis to translate the steady gene activation patterns into flower morphology (Coen & Meyerowitz, 1991; Meyerowitz, 1994). Since we lack a similar morphogenetic model for root hair formation, we used the activation state of the most downstream elements of the regulatory network (RHD6 and AXR2) to establish cell fate (*i.e.*, trichoblasts or atrichoblasts). Second, the flower models

were treated as autonomous dynamical systems in which no links with extracellular signals were postulated. In this case, however, we postulate a link with external signals via the response of the elements representing *CPC* and ETR1. Such responses are incorporated by fixing the initial states of these two elements in the network, in such a way that the appearance of root hairs is determined by the position of epidermal cells relative to that of cortical cells. This is the first dynamical model that is able to describe the effect of mutations as well as pharmacological treatments on the cellular patterning in the root epidermis.

meristem. Along this gradient, from the root tip to the plant base four regions have been defined: the division, the slow elongation, the fast elongation, and the differentiation zones (Berger *et al.*, 1998b). Each cell file is maintained by elongation and subsequent anticlinal division (radial, with the division plane perpendicular to the growth axis) of initial cells. There are four types of initial cells: epidermis/lateral root cap, cortex/endodermis, vascular and columella initials. While initial cells only divide, the resulting daughter cells are able to divide and differentiate into one of the mature cell types forming the different

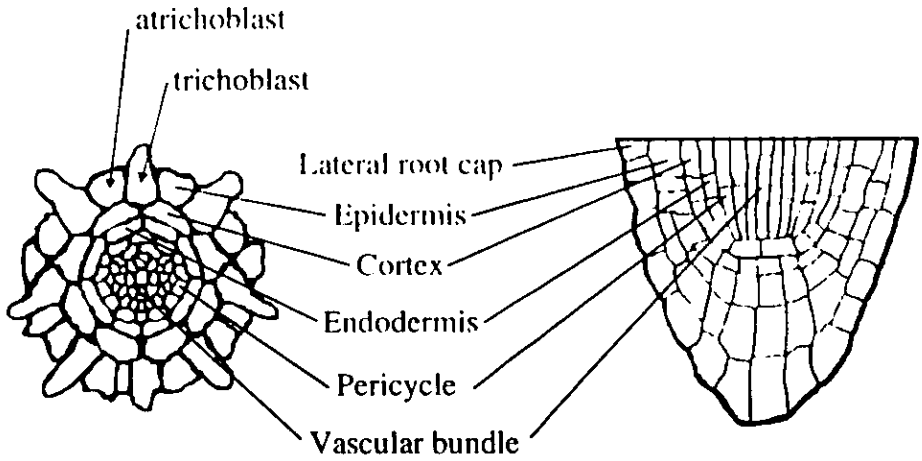


Figure 1

THE ROOT OF *Arabidopsis*

The mature root of *Arabidopsis* consists of single concentric layers of cells, which from the periphery to the center form the lateral root cap, epidermis, cortex, endodermis and pericycle, surrounding a vascular bundle (Fig. 1). Each tissue layer is formed by cells arranged in parallel columns called files. Mature cells are located distal to the root tip, while immature cells are near the root

root tissues (Dolan *et al.*, 1993; 1994). Epidermal cells undergo a further process of differentiation. Cells that contact the anticlinal cell wall between two cortical cells differentiate into root hair cells (trichoblasts). In contrast, developing epidermal cells contacting the wall of a single cortical cell differentiate into mature hairless cells (atrichoblasts). The hair is a tubular projection that develops in trichoblasts, but its presence is not the only morphological difference with atrichoblasts. From a very early stage of development trichoblasts have a more

densely staining cytoplasm, a reduced vacuolation and a smaller size than atrichoblasts (Berger *et al.*, 1998b; Dolan *et al.*, 1994).

MOLECULAR AND GENETIC BASIS OF THE REGULATORY NETWORK MODEL

We present a network model that integrates most of the published genes involved in root epidermis development for which regulatory interactions could be inferred. We first review the experimental evidence that supports the proposed regulatory network architecture. The model includes genes that encode transcription factors and regulatory proteins. In some cases, transcriptional and post-transcriptional regulation could not be distinguished based on available data. Therefore, if an element of the network model is "off", it represents either a null transcriptional gene activity or an inactive protein. Despite being biologically relevant, the ambiguity on the molecular mechanism does not affect the dynamical result, because the particular nature of functionality (or lack of it) is irrelevant for the study of the collective behavior of interdependent elements. For a review on the known molecular nature of many mutations see Van der Berg *et al.* (1998). Throughout the text, genes coding for known transcription factors are in italics. Wild type and mutated genes are indicated with uppercase and lowercase letters, respectively.

TRANSPARENT TESTA GLABRA (*TTG*) is a gene coding for a WD40 protein involved in root hair development, and its loss-of-function mutation results in plants with hairs throughout the epidermis (Galway *et al.*, 1994; Walker *et al.*, cited in Hung *et al.*, 1998). *TTG* is an important activator of *GLABRA2* (*GL2*), of which its mRNA is expressed inside atrichoblasts in wild type plants. Like *ttg* plants, *gl2* mutants have hairs throughout the epidermis. However, *gl2* epidermal cells can be distinguished as trichoblasts or atrichoblast based on vacuolation and cytoplasmic density, while *ttg* epidermal cells cannot (Galway *et al.*, 1994; Masucci *et al.*, 1996; Masucci & Schiefelbein, 1996; Schiefelbein *et al.*, 1997). *GL2* encodes a homeodomain protein that probably functions as a transcription factor of genes involved in determining the hairless cellular fate. One important activator of *GL2* is the gene *TTG*, because *GL2* expression is markedly reduced in *ttg* mutants (Di Cristina, 1996; Hung *et al.*, 1998). However, the activation of *GL2* by *TTG* seems to be mediated by another gene product. In *Arabidopsis*, over-expression of the corn *R* cDNA (a *myc* gene), under control of the cauliflower mosaic virus 35S promoter, suppresses the *ttg* phenotype (Galway *et al.*, 1994). Such information suggests that an *Arabidopsis R* homolog is acting downstream of *TTG*. Moreover, transgenic *35S::R GL2::GUS* plants show *GUS* activity throughout the epidermis, lateral root cap cells and cortical cells; while *35S::R gl2* plants show a *gl2* phenotype (Hung *et al.*, 1998). All this evidence supports that *TTG* activates an *Arabidopsis R* homolog, which in turn activates *GL2* (Fig. 2a).

TTG is not the only regulator of *GL2*: *GL2::GUS ttg* seedlings exhibit a significant reduction, but not a total loss, of *GUS* activity. Interestingly, the remaining *GL2::GUS* activity is found in the atrichoblasts, where *GL2* it is normally expressed (Hung *et al.*, 1998). Therefore, the *ttg* mutation affects the level of *GL2* expression but not its spatial localization. These results suggest that an additional *TTG*-independent pathway restricts the *GL2* expression to atrichoblasts. Such alternative pathway might be mediated by *CAPRICE* (*CPC*) that is a putative transcription factor expressed in trichoblasts, as revealed by GUS staining (Wada *et al.*, 1997) and preliminary *in situ* hybridization experiments (Wada *et al.*, 1998). The *cpc* mutants express *GL2* in almost all epidermal cells and have few irregularly distributed root hairs, while *35S::CPC* plants have root hairs in all epidermal cells like the *gl2* or *ttg* mutant phenotypes. Also, *cpc gl2* double mutants have the same phenotype as *gl2* single

mutants (Wada *et al.*, 1997). Together, all this evidence indicates that *CPC* is a negative regulator of *GL2* (Fig. 2a). *CPC* and *TTG* regulate *GL2* via two different pathways, because *cpc ttg* roots have an intermediate phenotype compared to each of the single mutants (Wada *et al.*, 1997).

