

# Predecir el comportamiento de una célula

Un modelo para una nueva especie, *Mycoplasma pneumoniae*, tal vez permita augurar la respuesta de un organismo ante cambios ambientales y genéticos

MARIA LLUCH SENAR

**Hace diez años**, cuando empecé a trabajar con *Mycoplasma*, sentía fascinación por el modo en que unos microorganismos que podían vivir fuera de una célula huésped, sin pared celular y con genomas tan reducidos, realizaban funciones vitales y complejas, como metabolizar nutrientes o regular la expresión génica, entre otras. Además, algunos de los miembros de la familia, como *M. genitalium* y *M. pneumoniae*, causan enfermedades comunes en humanos, como la uretritis y la neumonía atípica, respectivamente, males no muy graves pero que pueden provocar complicaciones importantes en pacientes inmunodeprimidos. Pese a la aparente simplicidad de estas bacterias, las conexiones entre los diferentes procesos que les permitían desempeñar estas funciones seguían siendo un misterio en aquel tiempo.

En 2007, en nuestro laboratorio dirigido por Luis Serrano, del Centro de Regulación Genómica, se planteó la siguiente pregunta: si lográsemos cuantificar todas las proteínas y reacciones metabólicas y definir las funciones de todos los genes que constituyen el genoma de un organismo, ¿podríamos integrar toda esta información en un modelo matemático que permitiese representar en un ordenador el comportamiento de la célula y predecir su respuesta frente a diferentes perturbaciones?

La era de las técnicas «-ómicas» supuso una revolución en la obtención de información cuantitativa de estos procesos celular-

res. *M. pneumoniae* se seleccionó como modelo de estudio en nuestro laboratorio por su reducido genoma. Desde 2009 nuestro grupo ha caracterizado esta bacteria por lo que respecta a su composición total de ADN (genoma), ARN (transcriptoma), proteínas (proteoma) y metabolitos (metaboloma). Se ha obtenido así información sobre la metilación del ADN, los niveles de ARN y proteínas y un mapa metabólico de las diferentes reacciones bioquímicas que se producen dentro de la célula.

La recopilación de todos estos datos supuso un reto y una enorme contribución para la comunidad científica, pero solo representaba el principio de una nueva etapa. Llegados a este punto, los esfuerzos de nuestro grupo se centraron en la interpretación de su significado biológico, con el fin de convertir toda esta información cuantitativa en un modelo matemático, en una célula «in silico». Judith Wodke y Tobias Mayer, de nuestro grupo, desarrollaron hace un año el primer modelo del metabolismo de *M. pneumoniae* que permitía predecir el efecto de la supresión de uno o varios genes en el crecimiento del microorganismo.

Paralelamente, el equipo de Markus Covert, en la Universidad Stanford, revolucionó el campo de la modelización cuando publicó lo que definió como el primer modelo computacional de un organismo vivo, en concreto, de *M. genitalium*. Dado que no se dis-

poderoso para dar sentido a esos datos desordenados y fragmentados. Por su naturaleza, una simulación permitiría deducir una serie de hipótesis sobre lo que ocurre dentro de una célula, las cuales se podrían transformar en algoritmos matemáticos. Los esquemas que suelen aparecer en los artículos donde se muestra que el factor X regula al gen Y «de algún modo» no son suficientemente precisos para un programa informático. Los programadores expresan esos procesos con ecuaciones (uno de los ejemplos más simples es  $Y = aX + b$ ), aunque tengan que obtener de meras estimaciones los valores de las variables, como  $a$  y  $b$ . Esta necesidad de precisión revela, en última instancia, los experimentos de laboratorio que deben emprenderse para llenar las lagunas de conocimiento sobre las velocidades de reacción u otras magnitudes.

