

Redes, modelos y enseñanza de las biomatemáticas

♦ José Díaz
Elena Álvarez



El estudiante que elige una carrera como biología es generalmente una persona que no tiene interés alguno en las matemáticas. De hecho, llega con la idea preconcebida y errónea de que las matemáticas no le servirán para nada. En varias universidades donde se ofrece la carrera de biología, las matemáticas han sido borradas de la currícula de los estudiantes, dejando en ocasiones tan sólo algún curso de álgebra básica y de bioestadística.

El tiempo está demostrando que este es un grave error, ya que el estudiante de biología del siglo XXI deberá adquirir una sólida formación en matemáticas, física y computación, junto a una sólida formación como biólogo, ya sea en la investigación de campo o de laboratorio. Incluso áreas tan tradicionalmente alejadas de las matemáticas, como la botánica y la zoología, comienzan a necesitar de trabajos interdisciplinarios con un elevado contenido matemático y computacional.

Máquina de Turing

La trama de miles de reacciones químicas que suceden en la célula está organizada en redes de interacciones en donde las asas de retroalimentación mantienen la estabilidad de la misma. Estas redes poseen propiedades matemáticas que caracterizan su estructura y han sido ampliamente analizadas

por el grupo de Barabasi en el Centro de Investigación sobre Redes Complejas de la Universidad de Notre Dame. Los trabajos de este grupo han dejado en claro que hay moléculas que poseen ciertas características especiales que les permiten jugar un papel central en la coordinación de las actividades celulares. Se puede citar a las proteínas Ras y p53, entre otras, como ejemplos de moléculas proteicas que funcionan como nodos claves o interruptores y cuyos cambios conformacionales determinan el tipo de respuesta que la célula tiene. A estas moléculas altamente conectadas con diferentes vías de señalización se les denomina *hubs*. La red dinámica celular está conformada por estos nodos que unen diferentes subsistemas o subredes parcialmente autónomos denominados módulos. Finalmente estos módulos, cuya conformación va desde unas cuantas hasta un gran número de asas de retroa-

♦ Profesor-Investigador, Facultad de Ciencias, UAEM
Investigadora, Instituto de Ecología



limentación, tanto positivas como negativas, son los responsables de que el comportamiento celular sea robusto o estable frente a perturbaciones internas o externas aleatorias, aunque no es tan robusto frente a perturbaciones dirigidas a estos nodos hiperconectados.

Stuart Kauffman, médico norteamericano, ha desarrollado desde la década de 1960 la idea de que la red de regulación genética de los organismos puede ser modelada como una Máquina de Turing binaria, donde el estado de inactividad de un gen está representado por el número 0 y el estado de actividad por el número 1. De esta forma, a partir de ciertas condiciones iniciales o conjunto de ceros y unos que caracterizan el estado inicial del sistema o el arreglo de activaciones genéticas, más un conjunto de reglas lógicas que guían las interacciones entre estos genes, es posible formular un modelo que representa la dinámica de la red bajo estudio. Dicha red puede llegar a un estado estacionario estable de punto fijo, o a un ciclo límite en el espacio de fase de dimensión “n”, donde “n” representa el número de genes en la red de interacciones. De acuerdo con Kauffman, cada punto estable representa un fenotipo celular hacia el cual evoluciona el sistema en su desarrollo. Las ideas de Kauffman están resumidas en su ya clásico libro *Origins of Order: Self-Organization and Selection in Evolution*.¹

Las ideas de Kauffman han dado lugar a un amplio campo de investigación donde confluye un gran número de especialistas de diferentes ramas de la ciencia que tratan de desentrañar los secretos que guarda el genoma de los organismos. Uno de los desarrollos teóricos más claros inspirados en la propuesta de Kauffman para estudiar la dinámica de la diferenciación molecular se llevó a cabo en un laboratorio mexicano. En el Laboratorio de Genética Molecular, Desarrollo y Evolución de Plantas del Instituto de Ecología de la UNAM, desde mediados de la década de 1990 se han desarrollado modelos de redes de regulación genética con el fin de estudiar la determinación de los órganos florales durante el desarrollo de la *Arabidopsis thaliana*.