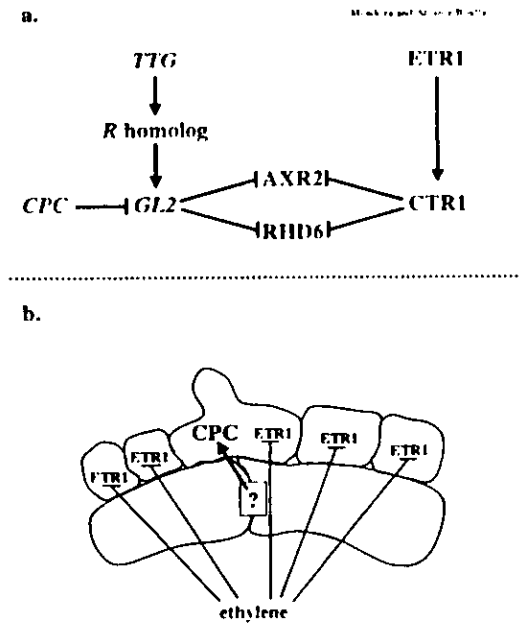


Figure 2

Members of the ethylene receptor family share sequence similarity, and mutations in any of them have a similar ethylene insensitive phenotype (Hua *et al.*, 1995; Hua & Meyerowitz, 1998). Analyses of single and multiple mutants showed that *ETHYLENE RESISTANT 1* (*ETR1*), *ETHYLENE RESISTANT 2* (*ETR2*), *ETHYLENE INSENSITIVE 4* (*EIN4*) and *ETHYLENE RESPONSE SENSOR 2* (*ERS2*) have redundant functions in the ethylene signaling pathway (Hua & Meyerowitz, 1998). The most studied member of such family is ETR1, which is a membrane protein very similar to the bacterial two-component histidine kinases (for recent reviews see Fluhr, 1998; Kieber, 1997). We have included only the ETR1 protein in our model because the ethylene-binding activity has been demonstrated solely for this receptor. However, it should be kept in mind that the *in vivo* response to ethylene is probably mediated by the formation of multiple complexes (reviewed in Johnson & Ecker, 1998). It is known that ETR1 acts upstream of *CONSTITUTIVE TRIPLE RESPONSE 1* (*CTR1*, Kieber, 1997), and ETR1 seems to be an activator of CTR1 (Fig. 2a) because loss-of-function mutants of the ethylene receptor family and *ctr1* mutants are similar (Hua & Meyerowitz, 1998). Finally, in contrast with the typical mechanism of hormone action, ethylene *represses* the activity of the receptor complex upon binding, which in turn activates CTR1 (Hua & Meyerowitz, 1998). Although the net effect of ethylene is the inhibition of CTR1, its specific mechanism of action is relevant for the specification of the model (see ahead).

The ethylene response pathway plays a major role in root hair appearance and development (Tanimoto *et al.*, 1995). Such pathway has been partially uncovered, and includes the gene *CTR1*. Plants with the *ctr1* mutation have ectopic root hairs; but in contrast to *ttg* and *gl2* plants, they have only a slight increase in the proportion of hair to non-hair cells (Dolan *et al.*, 1994). The *CTR1* gene encodes a putative serine-threonine kinase, and the phenotype of *ctr1* mutants suggests that the CTR1 protein negatively regulates the ethylene signaling pathway (Kieber *et al.*, 1993; Scheres, 1997). Also, it is known that CTR1 acts downstream of the ethylene receptor family. The family of genes encoding different forms of the ethylene receptor includes *ETR1*, *ETR2*, *EIN4*, *ERS1* and *ERS2*. *In vitro* assays show that CTR1 interacts with, at least, ETR1 and ERS1 proteins (Clark *et al.*, 1998).

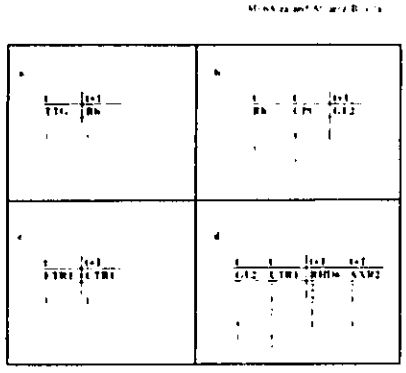
AUXIN RESISTANT2 (*AXR2*) and *ROOT HAIR DEFECTIVE6* (*RHD6*) are also involved in root hair development but they seem to be part of two different pathways downstream of *GL2*. Both single *axr2* and *rhd6* mutants have fewer root hairs than wild type plants, while double

axr2 rhd6 mutants do not have any root hair, suggesting that AXR2 and RHD6 form part of two separate pathways. However, the pathway including RDH6 seems to be more important than that of AXR2, because *rhd6* mutants have fewer hairs than *axr2* mutants (see ahead). Double *axr2 ttg* and *axr2 gl2* mutants have the same phenotype as *axr2* plants (Masucci & Schiefelbein, 1996), suggesting that AXR2 is downstream of the *TTG/GL2* pathway. We propose that *GL2* inhibits the AXR2 gene, or alternatively, inactivates the AXR2 product (Fig. 2a). This ambiguity is not problematic for defining the model because the mathematical treatment is identical in both cases. In contrast, *ttg rhd6* and *gl2 rhd6* mutants have more hairs than *rhd6*, but clearly fewer than *ttg* or *gl2* plants (Masucci & Schiefelbein, 1996). This partial suppression of the *ttg/gl2* phenotype indicates that RHD6 is downstream of *TTG* and *GL2*. However, *gl2 rhd6* mutants have more hairs than *rhd6*, also supporting that there is an RHD6-independent pathway altering root hair appearance. To explain the phenotypes described above we propose that RHD6 is inhibited by *GL2* but in a separate pathway from that of AXR2 (Fig. 2a). However, the inhibition of RHD6 by *GL2* seems to be weaker than that of AXR2, because the difference in the number of root hairs between *axr2* and *axr2 gl2* mutants is larger than the difference between *rhd6* and *rhd6 gl2* mutants. Finally, the fact that *GL2::GUS* expression is not altered by mutations in AXR2, RHD6, or even *GL2* (Masucci & Schiefelbein, 1996) suggests that there is no feedback from RHD6 or AXR2 to *GL2*.

RHD6 and AXR2 seem to connect the transcriptional regulation pathway determining cell fate in the root epidermis (*i.e.* trichoblasts or atrichoblasts) to the ethylene response pathway via *GL2*. ACC, the direct precursor of ethylene, induces the appearance of root hairs in a dose-dependent manner in wild type plants (Tanimoto *et al.*, 1995). The hairless *rhd6* and *axr2* single mutants almost recover the wild type phenotype when treated with ACC. However, ACC addition does not induce root hair formation in *rhd6 axr2* double mutants (Masucci & Schiefelbein, 1996). This data suggests that RHD6 and AXR2 are downstream of the ethylene response pathway. Such a pathway from the ethylene receptor to RHD6 and AXR2 should be independent from *GL2*, because ACC treatment does not alter *GL2* expression (Masucci & Schiefelbein, 1996). We propose that CTR1 represses AXR2 and RHD6 independently of *GL2* (Fig. 2a). We do not know, however, if such negative regulation is exerted at a transcriptional level or directly on the protein.

