

José Dávila-Velderrain* y Elena Álvarez-Buylla Rocas**

Esquemas de causación lineales en biología postgenómica: la subliminal y conveniente suposición del mapeo uno a uno entre genotipo y fenotipo

Resumen | En este ensayo cuestionamos la validez de algunos supuestos básicos en biología molecular y evolución sobre la base de datos experimentales recientes y a través de los lentes de las perspectivas no lineal y de sistemas. Enfocamos nuestra discusión sobre dos fundamentos bien establecidos de la biología: el flujo de información en biología molecular (es decir, su dogma central), y el paradigma de la vía de señalización “causal” lineal. Según ambos paradigmas, la suposición subliminal de un mapeo genotipo-fenotipo (one-to-one genotype-phenotype mapping GPM) uno a uno constituye una hipótesis de trabajo subyacente en muchos casos. Preguntamos si esto es empíricamente sostenible en el marco de la biología postgenómica. Llegamos a la conclusión de que, cuando se asume el concepto de redes complejas y procesos dinámicos en el comportamiento celular —un punto de vista que ahora ha sido validado empíricamente— ya no puede sostenerse el mapeo uno a uno. Ofrecemos la hipótesis que semejante suposición subliminal y a veces explícita debe sostenerse, hasta cierto punto, porque es conveniente para la apropiación privada y el mercadeo de los descubrimientos científicos. Sería de esperar que nuestra discusión contribuya a allanar la transición en curso hacia una biología de sistemas más integradora, explicativa, cuantitativa y multidisciplinaria. Esta última, a su vez, rendirá probablemente desarrollos médicos y agrícolas más preventivos y sustentables, respectivamente, que un enfoque reduccionista.

Linear Causation Schemes in Post-genomic Biology: The Subliminal and Convenient One-to-one Genotype-Phenotype Mapping Assumption

Abstract | In this essay we question the validity of basic assumptions in molecular biology and evolution on the basis of recent experimental data and through the lenses of a systems and nonlinear perspective. We focus our discussion on two well established foundations of biology: the flow of information in molecular biology (i.e. the central dogma of molecular biology) and the “causal” linear signaling pathway paradigm. Under both paradigms the

* Instituto de Ecología-Universidad Nacional Autónoma de México. Correo electrónico: jdjosedavila@gmail.com

** Centro de Ciencias de la Complejidad-Universidad Nacional Autónoma de México. Correo electrónico: eabuylla@gmail.com

subliminal assumption of a one-to-one genotype-phenotype mapping (GPM) constitutes an underlying working hypothesis in many cases. We ask if this is empirically sustainable in post-genomic biology. We conclude that, when embracing the notion of complex networks and dynamic processes governing cellular behavior—a view now empirically validated—one-to-one mapping can no longer be sustained. We hypothesize that such subliminal and sometimes explicit assumptions may be upheld, to a certain degree, because they are convenient for the private appropriation of scientific discoveries. Hopefully, our discussion will help smooth the undergoing transition towards a more integrative, explanatory, quantitative and multidisciplinary systems biology. The latter will likely also yield more preventive and sustainable medical and agricultural developments, respectively, than a reductionist approach.

Palabras clave | biología postgenómica – mapeo genotipo-fenotipo – determinismo genético – flujo de información genética

Keywords | post-genomic biology – genotype-phenotype mapping – genetic determinism – flow of genetic information

Introducción

LA PRÁCTICA DE LA CIENCIA surge, en su mayor parte, de los consensos. El progreso científico, sin embargo, también está sostenido por el desafío continuo a las ideas aceptadas. De tanto en tanto se producen rupturas de acuerdos tácitos, y entonces —dicen algunos— ocurre una transición, un llamado cambio de paradigma (Kuhn 2012). En las últimas décadas, varios autores han venido discutiendo la posibilidad de un cambio de paradigma en biología, dada la aparente crisis de algunos de sus principios fundacionales (Wilkins 1996; Strohmman 1997; O'Malley y Boucher 2005). En este trabajo, en cambio, nos interesa transmitir que una gran proporción de la investigación biológica aceptada asume subliminalmente ciertas suposiciones que, en esta época postgenómica, son empíricamente insostenibles. Algunas de estas suposiciones están tan profundamente arraigadas que todavía impregnan el diseño, la interpretación y la descripción de una amplia gama de investigaciones biológicas a nivel molecular aunque, si se le enfrentara directamente con el tema, cualquiera les restaría importancia. De manera rutinaria buscamos mutaciones simples “causales”, responsables de fenotipos complejos y suponemos que, encontrando la base molecular de una mutación correlacionada con una condición particular, ya hemos explicado el surgimiento de esta última. Lo importante de esto es que esta lógica da por hecho que en la mayoría de los casos una relación de uno-a-uno será posible. Por extensión de estas suposiciones definimos las vías de señalización como entidades