Estos modelos, que en sus versiones más recientes han incorporado hasta quince genes florales de *Arabidopsis thaliana*, han demostrado que la determinación de las células primordiales que darán lugar a sépalos, pétalos, estambres y carpelos es independiente de las condiciones iniciales de la red y es robusta a perturbaciones de las reglas lógicas de interacción entre genes.

Esto se demostró con un modelo discreto de tres estados simulado como una Máquina de Turing en un espacio de quince dimensiones. Esta red genética alcanza un conjunto de diez estados estacionarios estables o “atractores puntuales”, cuando se exploran todas las condiciones iniciales posibles. A

¹ Stuart A. Kauffman, *Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press, Oxford, 1993.

cada atractor de punto fijo se llega a partir de subconjuntos de estos estados iniciales, los cuales se conocen como “cuencas de atracción”. Cuatro de estos puntos fijos corresponden a combinaciones de activaciones de genes que experimentalmente se han encontrado en las células primordiales y que darán lugar a los sépalos, pétalos, estambres y carpelos.

Sin embargo, como todo modelo teórico, el tratamiento como una Máquina de Turing de una red genética tiene limitaciones, en este caso, para describir la dinámica espacio-temporal de determinación de los estados de activación genética que caracterizan los distintos segmentos en que se diferencia el primordio floral. Por ejemplo, este modelo discreto para una célula única no nos permite estudiar la secuencia con que los arreglos de activación genética, que corresponden a cada uno de los tipos celulares primordiales de la flor, aparecen en el tiempo. Este grupo de investigación está desarrollando actualmente modelos que permitan explorar las dinámicas espacio-temporales.

Otro aspecto importante de las respuestas celulares de las plantas que se han modelado en este grupo de investigación es la activación de los genes que subyacen, tras la triple respuesta de las plan-

tas “maltratadas”, a la fitohormona gaseosa etileno. Esta triple respuesta incluye el acortamiento y engrosamiento de la raíz, incremento en el número y grosor de los pelos radicales y formación de una especie de gancho en la punta de las plántulas. Tal respuesta al etileno parece estar controlada por una cascada de transducción de señales en la que está involucrado el gen ERF1 (*Ethylene Response Factor 1*), el cual a su vez activa una serie de genes que generan la triple respuesta total de la planta.

Los resultados experimentales encontrados por Chen y Bleecker indican que la curva de dosis-respuesta al etileno tiene una forma sigmoidea para las características fenotípicas de la variedad silvestre de *Arabidopsis*, pero que dicha forma está muy alargada entre las zonas de respuesta subumbral y de saturación de la curva.² El ajuste de esta curva indica que es inapropiado utilizar un modelo booleano del tipo de Máquina de Turing para explicar la respuesta fenotípica de las plantas al etileno.

Con el fin de desarrollar un modelo que explique toda esta gama de respuestas fenotípicas a la acción del etileno, se desarrolló un modelo matemático de esta vía de señalización,³ pero con la variante de utilizar un modelo de tipo Markov para

² Q. G. Chen y A. B. Bleecker, “Analysis of Ethylene Signal-Transduction Kinetics Associated with Seedling-Growth Response and Chitinase Induction in Wild-Type and Mutant *Arabidopsis*”, *Plant Physiology*, vol. 108, 1995, p. 597.

³ José Díaz y Elena R. Álvarez-Buylla, “A model of the ethylene signaling pathway and its gene response in *Arabidopsis thaliana*: pathway cross-talk and noise-filtering properties”, *Chaos. An interdisciplinary journal of nonlinear science*, vol. 16, ms. 023112, 11 de mayo de 2006, pp. 1-16.



describir el proceso de activación y desactivación de los genes. En este tipo de modelación nos preguntamos sobre la probabilidad de que cierto gen esté en el estado activo al tiempo $t + \Delta t$, dado que estuvo activo al tiempo “ t ” por acción de la fitohormona etileno. Este modelo describe adecuadamente el comportamiento observado por Chen y Bleecker y proporciona, además, una serie de predicciones acerca del comportamiento dinámico de esta vía de señalización que puede ser sometida a verificación experimental.