Abnormal cell divisions and laser cell ablation experiments show that the expression of *GL2* is a default state in the root epidermis; thus the atrichoblast cellular fate does not seem to involve any extracellular signal. The files of epidermal cells are created by transverse anticlinal divisions and further enlargement of an epidermis initial, keeping a constant number of epidermal files. However, sometimes an epidermal cell (often a trichoblast) divides longitudinally originating an extra file. Trichoblasts do not express *GL2* before an abnormal division; however, in the resulting daughter cell that overlies only one cortical cell *GL2* becomes active (Fig. 5 in Berger *et al.*, 1998b). Since the expression patterns of *GL2* and *CPC* are complementary (Wada *et al.*, 1997), the latter should not be active in the newly formed atrichoblast. In this new atrichoblast, either a repressor of *CPC* could become active, or an activator could be absent. Some indirect evidence supports the second possibility. Developing atrichoblasts, isolated from trichoblasts, other atrichoblasts or cortical cells, continue with their normal division and development into mature non-hair epidermal cells, as evidenced by molecular and morphological markers (Berger *et al.*, 1998a). In contrast, isolated trichoblasts do not always continue with their normal development. Consequently,

the non-hair cell fate (*i.e.*, active *GL2* and inactive *CPC*) does not require a positional information and thus might be the default genetic and developmental state. Under such scenario, the development of a hair cell involves an external signal that initiates the expression of *CPC* with the consequent inhibition of *GL2*.



The network model incorporates the effect of extra-cellular signals that modify the activation state of two network elements: *CPC* and ETR1. The initial states of these determine which attractor is reached. Each attractor is a particular pattern of stable gene activation, including AXR2 and RHD6. Moreover, the activation state of AXR2 and RHD6 can be translated into a particular phenotype. We determined the initial state of *CPC* based on the location of cells. Epidermal cells that contact two cortical cells express *CPC* (initial state = 1) while those that contact only one cell do not express this gene (initial state = 0). Also, we assumed that the ethylene supply, and therefore the ethylene receptor activity, are maintained constant and uniform throughout the epidermis (Fig. 2b, and see discussion ahead). In our model we set ETR1=1 to simulate the wild type activation level of ethylene receptor in both trichoblasts and atrichoblasts. On the contrary, to simulate pharmacological treatments we have set ETR1 initial states to 0 and 2, to model the addition of ethylene (or ACC) and AVG, respectively. This is because ethylene inhibits the activity of its own receptor, while AVG inhibits ethylene synthesis and thus leads to an over-activation of the ethylene receptor. Based on experimental evidence at hand, we assumed that the ethylene supply and the ethylene receptor activity are maintained constant and uniform throughout the epidermis (Fig. 2b, and see discussion ahead).

Finally, according to the available molecular evidence in both normal and abnormal division patterns, *TTG* should be constitutively expressed in the root epidermis, while *CPC* expression should be induced only in cells overlying two cortical cells (see Fig. 2b). Our model incorporates both hypotheses. The constitutive expression of *TTG* could be maintained by a positive self-regulation after its induction during embryonic development. One possible way to attain the localized *CPC* activation is by means of a diffusible signal such as ethylene, capable of crossing between the junction of longitudinal anticlinal cell walls, but not through them (Tanimoto *et al.*, 1995). However, a more plausible mechanism could involve a molecule specifically expressed in the anticlinal cell walls, capable of inducing the activation of *CPC* in the contacting epidermal cells.

SIMULATIONS

Fig. 3 shows the transition tables for each one of the elements included in the regulatory network. Half of the network elements are binary variables, while network elements involved in the ethylene-response pathway are modeled as three-state variables. The three-state elements enable the simulation of the differential response of roots to ACC (the direct precursor of ethylene) or AVG (an

inhibitor of ethylene synthesis) treatments. All the tables in Fig. 3 could be united into a single transition table for the entire network, thus defining the entire dynamical behavior of the system. However, this is not practical because a complete table would have 1296 rows. Nevertheless, we present in Table 1 the final activation states associated with every initial state (*i.e.*, attractors and their basins of attraction), excluding the transient states. In the model, six basins of attraction are reached. Two of them correspond to the wild type stable activation states of atrichoblasts and trichoblasts. The other four correspond to altered patterns of root hair formation due to experimentally induced changes in ethylene availability. For comparison with the "normal" model, we also include in Table 1 the attractors of all the single loss-of-function mutants. Such mutants might be correlated with a particular phenotype with the aid of Table 2 (see ahead). Multiple mutants are not shown, but they can be derived from the transition tables in Fig. 3.

The initial state of *CPC* and ETR1 alone determine which attractor is reached by the model. Two of the final stable states correspond to the expression patterns observed in wild type trichoblasts and atrichoblasts. To simulate epidermal cells overlying a longitudinal anticlinal wall, *CPC* and ETR1 initial states are set to 1, and for the epidermal cells contacting only one cortex cell *CPC* and ETR1 are set to 0 and 1, respectively. In the first type of cell, which corresponds to a trichoblast, the attractor reached is 11101121; while in the second one that corresponds to an atrichoblast, the attractor is 11011100 (Table 1). The first attractor (11101121) includes an active CPC variable, RHD6 equals 2 and AXR2 equals 1. Based on Table 2, these activation states of RHD6 and AXR2 yield a phenotype where 90% of the cells in a file develop a hair. In the second attractor (11011100), CPC, RDH6 and AXR2 are equal to 0. In this particular case, according to Table 2, none of the cells in the file with that stable activation state develop a hair. These results of the model match the phenotype of wild type roots in which most trichoblasts develop a hair while no atrichoblast do.

The other four attractors correspond to the simulation of pharmacological treatments on the root under a wild type genetic background. As mentioned earlier, we may simulate ethylene or ACC treatment by setting the variable ETR1 to zero. In this case the attractors 11100022 and 11010011 are reached in the trichoblasts and atrichoblasts, respectively. The extra availability of ethylene on the root induce all the cells that contact anticlinal walls to develop a hair. In contrast, only 30% of the cell files that contact a periclinal wall are predicted to develop a hair. Similarly, in order to model AVG treatment we set the initial state of the variable ETR1 to 2. In this case the attractors reached are 11102210 and 11012200 for trichoblasts and atrichoblasts, respectively. As shown in Table 2, these attractors correspond to 20 and 0% of cells in a file developing a hair, respectively. The model seems to describe accurately the change in the hair to non-hair ratio due to a change in ethylene availability (Table 3).

Our regulatory network model, in contrast to previous root epidermis models (Scheres, 1996; 1997; Tanimoto *et al.*, 1995), can be used also to simulate gain- and loss-of-function mutations. In our model, mutations in all members of the ethylene receptor family (ETR1 = 0) cause the inactivation of CTR1 (= 0). In this case, the attractors reached by the model are 11100022 and 11010011, for cells overlying two and one cortical cells, respectively. According to Table 2, the phenotype corresponding to these two attractors are 100% of hair cells if over the anticlinal wall, and 30% of hair cells if not. The attractors of other mutations are also shown in Table1. Simulation results for other mutants are shown in Table 3, where we compare the attractors predicted by our model for trichoblasts and

atrichoblasts with published mutants and pharmacological treatments. Comparison with published results show that there is an overall good agreement among those and the simulated results. Finally, we summarize some simulation predictions for novel mutants and pharmacological treatments in Table 4.