Al mismo tiempo, quedaba claro que, una vez se hubiera verificado la exactitud de los modelos, estos podrían sustituir algunos experimentos, lo que limitaría el costoso trabajo de laboratorio a preguntas que no pueden responderse solo con simulaciones. Y los experimentos simulados que generasen resultados sorprendentes ayudarían a los expertos a priorizar sus investigaciones y a aumentar el ritmo de los descubrimientos. De hecho, los modelos representan un instrumento tan tentador para distinguir entre causa y efecto que, en 2001, Masaru Tomita, de la Universidad de Japón, afirmó que la simulación de una célula completa constituía «uno de los grandes retos del siglo XXI».

Cuando todavía realizaba el doctorado, me quedé impresionado por los primeros resultados obtenidos por los principales

modelizadores de células de aquel tiempo y me obsesioné con este gran reto. Incluso cuando monté mi propio laboratorio y me centré en el desarrollo de técnicas de imagen para células individuales, la idea seguía en mis pensamientos. Y entonces, en aquella vuelta a casa en bicicleta de aquel febrero, vi el camino para abordarla.

## DOS ENFOQUES DECISIVOS

Antes de simular el ciclo de vida de una especie microbiana con la exactitud suficiente como para replicar sus comportamientos complejos y realizar así nuevos descubrimientos, tendríamos que resolver tres problemas. Primero, necesitábamos codificar todas las funciones importantes (como el flujo de energía, nutrientes y productos de las reacciones en la célula —su metabolismo—, la síntesis y degradación de ADN, ARN y proteínas, o la actividad de una infinidad de enzimas) en fórmulas matemáticas y algoritmos informáticos. Segundo, debíamos construir un armazón en el que integrar todas esas funciones. El último problema era, desde muchos puntos de vista, el más difícil: establecer límites superiores e inferiores a los aproximadamente 1700 parámetros del modelo de manera que adoptasen valores que se ajustaran a la realidad (o que al menos se situaran en el rango adecuado).

Comprendí que, por muy a fondo que revisáramos toda la bibliografía acerca de *M. genitalium* y sus parámetros (Karr, Sanghvi y yo terminamos pasando dos años seleccionando datos de casi 900 artículos científicos), en algunos casos tendríamos que hacer estimaciones o emplear valores experimentales de

ponía de los datos experimentales requeridos para el modelo de *M. genitalium*, los autores tuvieron que recurrir a parámetros pertenecientes a otras bacterias filogenéticamente próximas, aunque ello no siempre fue posible. Pese al increíble logro que ha supuesto el desarrollo de un modelo con una estructura matemática tan coherente, una de las principales críticas que ha recibido es que no presenta suficiente robustez, dado que los parámetros que permiten hacer las predicciones no corresponden a la bacteria modelizada.

Por este motivo, en tiempo reciente ambos laboratorios hemos unido nuestros esfuerzos con el objetivo de emplear la estructura matemática del modelo de *M. genitalium* y adaptarla para construir un modelo de *M. pneumoniae* que integre todos los parámetros de los diferentes módulos, determinados experimentalmente.

Esperamos que este modelo pueda predecir la respuesta a diferentes alteraciones o perturbaciones, ya sean genéticas (eliminación o sobreexpresión de genes, introducción de genes nuevos) o ambientales (suministro de antibióticos, cambios en el medio de cultivo). La capacidad de predecir el comportamiento de un ser vivo en el ordenador supondría un avance en el campo de la biología sintética. En el caso concreto de *M. pneumoniae*, permitiría su utilización como vehículo inteligente para el tratamiento de enfermedades. La posibilidad en un futuro de modelizar seres vivos más complejos abriría nuevos horizontes en las aplicaciones basadas en microorganismos, como la fabricación de biocombustibles, la química limpia o la biorremediación.

**María Lluch Senar** es investigadora del Grupo de Diseño de Sistemas Biológicos en el Centro de Regulación Genómica de Barcelona.

## Las simulaciones que generasen resultados ayudarían a priorizar las investigaciones futuras y favorecerían los descubrimientos

tipos de bacterias muy diferentes (entre ellas *Escherichia coli*) para obtener ciertos datos, como el tiempo promedio que los transcritos de ARN se mueven por la célula antes de que las enzimas los fragmenten para reciclar sus componentes. Sin un modo de acotar y comprobar esas estimaciones, no teníamos ninguna esperanza de éxito.