Este ejemplo muestra que la modelación del proceso de expresión y regulación genética como una Máquina de Turing no siempre resulta adecuada, sobre todo cuando se trata de procesos con modulación temporal continua. Sin embargo, los modelos discretos son los más sencillos desde el punto de vista computacional, lógico y matemático, y parecen muy adecuados para modelar la lógica de regulación genética que subyace tras la diferenciación celular. Por otro lado, los métodos estocásticos de modelación permiten explicar las características fenotípicas dependientes de tiempo y espacio y, además, efectuar predicciones cuantitativas más precisas que pueden ser verificadas experimentalmente. Estos métodos también nos están permitiendo formular modelos de redes dinámicas para estudiar la secuencia temporal de estados estacionarios.

En conclusión, se deja en claro cómo la biología de las células vegetales está avanzando también en parte gracias al uso de herramientas formales y

computacionales que en este siglo XXI seguramente tendrán un auge mayor.

Utilidad de las matemáticas para los biólogos

Como se ha visto, desde los intentos iniciales de Nicolás Rashevsky hasta los modelos matemáticos de hoy en día, el componente matemático, físico y computacional en el estudio de los procesos biológicos está tomando un papel central en el desarrollo de la biología, dando forma a lo que poco a poco se va conformando como una nueva rama sólida de esta ciencia que bien podría llamarse biología teórica y computacional.

De esta forma, la biología del siglo XXI se desarrollará como resultado de la interacción de sus dos brazos, el experimental y el teórico, que actuarán conjuntamente y en continua retroalimentación para desentrañar los todavía numerosos misterios que guarda el estudio de los sistemas vivos. Por lo tanto, es un error que el estudio de la física, las matemáticas y la computación sea excluido del currículo de los biólogos u otros científicos de áreas afines. No cabe duda que los mejores teóricos serán aquellos biólogos que manejen estas herramientas formales con gran destreza y naturalidad y a la vez posean la intuición para hacer preguntas biológicas relevantes que ayuden a avanzar en el entendimiento de los sistemas vivos. Las herramientas formales y de simulación serán otras tantas que los biólogos tendrán que manejar, al igual que el microscopio y la cámara electroforética, incluyendo otras para analizar las bases moleculares del comportamiento celular.

Dentro de la currícula del biólogo es necesario incluir cursos de matemáticas que le permitan desarrollar un pensamiento cuantitativo que, aunado al desarrollo de su intuición biológica, le proporcionen las herramientas necesarias para contribuir al desarrollo de la biología del siglo XXI. En primer lugar, debe incluirse un curso de álgebra general, tal vez como propedéutico, para subsanar las lagunas que puedan tener los alumnos aspirantes a ingresar a la escuela profesional. Posteriormente, debe impartirse un buen curso de cálculo, tanto diferencial como integral, que le proporcione las herramientas necesarias para la modelación de procesos continuos que cambian en tiempo y espacio. Este curso debe darse desde una perspectiva moderna, con fuerte apoyo computacional que le permita al alumno expresar los aspectos fundamentales de un problema en forma matemática, es decir, modelar el proceso y observar su variación temporal en tiempo real. Además, el uso de la computadora permitirá que el alumno aprenda a resolver numéricamente ciertos problemas complejos, con lo cual adquirirá la habilidad requerida para efectuar la simulación de procesos más complejos con significado biológico.