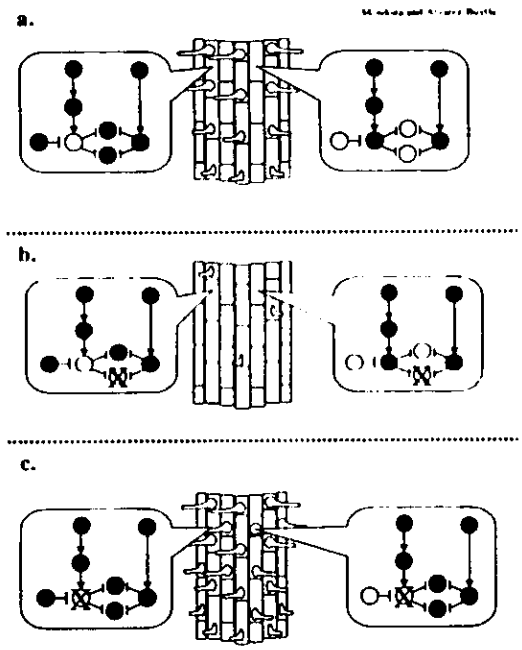


Figure 4

Finally, it is necessary to make clear how Table 2 was constructed since it links the activation state of the genetic network with a root phenotype. To obtain the phenotype associated to each attractor of the network model, either under wild type or mutant backgrounds, we used the activation states of the most downstream genes in the model (*RHD6* and *AXR2*). We have inferred the effect of these two genes on root hair development based the phenotypes of their single and double mutants, *i.e. rhd6*, *axr2* and *rhd6 axr2* plants. In Table 2 we propose a correlation between discretized values of the RHD6 and AXR2 variables with the percentage of hair-bearing cells in an epidermal file. Even though the expression levels of *RHD6* and *AXR2* are not known, Table 2 reflects the fact that these genes collaborate positively to the development of root hairs and that their combined mutation result in a total loss of root hairs (Masucci and Schiefelbein, 1994; 1996). Due to the scarce molecular evidence, and the large variability of the published results regarding hair count in the reported mutants, we did not elaborate Table 2 by using any optimization procedure. We rather chose values that roughly approximate the available information for single and multiple mutants involving *RHD6* and *AXR2* genes. The rightmost column of Table 2 shows the percentage of cells in a file that develop a hair, given that the activation state of the genetic network should be identical for all the cells forming that particular file. Notice that the response in Table 2 is not all-or-none, there is a gradual change in the appearance of root hairs depending on the values of variables AXR2 and RHD6. Moreover, the precise location of the hairs inside a file is not specified, thus describing a partially stochastic process. Fig. 4 exemplifies the gradual response in root hair formation under different genetic backgrounds.

There are two sources of variation that prevented us to make a close fit of values in Table 2 with experimental results. The first is the variability in the number of hairs per file, while the second is the variability of the number of the cell files themselves. While the number of epidermal files is maintained constant during root growth, their number varies from individual to individual. The ratio of trichoblasts to atrichoblasts normally ranges between 0.5 and 1 (Dolan *et al.*, 1993; 1994). For such reasons, authors report the number of hair cells as the fraction of the total number of epidermal cells, regardless of the number of cell files. This explanation is relevant to compare the numerical results of our model with those results already reported. For example, the rightmost column of Table 3 along the wild type line indicates that 42% of epidermal cells present hairs. It is not reported, however, the average number of files per individual. If the

quantification was made on plants having a trichoblast to atrichoblast ratio of 1, then the results would imply that 84% of the cells in a trichoblasts file present a hair. However, if the results were obtained with plants with a trichoblast to atrichoblast ratio of 0.43, then the same reported value should imply that all cells in the trichoblast file present a hair. Therefore, it is important to have in mind that our model presents results in a different way with respect of the results reported by the experimentalists.

DISCUSSION

Cellular interactions are important during plant morphogenesis. However, the precise molecular mechanisms entailed by such interactions are only starting to be understood. Such mechanisms include how cells react to signals, and how these are integrated among cells during morphogenesis (Meyerowitz, 1997). Integrative dynamical modeling approaches will become useful to make progress in this area. In this paper we have proposed a model for a mechanism by which cells transduce extra-cellular signals, via the genetic regulatory network that controls cell fate determination in the root epidermis.

We presented a simple discrete-state model. However, theoretical studies show that the number and type of stable states found in discrete systems are the same to those attained with appropriate equivalent continuous models (Bagley & Glass, 1996; Glass, 1975; Glass & Kauffman, 1973). It is noteworthy, that in contrast to our previous *Arabidopsis* network models, this one does not include any feedback loops. The presence of functional positive feedback loops is a pre-requisite to obtain multistationarity, which is necessary to attain different stable states that represent the genetic activation patterns observed in different cell types (Thieffry *et al.*, 1995; Thomas, 1991). In the network model presented here, cellular differentiation depends upon the initial activation states of two elements of the model, *CPC* and ETR1, which are assumed to depend on a signaling mechanism external to the network. This mechanism does not preclude the possibility that future studies uncover new genes and interactions including some positive functional loops, which may reinforce the cell patterning process simulated by the model that we present here.

Our model makes two basic assumptions concerning the signals involved in cell fate determination in the root epidermis. Both are based on experimental data that show that the contact of epidermal cells with a longitudinal anticlinal cortex wall is critical for root hair development (Scheres, 1996; 1997), and that the default state seems to be the non-hair cell fate (Berger *et al.*, 1998a). With such information we established that the initial state of *CPC* depends on the position of the epidermal cell relative to cortex cells. A similar mechanism has been proposed before by Scheres (1997). This assumption implies that an unknown signal activates *CPC* in those cells that will become trichoblasts. A cell wall component, for example a glycoprotein, rather than a diffusible molecule could be involved in such mechanism. A key prerequisite for that component is its localization along the longitudinal anticlinal walls of cortex cells. This type of expression pattern has been documented already in cells of *Arabidopsis* roots for some carbohydrates (Freshour *et al.*, 1996).

Second, we assume in our model that ethylene is a signaling molecule uniformly available to all epidermal cells. Thus, ethylene does not convey by itself positional information during sub-specification of epidermal cell-types. In contrast, some previous models assume that ethylene (or ACC) diffuses only through longitudinal anticlinal walls of cortical cells, inducing root hair development only in epidermal cells that are in direct contact with the cleft (Scheres, 1997; Tanimoto, 1995). The fact

that in the absence of *TTG* and *GL2*, hair formation in AVG-treated roots occurs in the right place, partially supports the pre-patterning role of ethylene. However, there are two experimental results that challenge such hypothesis. The first one comes from a trichoblast that divides longitudinally to yield two daughter cells, from which one becomes an atrichoblast (Berger *et al.*, 1998b). If a diffusible signal determined the trichoblast fate, then both daughter cells should attain this fate. One could still argue for a diffusible signal that could only maintain the trichoblast fate in one cell and not the other, but this mechanism does not seem a plausible one. The second piece of evidence is that an atrichoblast remains as such even if surrounding trichoblasts, atrichoblasts or cortical cells are ablated (Berger *et al.*, 1998a). With such experimental manipulation any gradient of a diffusible morphogen should have been destroyed thus increasing ethylene availability. This in turn would have caused that isolated atrichoblast differentiate into a trichoblast.