En ese momento de inspiración en 2008, me di cuenta de que la modelización de un organismo unicelular (y no de un grupo de células, como en casi todos los estudios previos) podría aportarnos esas cotas que necesitábamos. Consideremos el crecimiento y la reproducción. Una gran población de células crece paulatinamente. El nacimiento y la muerte de una sola de ellas no afecta mucho al conjunto. Pero para un organismo

unicelular, la división representa un suceso extraordinario. Antes de que se divida en dos, debe doblar su masa, y no solo la total. La cantidad de ADN, la membrana celular y cada tipo de proteína necesaria para la supervivencia deben duplicarse. Si el modelo se limita a simular una sola célula, el ordenador puede contar y seguir cada molécula durante todo su ciclo de vida. Y cuando la célula se convierte en dos, puede comprobar si todos los números encajan.

Además, un organismo unicelular se reproduce a un ritmo más o menos establecido. De este modo, *M. genitalium* se divide cada nueve o diez horas en el laboratorio. Rara vez tarda menos de seis o más de quince. El hecho de que tenga que duplicar todo su contenido en este intervalo de tiempo nos permitiría elegir posibles intervalos de numerosas variables que de otra manera quedarían sin determinar, como las que controlan el momento en el que comienza la replicación del ADN.

Reuní un equipo de físicos, biólogos, modelizadores, e incluso un antiguo ingeniero informático de Google, para debatir sobre las estrategias matemáticas que debíamos emplear. Michael Shuler, ingeniero biomédico de la Universidad de Cornell y pionero en simulación celular, había construido modelos impresionantes a partir de ecuaciones diferenciales ordinarias. Bernhard Palsson, que dirigió mis estudios en San Diego, había desarrollado una potente técnica, el análisis de balance de flujos, que resultaba adecuada para modelizar el metabolismo. Pero otros investigadores habían demostrado que el azar desempeña un papel importante en la transcripción génica. Además, la división celular conlleva obviamente un cambio en la geometría de la membrana. Esos otros métodos no podrían tener en cuenta tales aspectos. Ya como doctorando me di cuenta de que una sola técnica no puede modelizar todas las funciones de una célula; de hecho, mi tesis describe un método para acoplar dos modelos matemáticos diferentes en un solo simulador.

Decidimos, por tanto, crear un modelo de una célula a partir de 28 módulos, cada uno de los cuales empleaba el algoritmo que mejor se adaptaba al proceso biológico y al grado de conocimiento que poseíamos de él. Pero con esta estrategia obtuvimos una amalgama de procedimientos matemáticos. Necesitábamos unirlos en un conjunto coherente.

Pensé entonces en un curso de diseño de industrias químicas que seguí cuando estudiaba la carrera. Para el proyecto final del curso usamos un potente simulador, denominado HYSYS, para hacer un boceto de una gran refinería. HYSYS permitía diseñar un sistema en el que cada reacción principal ocurría en depósitos separados. Mediante tuberías se conectaban las salidas de un depósito con las entradas de otros. Esta estructura combinaba numerosos tipos de reacciones químicas en un sistema ordenado y predecible.

Se me ocurrió que ese enfoque, con alguna modificación, podría valer para nuestro simulador celular si se asumía una importante simplificación: que aunque todos esos procesos biológicos se produzcan a la vez en una célula viva, sus acciones son independientes, a efectos prácticos, en periodos de menos de un segundo. Partiendo de esta suposición, podríamos dividir la vida de una célula en intervalos de un segundo y ejecutar en ese tiempo cada uno de los 28 módulos, en orden, antes de actualizar el conjunto de las variables celulares. El modelo reflejaría toda la interconectividad de la bioquímica (como la dependencia de la transcripción génica y la síntesis de ADN con la energía y los nucleótidos producidos por el metabolismo), pero solo en escalas temporales superiores a un segundo.