Posteriormente, un buen curso de álgebra lineal incrementará las habilidades cuantitativas y abstractas que le permitirán modelar en forma adecuada problemas relacionados con redes genéticas y metabólicas. El uso de la computadora en este curso permitirá que el alumno se familiarice con el álgebra de matrices, el uso de determinan-

tes y el cálculo de eigenvalores y eigenvectores, con lo cual podrá resolver problemas relacionados con la determinación de las cuencas de atracción y puntos estacionarios de redes genéticas regulatorias discretas (booleanas, principalmente).

A continuación, un curso de ecuaciones diferenciales, con fundamentos de álgebra lineal, le permitirá al alumno adquirir las herramientas necesarias para modelar procesos biológicos que varían en el tiempo y espacio en forma continua. En este curso aprenderá los fundamentos del análisis de estabilidad de sistemas lineales y no lineales, la representación de estos procesos en un espacio de fases continuo y la modelación y simulación de procesos biológicos complejos usando adecuadamente las herramientas numéricas disponibles.

El curso tradicional de probabilidad y estadística que se imparte a los biólogos debe dividirse en dos. En el primer curso deben cubrirse los fundamentos de probabilidad, con énfasis en una introducción a procesos estocásticos, con el fin de que, con el auxilio computacional adecuado, el alumno pueda modelar procesos biológicos estocásticos, como procesos de Markov, procesos bayesianos y otros de tipo probabilístico complejo que ocurren frecuentemente en áreas de la biología, como la genética, la ecología, la biología celular y la teoría de evolución, entre otras. En el curso de estadística, una vez separado del curso de probabilidad, podrán estudiarse más a fondo los cimientos del diseño y el análisis estadístico de experimentos.



Finalmente, en el caso de aquellos estudiantes con aptitud teórica, esta serie de cursos básicos se podrá complementar con cursos optativos adecuados. Por ejemplo, en la Facultad de Ciencias de la Universidad Autónoma del Estado de Morelos (UAEM) se ofrecen cursos como biología celular computacional, sistemas dinámicos, series de tiempo aplicadas al análisis de datos electroencefalográficos, biología matemática, entre otros. Estos cursos pueden ofrecerse tanto en el ámbito de licenciatura como de posgrado, en apoyo a otras escuelas y facultades que lo requieran.

Sin embargo, un problema fundamental en la enseñanza de las matemáticas para el área biológica es la falta de personal adecuado para impartirlas. Se requiere de profesores especializados con perfil multidisciplinario que puedan encargarse de ofrecer estos cursos. El número de profesionales con estas características es muy reducido en el país actualmente. Sin embargo, en el estado de Morelos hay un grupo capacitado para efectuar esta tarea. En el Instituto de Ciencias Físicas de la Universidad Nacional Autónoma de México (UNAM), *campus* Cuernavaca, se encuentran los doctores Aldana y Martínez Mekler, del Departamento de Dinámica no Lineal. En la Facultad de Ciencias de la UAEM se

encuentran los doctores Atakishiyeva, Baier y Díaz, entre otros. En entidades como el Distrito Federal el número de investigadores es más grande y existen unidades de investigación y enseñanza de las biomatemáticas en la Facultad de Ciencias, en el Instituto de Ecología y en el Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS) de la UNAM. En la Universidad Autónoma Metropolitana, unidad Iztapalapa, se encuentra el doctor Santiago, biólogo y matemático. En el IPN hay grupos de investigación en esta área, especialmente en la Escuela Superior de Física y Matemáticas (ESFM) y en el Centro de Investigación y Estudios Avanzados (Cinvestav). En Guanajuato hay especialistas en el área, principalmente en el Centro de Investigación en Matemáticas AC (Cimat). El número de grupos de investigación en las áreas de biología teórica, biomatemáticas y sistemas dinámicos es más reducido en otras partes de la República; sin embargo, existe ya una base de investigación que puede ser detonante tanto para impulsar el desarrollo de la biología del siglo XXI en México para formar el cuadro de biólogos teóricos que se encargarán de hacer crecer y florecer este aspecto de la ciencia de la vida durante los años venideros.