Previous authors have proposed static schematic representations of the gene interactions involved in cell-type determination in the *Arabidopsis* root epidermis (Scheres & Wolkenfelt, 1998; Schiefelbein, 1998; Schiefelbein *et al.* 1997). These authors have proposed the same architecture of interactions among the genes encoding transcription factors (*CPC*, *TTG*, *GL2*, and R homolog) that we present here. However, we propose that the ethylene response pathway is independent of the *GL2* pathway, although both converge in the negative regulation of downstream elements such as AXR2 and RHD6 (see Fig. 2a). In contrast, most previous authors have postulated that the ethylene response pathway is downstream and repressed by *GL2* (see for example Schiefelbein *et al.*, 1997). However, loss-of-function mutations, either on the ethylene receptor family or on *CTR1* (Hua & Meyerowitz, 1998; Kieber *et al.*, 1993) develop root hairs in all epidermal cells, evidencing that the ethylene response pathway is active even in cells expressing *GL2*. Therefore, *GL2* cannot be a repressor of the ethylene pathway (see Fig. 2a). Moreover, our architecture is also compatible with phenotypes of *ttg* and *gl2* mutants treated with the ethylene synthesis inhibitor, AVG (Masucci & Schiefelbein, 1996). It is noteworthy that in some previous papers these experiments have been used to postulate that the ethylene response pathway is downstream of *GL2* (Scheres, 1997; Schiefelbein *et al.*, 1997). Finally, loss-of-function mutants of the ethylene receptor family, namely the *etr1 etr2 ein4 ers2* quadruple mutants, have a constitutive ethylene response phenotype, like that of *ctr1* mutants (Hua & Meyerowitz, 1998). This phenotype is also compatible with the model presented here, but not for previous ones (see Tanimoto *et al.*, 1995). In brief, our model is a new proposition of the regulatory interactions between genes and proteins controlling root hair development. Notably, it establishes that the transduction pathway of ethylene is parallel, and not downstream, of the group of transcription factors that control root hair appearance. And finally, it is a dynamic model that can be used to describe and predict the stable expression patterns, under wild type, mutants and pharmacologically treated roots, of a group of proteins and molecules controlling root hair development.

The agreement of published and simulated results suggests that this first dynamical model captures key elements involved in the genetic regulation of cell fate determination in the epidermis. Notice the overall agreement among results of the model and the published experimental data presented in Table 3. It is important to consider that the experimental results have a large standard deviation that we did not included in the table (see for example, Masucci and Schiefelbein, 1994; 1996). Still, the phenotypes predicted by the model lie most of the times near the average observed phenotypes but for two cases. The model does not reproduce the

reduction in the number of root hairs observed for the double mutants *ttg axr2* and *gl2 axr2*. Notice, however, that the percentage of ectopic hairs is more accurate. A partial solution for the discrepancies would be solved by including a non-zero activity for *GL2* in the absence of the normal activation coming from *TTG*. Even though there is experimental evidence for such activity (Hung *et al.*, 1998), we did not include that characteristic because it seems to come from an unidentified gene not accounted for in our model, as discussed in a previous section. Such *TTG*-independent *GL2* activation might contribute to a reduction in the *RHD6* activity with a concomitant reduction in the predicted number of hairs in the *ttg axr2* mutant, thus approaching better the published results. Nevertheless, the incorporation of such influence over *GL2* would not explain the discrepancies of the model with the *gl2 axr2* mutants. In this particular case, the presence of an inhibitory pathway over *RHD6* independent of *TTG/GL2* might explain the inaccuracy of the model. In this case we do not posses evidence of such a pathway, but it might be used for the experimentalist as a clue for future work.

We still need much more molecular and genetic data to further elaborate the model. For example, it is known that the gene *ROOT HAIRLESS 1 (RHL1)* is an important gene involved in root hair formation (Schneider *et al.*, 1998). Nevertheless, we were unable to include this and other genes because the appropriate multiple mutations and molecular evidence are still unavailable. Furthermore, available data already suggests that there are missing genes in the network proposed here. For example, the phenotype of *cpc ttg* double mutants suggests that *GL2* has a basal expression level in the absence of the inhibition of *CPC* and the activation of *TTG/R*-homolog pathway. This suggests that there is another activator of *GL2* yet to be discovered.

There is another relevant issue regarding the interpretation of our model. The network includes many of the regulatory genes that control the appearance of root hairs, but does not incorporate the most downstream target genes and biochemical pathways that directly affect the morphological cellular processes underlying hair formation. Thus, a stable activation pattern of the network represents the activity only of some key regulatory genes. These regulatory genes trigger the activation and inhibition of many non-modeled downstream pathways, which in turn determine the cellular fate of the epidermal cell. For example, the activity of the *CPC* network element results in the inhibition of *GL2*, causing the de-repression of AXR2 and RHD6, which form part of the proteins leading to the appearance of a root hair. In other words, the molecular events determine the cellular fate. In contrast, some authors imply in their schematic representations that the cellular fate determines the activation state of some of the genes that are themselves responsible for cell fate determination (see for example Fig. 1 of Schiefelbein, 1998), such circular schematic models may lead to considerable confusion.

Network models like the one presented here are not appropriate for analyzing transient states of cell differentiation, nor the explicit spatial and cellular context of development. Cell differentiation in the root epidermis is a continuous process. Macroscopic differences between trichoblasts and atrichoblasts appear early during development (Dolan *et al.*, 1993; 1994). As development continues, epidermal cells stop dividing and enlarge differentially. Finally, epidermal cells cease to grow and trichoblasts develop root hairs, thus defining the differentiation root zone. At this stage, epidermal cells reach a stable state and are considered to be mature. Different genes seem to control different stages of this process. For example, while *TTG* seems to act early,

ethylene seems to act late (see Scheres, 1996; 1997). However, once cells reach the stable mature state of differentiation, the activity of such genetic products should also reach a stable steady state. Discrete network models are suitable for the prediction of such stable genetic regulatory patterns observed in mature cells. In this case, those steady states correspond to the genetic profiles of trichoblasts and atrichoblasts in the root epidermis. Therefore, the morphological interpretation of our genetic regulatory network corresponds only to the differentiation region of the root. Consequently, our model may not simulate the transitory stages of development and does not consider explicitly the gradient of cellular differentiation that exists from the root meristem towards the hypocotyl.

The rapid progress in genetic and experimental approaches will eventually provide the data to detail the molecular pathways and cellular mechanisms that determine cell fate, cell shape and cell division during organ development. Also, they will help to understand how these pathways are integrated and modulated by external signals. However, formal approaches will become necessary to provide rigorous spatio-temporal frameworks to integrate data, and also to elaborate predictions that link the genetic circuitry underlying development with the final morphology of an organism. Network models are useful to this end by integrating these models with others that explicitly consider the spatial and cellular scales. In particular, Mjolsness *et al.* (1991) proposed a system made of a neural network coupled to a grammar to model cellular differentiation. However, those authors recognized that they do not have yet a specific biological system in which to apply their methodology. In this paper, we showed that since cellular fate in the root is largely determined by positional information, it is possibility of developing spatially explicit models using computational tools such as cellular automata (Ermentrout & Edelstein-Keshet, 1993) or L-systems (Prusinkiewicz & Lindenmayer, 1990).

ACKNOWLEDGEMENTS

Thanks to Francisca Acevedo for her comments in the preparation of this manuscript. This work was supported by an UNAM scholarship to L.M., and a grant from the Mexican Council of Science and Technology (CONACYT 3667 P-N) to E.R.A.B.

REFERENCES

Bagley, R.J. & Glass, L. (1996). Counting and classifying attractors in high dimensional dynamical systems. *J. theor. Biol.* **183**, 269-284.

Berger, F.; Haseloff, J.; Schiefelbein, J. & Dolan, L. (1998a). Positional information in root epidermis is defined during embryogenesis and acts in domains with strict boundaries. *Curr. Biol.* **8**, 421-430.

Berger, F.; Hung, C.Y.; Dolan, L. & Schiefelbein, J. (1998b). Control of cell division in the root epidermis of *Arabidopsis thaliana*. *Dev. Biol.* **194**, 235-245.

Bodnar, J.W. (1997). Programming the *Drosophila* embryo. *J. theor. Biol.* **188**, 391-445.

Clark, K.L.; Larsen, P.B.; Wang, X., & Chang, C. (1998). Association of the *Arabidopsis* CTR1 Raf-like kinase with the ETR1 and ERS ethylene receptors. *Proc. Natl. Acad. Sci. USA* **95**, 5401-5406.

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

Di Christina, M.; Sessa, G.; Dolan, L.; Linstead, P.; Baima, S.; Roberti, I. & Morelli, G. (1996). The *Arabidopsis* Athb-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development. *Plant J.* **10**, 393-402.

Dolan, L.; Janmaat, K.; Willemsen, V.; Linstead, P.; Poethig, S.; Roberts, K. & Scheres, B. (1993). Cellular

organization of the *Arabidopsis thaliana* root. *Development* **119**, 71-84.

Dolan, L.; Duckett, C.M.; Grierson, C.; Linstead, P.; Schneider, K.; Lawson, E.; Dean, C.; Poethig, S. & Roberts, K. (1994). Clonal relationships and cell patterning in the root epidermis of *Arabidopsis*. *Development* **120**, 2465-2474.

Ermentrout, G.B. & Edelstein-Keshet, L. (1993). Cellular automata approaches to biological modeling. *J. theor. Biol.* **160**, 97-133.

Fluhr, R. (1998). Ethylene perception: from two-component signal transducers to gene induction. *Trends Plant Sci.* **3**, 141-146.

Freshour, G.; Clay, R.P.; Fuller, M.S.; Albersheim, P.; Darvill, A.G. & Hahn, M.G. (1996). Developmental and tissue-specific structural alterations of the cell-wall polysaccharides of *Arabidopsis thaliana* roots. *Plant Physiol.* **110**, 1413-1429.

Galway, M.E.; Masucci, J.D.; Lloyd, A.M.; Walbot, V.; Davis, R.W. & Schiefelbein, J.W. (1994). The *TTG* gene is required to specify epidermal cell fate and cell patterning in the *Arabidopsis* root. *Dev. Biol.* **166**, 740-754.

Glass, L. (1975). Classification of biological networks by their qualitative dynamics. *J. theor. Biol.* **54**, 85-107.

Glass, L. & Kauffman, S.A. (1973). The logical analysis of continuous, non-linear biochemical control networks. *J. theor. Biol.* **39**, 103-129.

Hua, J.; Chang, C.; Sun, Q. & Meyerwitz, E.M. (1995). Ethylene insensitivity conferred by *Arabidopsis ERS* gene. *Science* **269**, 1712-1714.

Hua, J. & Meyerowitz, E.M. (1998). Ethylene responses are negatively regulated by a receptor gene family in *Arabidopsis thaliana*. *Cell* **94**, 261-271.

Hung, C.Y.; Lin, Y.; Zhang, M.; Pollock, S.; Marks, M.D. & Schiefelbein, J. (1998). A common position-dependent mechanism controls cell-type patterning and *GLABRA2* regulation in the root and hypocotyl epidermis of *Arabidopsis*. *Plant Physiol.* **117**, 73-84.

Johnson, P.R. & Ecker, J.R. (1998). The ethylene gas signal transduction pathway: a molecular perspective. *Annu. Rev. Genet.* **32**, 227-254.

Kauffman, S.A. (1993). *The origins of order: self-organization and selection in evolution*. Oxford University Press.

Kieber, J.J. (1997). The ethylene response pathway in Arabidopsis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**, 227-296.

Kieber, J.J.; Rothenberg, M.; Roman, G.; Feldmann, K.A. & Ecker, J.R. (1993). *CTR1*, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the Raf family of protein kinases. *Cell* **72**: 427-441.

Masucci, J.D.; Rerie, W.G.; Foreman, D.R.; Zhang, M.; Galway, M.E.; Marks, M.D. & Schiefelbein, J.W. (1996). The homeobox *GLABRA2* is required for position-dependent cell differentiation in the root epidermis of *Arabidopsis thaliana*. *Development* **122**, 1253-1260.

Masucci, J.D. & Schiefelbein, J.W. (1996). Hormones act downstream of *TTG* and *GL2* to promote root hair outgrowth during epidermis development in the *Arabidopsis* root. *Plant Cell* **8**, 1505-1517.

Mendoza, L. & Alvarez-Buylla, E.R. (1998). Dynamics of the genetic regulatory network for *Arabidopsis thaliana* flower morphogenesis. *J. theor. Biol.* **193**, 307-319.

Mendoza, L.; Thieffry, D. & Alvarez-Buylla E.R. (1999). Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis. *Bioinformatics*. **In press**.

Mjolsness, E.; Sharp, D.H.; & Reinitz, J. (1991). A connectionist model of development. *J. theor. Biol.* **152**: 429-453.

Meyerowitz, E.M. (1994). The genetics of flower development. *Sci. Am.* **271**, 40-47.

Meyerowitz, E.M. (1997). Plants and the logic of development. *Genetics* **145**, 5-9.

Pigliucci, M. (1996). How organisms respond to environmental changes: from phenotypes to molecules

(and vice versa). *Trends Ecol. Evol.* **11**, 168-173.

Prusinkiewicz, P. & Lindermayer, A. (1990). *The algorithmic beauty of plants*. New York: Springer-Verlag.

Reinitz, J. & Sharp, D.H. (1995). Mechanism of formation of eve stripes. *Mech. Dev.* **49**, 133-158.

Scheres, B. (1996). Embryo patterning genes and reinforcement cues determine cell fate in the *Arabidopsis thaliana* root. *Seminars Cell & Dev. Biol.* **7**, 857-865.

Scheres, B. (1997). Cell signaling in root development. *Curr. Op. Genet. & Develop.* **7**, 501-506.

Scheres, B. & Wolkenfelt, H. (1998). The *Arabidopsis* root as a model to study plant development. *Plant Physiol. Biochem.* **36**, 21-32.

Schiefelbein, J. (1998). Cell-fate specification in the root epidermis. *Trends Plant Sci.* **3**, 3-4.

Schiefelbein, J.W.; Masucci, J.D. & Wang, H. (1997). Building a root: the control of patterning and morphogenesis during root development. *Plant Cell* **9**, 1089-1098.

Schneider, K.; Mathur, J.; Boudonck, K.; Wells, B.; Dolan, L. & Roberts, K. (1998). The *ROOT HAIRLESS 1* gene encodes a nuclear protein required for root hair initiation in *Arabidopsis*. *Genes Dev.* **12**, 2013-2021.

Tanimoto, M.; Roberts, K. & Dolan, L. (1995). Ethylene is a positive regulator of root hair development in *Arabidopsis thaliana*. *Plant J.* **8**, 943-948.

Thieffry, D.; Snoussi, E.H.; Richelle, J. & Thomas, R. (1995). Positive loops and differentiation. *J. Biol. Syst.* **3**, 457-466.

Thomas, R. (1991). Regulatory networks seen as asynchronous automata: a logical description. *J. theor. Biol.* **153**, 1-23.

Van der Berg, C.; Weisbeek, P. & Scheres, B. (1998). Cell fate and cell differentiation status in the *Arabidopsis* root. *Planta* **205**, 483-491.

Wada, T.; Tachibana, T.; Shimura, Y. & Okada, K. (1997). Epidermal cell differentiation in *Arabidopsis* determined by a Myb homolog, *CPC*. *Science* **277**, 1113-1116.

Wada, T.; Tachibana, T.; Shimura, Y. & Okada, K. (1998). Expression pattern of CPC and its role in the differentiation of root hairs. In: *Proceedings of the 9th International Conference on Arabidopsis Research*. Pp 487.

Weigel, D. & Meyerowitz, E.M. (1993). Activation of floral homeotic genes in *Arabidopsis*. *Science* **261**, 1723-1726.

TABLE 1. Basins of attraction and attractors of the network model.

Mutants	Basins of attraction (initial states)	Attractors (final steady states)
Wild type	xx0x1xxx	11011100 *
	xx1x1xxx	11101121 **
	xx0x xxx	11010011 @
	xx1x xxx	11100022 @@
	xx.x2xxx	11.122.0 #
	xx1x2xxx	11102210 ##
ttg	xx0x1xxx	00001121 *
	xx1x1xxx	00101121 **
	xx0x0xxx	00000022 @
	xx1x xxx	00100022 @@
	xx.x2xxx	00002210 #
	xx1x2xxx	00102210 ##
R homolog	xx0x1xxx	10001121 *
	xx1x1xxx	10101121 **
	xx0x0xxx	10000022 @
	xx1x xxx	10100022 @@
	xx.x2xxx	10002210 #
	xx1x2xxx	10102210 ##
cpc	xxxx1xxx	11011100 *
	xxxx0xxx	11010011 @
	xxxx2xxx	11012200 #
gl2	xx0x1xxx	11001121 *
	xx1x1xxx	11101121 **
	xx0x0xxx	11000022 @
	xx1x xxx	11100022 @@
	xx0x2xxx	11002210 #
	xx1x2xxx	11102210 ##
etr1	xx0xxxxx	11010011 *
	xx1xxxxx	11100022 **
ctr1	xx0x1xxx	11011011 *
	xx1x1xxx	11101022 **
	xx0x0xxx	11010011 @
	xx1x0xxx	11100022 @@
	xx0x2xxx	11012011 #
	xx1x2xxx	11102022 ##
rhd6	xx0x1xxx	11011100 *
	xx1x1xxx	11101121 **
	xx0x0xxx	11010001 @
	xx1x0xxx	11100002 @@
	xx0x2xxx	11012200 #
	xx1x2xxx	11102200 ##
axr2	xx0x1xxx	11011100 *
	xx1x1xxx	11101120 **
	xx0x0xxx	11010010 @
	xx1x0xxx	11100020 @@
	xx0x2xxx	11012200 #
	xx1x2xxx	11102210 ##

We present simulation results for wild type, and the loss-of-function mutants of all the network elements (left column). The central column contains the basins of attraction (initial states) that lead to the corresponding attractors (final steady states) in the right column. The state of the entire network is represented by a vector containing the activation states of variables TTG, R homolog (Rh), CPC, GL2, ETR1, CTR1, RHD6 and AXR2, always in this order. In the basins of attraction the initial state of the variables can be 0, 1, 2 or X (indicating any of the possible values). In the attractors, variables attain an activation state of 0, 1 or 2. Attractors marked with *, @ and #, are those obtained in the simulations under normal ethylene, ACC and AVG treatments, respectively. Single and double symbols indicate attractor obtained in epidermal cells overlying one or two cortical cells, respectively. Thus, the genetic activation patterns of wild type atrichoblasts and trichoblasts are marked with asterisks in the "wild type" row.

TABLE 2. Phenotypic effect on root hair development of all the possible activation states of RHD6 and AXR2, the two most downstream elements of the network model.

RHD6	AXR2	% of hair-bearing cells in a file
0	0	0
0	1	10
0	2	70
1	0	20
1	1	30
1	2	90
2	0	70
2	1	90
2	2	100

Not all cells in a file develop a hair; the table presents the percentage of cells in a particular file that develops a root hair depending on the activation state of AXR2 and RHD6, which can be 0, 1 or 2. The table assumes that all the cells of a particular cell file have the same activation state of the corresponding regulatory network. Notice that the appearance of root hairs is a gradual response to the activity of AXR2 and RHD6. Figure 4 contains a schematic representation of the root hair morphology under different genetic backgrounds.

Table 2

TABLE 3. Comparison of network simulation predictions with experimental published data.

Mutant	Cells over two cortical cells (%)	Cells over one cortical cell (%)	Percentage of hair-cells with respect to the WT model	Experimental evidence. Average percentage of hair-cells with respect to wild type (WT) unless otherwise stated.
WT	11101121 (90)	11011100 (0)	0% ectopic.	42% of epidermal cells have a hair, 1% ectopic (Masucci and Schiefelbein, 1996) 1.75% ectopic (Di Cristina *et al.*, 1996). 0% ectopic (Tanimoto *et al.*, 1995).
WT + ACC	11100022 (100)	11010011 (30)	144.4-177.7%. 23-37.5% ectopic.	126 % than without ACC. 22% ectopic (Masucci and Schiefelbein, 1996) 35-50% ectopic (Tanimoto *et al.*, 1995).
WT + AVG	11102210 (20)	11012200 (0)	22.2% 0% ectopic.	26% than without AVG. 0% ectopic (Masucci and Shiefelbein, 1996). 0-15% than without AVG. 0% ectopic (Tanimoto *et al.*, 1995).
ttg	00101121 (90)	00001121 (90)	200-300%. 50-66.6% ectopic.	223%. 51% ectopic (Masucci and Schiefelbein, 1996) 227% (Wada *et al.*, 1997). 200% (Hung *et al.*, 1998).
cpc	11011100 (0)	11011100 (0)	0%	24% (Wada *et al.*, 1997).
35S::CPC	11101121 (90)	11101121 (90)	200-300%. 50-66.6% ectopic.	312% than WT (Wada *et al.*, 1997).
gl2	11101121 (90)	11001121 (90)	200-300%. 50-66.6% ectopic.	230%. 53% ectopic (Masucci and Schiefelbein, 1996) 250% (Wada *et al.*, 1997) 210% (Hung *et al.*, 1998).
ctr1	11101022 (100)	11011011 (30)	144.4-177.7%. 23-37.5% ectopic.	147%. 32% ectopic (Masucci and Shiefelbein, 1996)
rhd6	11101101 (10)	11011100 (0)	11.1% 0% ectopic.	9%. 3% ectopic (Masucci and Schiefelbein, 1994, 1996)
rhd6 + ACC	11100002 (70)	11010001 (10)	88.8-100%. 12.5-22.2% ectopic.	800% than without ACC. 61% than WT + ACC (Masucci and Schiefelbein, 1994, 1996).
axr2	11101120 (70)	11011100 (0)	77.7% 0% ectopic.	64%. 0% ectopic (Masucci and Schiefelbein, 1996).
axr2 + ACC	11100020 (70)	11010010 (20)	100-122.2%. 22.2-36.3% ectopic.	64% than WT + ACC. 114 % than without ACC (Masucci and Schiefelbein, 1996)
ttg cpc	00001121 (90)	00001121 (90)	200-300%. 50-66.6% ectopic.	150% (Wada *et al.*, 1997).
ttg gl2	00101121 (90)	00001121 (90)	200-300%. 50-66.6% ectopic.	220%. 56% ectopic (Hung *et al.*, 1998).
cpc gl2	11001121 (90)	11001121 (90)	200-300%. 50-66.6% ectopic.	232% (Wada *et al.*, 1997)
ttg rhd6	00101101 (10)	00001101 (10)	22.2-33.3%. 50-66.6% ectopic.	54%. 13% ectopic (Masucci and Schiefelbein, 1996)
ttg rhd6 + ACC	00100002 (70)	00000002 (70)	155.5-233.3%. 50-66.6% ectopic.	170% (Masucci and Schiefelbein, 1996).
ttg axr2	00101120 (70)	00001120 (70)	155.5-233.3%. 50-66.6% ectopic.	59%. 40% ectopic (Masucci and Schiefelbein, 1996)
gl2 rhd6	11101101 (10)	11001101 (10)	22.2-33.3%. 50-66.6% ectopic.	40%. 4% ectopic (Masucci and Schiefelbein, 1996).
gl2 rhd6 + ACC	11100002 (70)	11000002 (70)	155.5-233.3%. 50-66.6% ectopic.	523% than without ACC (Masucci and Schiefelbein, 1996)
gl2 axr2	11101120 (70)	11001120 (70)	155.5-233.3%. 50-66.6% ectopic.	66%. 25% ectopic (Masucci and Schiefelbein, 1996).
ctr1 rhd6	11101002 (70)	11011001 (10)	88.8-100%. 12.5-22.2% ectopic.	Like WT (Masucci and Schiefelbein, 1996).
rhd6 axr2	11101100 (0)	11011100 (0)	0%	0% (Masucci and Schiefelbein, 1996).
rhd6 axr2 + ACC	11100000 (0)	11010000 (0)	0%	0% (Masucci and Schiefelbein, 1996).

Mutants are listed in the left column. The second and third columns list the attractor predicted for cells contacting one or two cortical cells, respectively. The predicted percentage of hair bearing cells (see Table 2) is presented in parentheses. The third column presents the predicted phenotype as the percentage of hair cells with respect to the wild type model. Finally, the fourth column lists experimental results. Notice that many of the experimental results are reported as phenotypes with respect of pharmacological treatments. WT = wild type; ectopic hairs = root hair cells developed outside their normal position. Of the 24 phenotypes included in the table, the model describes accurately 15 of them, with the predicted number of hairs lying at ≤15% of the reported average. Seven of the phenotypes (*cpc*, *rhd6*, *rhd6*+ACC, *axr2*, *axr2*+ACC, *ttg cpc* and ttg *rhd6*) described only qualitatively by the model, with the numerical prediction at ≤50% of the reported average of root hairs. Finally, only two of the phenotypes (*ttg axr2* and *gl2 axr2*) are described badly by the model, the probable cause of such poor fitting and the solution are explained in the discussion.

TABLE 4. Predictions of the regulatory network for mutants and pharmacological treatments not reported yet.

Mutant	Cells over two cortical cells	Cells over one cortical cell	Predicted phenotype. Percentage of hair-cells with respect to wild type model.
ttg + ACC	00100022 (100)	00000022 (100)	222-333%. 50-66.6% ectopic.
ttg + AVG	00102210 (20)	00002210 (20)	44.4-66.6%. 50-66.6% ectopic.
cpc + ACC	11010011 (30)	11010011 (30)	66.6-100%. 50-66.6% ectopic.
cpc + AVG	11012200 (0)	11012200 (0)	0%.
35S::CPC + ACC	11100022 (100)	11100022 (100)	222-333%. 50-66.6% ectopic.
35S::CPC + AVG	11102210 (20)	11102210 (20)	44.4-66.6%. 50-66.6% ectopic.
gl2 + ACC	11100022 (100)	11000022 (100)	222-333%. 50-66.6% ectopic.
gl2 + AVG	11102210 (20)	11002210 (20)	44.4-66.6%. 50-66.6% ectopic.
ctrl + ACC	11100022 (100)	11010011 (100)	222-333%. 50-66.6% ectopic.
ctrl + AVG	11102022 (100)	11012011 (100)	222-333%. 50-66.6% ectopic.
rhd6 + AVG	11102200 (0)	11012200 (0)	0%.
axr2 + AVG	11102210 (20)	11012200 (0)	22.2%. 0% ectopic.
ttg cpc + ACC	00000022 (100)	00000022 (100)	222-333%. 50-66.6% ectopic.
ttg cpc + AVG	00002210 (20)	00002210 (20)	44.4-66.6%. 50-66.6% ectopic.
ttg gl2 + ACC	00100022 (100)	00000022 (100)	222-333%. 50-66.6% ectopic.
ttg gl2 + AVG	00102210 (20)	00002210 (20)	44.4-66.6%. 50-66.6% ectopic.
cpc gl2 + ACC	11000022 (100)	11000022 (100)	222-333%. 50-66.6% ectopic.
cpc gl2 + AVG	11002210 (20)	11002210 (20)	44.4-66.6%. 50-66.6% ectopic.
ttg rhd6 + AVG	00102200 (0)	00002200 (0)	0%.
ttg axr2 + ACC	00100020 (70)	00000020 (70)	155.5-233.3%. 50-66.6% ectopic.
ttg axr2 + AVG	00102210 (20)	00002210 (20)	44.4-66.6%. 50-66.6% ectopic.
gl2 rhd6 + AVG	11102200 (0)	11002200 (0)	0%.
gl2 axr2 + ACC	11100020 (70)	11000020 (70)	155.5-233.3. 50-66.6% ectopic.
gl2 axr2 + AVG	11102210 (20)	11002210 (20)	44.4-66.6%. 50-66.6% ectopic.
ctrl rhd6 + ACC	11100002 (70)	11010001 (10)	88.8-100%. 12.5-22.2% ectopic.
ctrl rhd6 + AVG	11102002 (70)	11012001 (10)	88.8-100%. 12.5-22.2% ectopic.
rhd6 axr2 + AVG	11102200 (0)	11012200 (0)	0%.
cpc gl2 + ACC	11000022 (100)	11000022 (100)	222-333%. 50-66.6% ectopic.
cpc gl2 + AVG	11002210 (20)	11002210 (20)	44.4-66.6%. 50-66.6% ectopic.
35S::CPC rhd6	11101101 (10)	11101101 (10)	22.2-33.3%. 50-66.6% ectopic.
35S::CPC rhd6 + ACC	11100002 (70)	11100002 (70)	155.5-233.3%. 50-66% ectopic.
35S::CPC rhd6 + AVG	11102200 (0)	11102200 (0)	0%.

The table contains the same kind of information as the first four columns in Table 3.

CAPTIONS

FIGURE 1. Schematic representation of transverse (left) and longitudinal (right) midsections of the *Arabidopsis thaliana* root. The primary root is formed by concentric layers of lateral root cap, epidermis, cortex, endodermis, and pericycle surrounding a vascular bundle. The transverse section represents a region where the lateral root cap detaches and root hairs begin to grow. Epidermal cells contacting one and two cortical cells differentiate into atrichoblasts and trichoblasts, respectively. Mature trichoblasts develop a root hair.

FIGURE 2. a. Architecture of the genetic regulatory network model for hair development in the *Arabidopsis* root epidermis. Arrows and blunt-lines represent activation and repression, respectively. **b.** Assumption on signals affecting the activity of the regulatory network. Ethylene availability, acting upon its receptor ETR1, is assumed to be uniform across epidermal cells under normal conditions and pharmacological treatments (thin lines). An unknown signal expressed along the longitudinal anticlinal cell walls of the cortex cells (thick line and question mark) is assumed to activate *CPC* in cells overlying two cortical cells.

FIGURE 3. Transition tables for the network variables: **a**. Rh (the *Arabidopsis* homolog of the corn R element), **b**. GL2, **c**. CTR1, and **d**. RHD6 and AXR2. Variable TTG is assumed to be constitutively active (see main text), or always inactive if mutated. The state of variable ETR1 reflects the level of available ethylene; under wild type conditions ETR1 is set to 1. To simulate treatments with increased ethylene or ACC, ETR1 is set to zero. Conversely, a decrease of ethylene availability, or the addition of AVG, is achieved by setting ETR1 to 2. Finally, CPC is set to 1 for cells in direct contact with two cortical cells, and set to 0 in those contacting only one cortical cell.

FIGURE 4. Schematic representation of root hair phenotype in: **a.** wild type plants, **b.** *rhd6* mutants, and **c.** *gl2* mutants. Steady state activation patterns of the genetic regulatory network in epidermal cells overlying two or one cortical cells are shown on the leftmost and rightmost columns, respectively. Root epidermis cell files are schematised in the central columns. Two mutants are exemplified: one that decreases (*rhd6*) and another that increases (*gl2*) the number of epidermal root hairs without changing the total number of epidermal cells. The topology of the genetic network corresponds to that presented in Fig. 2a. Network elements with an activation state equal to 0, 1 or 2 are represented as white, square-filled and black circles, respectively. Crosses indicate loss-of-function mutated genes.