autónomas que ordenan a la célula cómo debe comportarse bajo una determinada condición. Si surge una conducta patológica, buscamos la fuente de unas instrucciones incorrectas: el componente o la vía mutada. Automáticamente interpretamos cualquiera manifestación de un rasgo *aprendido*, por ejemplo, la resistencia a un fármaco, como la consecuencia de los principios de optimización de la adaptación (darwiniana) por medio de mutación y selección “aleatorias”. Este sesgo recurrente hacia las explicaciones *ad hoc* ¿se basa solamente en la plausibilidad dada por la evidencia, o es una simple consecuencia de una tradición ingenuamente heredada? Consideramos que la presentación explícita de algunas de las suposiciones mencionadas, a la luz de información empírica postgenómica, y a través del lente de una perspectiva no lineal, de la biología de sistemas, podrá contribuir a aclarar esta cuestión. Esto puede resultar de utilidad para los estudiantes de biología actuales y para aquellos científicos interesados en la investigación multidisciplinaria.

Un primer desvío necesario: ¿qué queremos decir con biología postgenómica? La disponibilidad de secuencias genómicas completas (y también transcriptomas, proteomas, metabolomas, etc.) impactó obviamente la investigación en biología, posibilitando nuevos niveles de interrogación, además de desenmascarar nuevas fuentes de apoyo (o rechazo) empírico para otros hechos hasta entonces supuestos. En nuestro caso, sin embargo, además del acceso a datos provenientes de toda la amplitud del genoma, también incluimos en el término *postgenómico* varios rasgos que caracterizan a la biología moderna: 1) abundancia de datos moleculares experimentales, 2) acceso a maneras sistemáticas para caracterizar estados fenotípicos celulares; y 3) una tendencia a generar datos cuantitativos y a formular modelos matemáticos/computacionales. En consecuencia, desde nuestro punto de vista la *biología postgenómica* es necesariamente multidisciplinaria, integradora, formal y cuantitativa.

La suposición más básica e ingenua: el GPM uno a uno

En la actualidad es común pensar en la relación entre genotipos y fenotipos en términos de algún tipo de mapeo complejo (Kauffman 1993; Mendoza y Álvarez-Buylla 1998; Wagner y Zhang 2011; 2014; Ho y Zhang 2014). Se puede seguir el rastro del concepto de un “mapa genotipo-fenotipo” hasta Alberch, quien propuso elegantemente un modelo basado en los principios de la dinámica de sistemas para expresar las limitaciones de lo que algunos llaman *determinismo genético*, es decir la suposición de que los genes determinan directamente los fenotipos (Alberch 1991). Sería igualmente limitado adoptar un *determinismo epigenético*. Cabe destacar que una suposición tan centrada en el gen como determinante absoluto constituye la base de las frecuentemente citadas metáforas

de “plano genético” o “programa genético” (Pigliucci 2010). Más aun, también implica una relación lineal entre genotipos y fenotipos; en otras palabras, un mapeo uno a uno. Este modelo simplista es atractivo, ya que abarca naturalmente una interpretación de causa y efecto, lo que lo hace intuitivamente interesante. Pero si reflexionamos sobre esta suposición de que *un genotipo genera específicamente un fenotipo particular*, tenemos que enfrentar el problema de cómo un punto de vista tan simplista puede cuadrar con cualquier observación. No obstante, este modelo uno a uno sigue siendo la base de la mayoría de los programas aceptados de desarrollo biomédico o biotecnológico (como, por ejemplo, los cultivos transgénicos).

También incluimos en el término postgenómico varios rasgos que caracterizan a la biología moderna: 1) abundancia de datos moleculares experimentales, 2) acceso a maneras sistemáticas para caracterizar estados fenotípicos celulares, y 3) una tendencia a generar datos cuantitativos y a formular modelos matemáticos/computacionales. En consecuencia, desde nuestro punto de vista la biología postgenómica es necesariamente multidisciplinaria, integradora, formal y cuantitativa

Un segundo desvío necesario: ¿cuáles genotipo y fenotipo? En la epistemología de la evolución y la biología, en general, es común hablar de genotipo y fenotipo como si fueran términos absolutos. Pero dichos términos pueden ser definidos en diferentes niveles y, en la práctica, las distinciones entre genotipo y fenotipo son apenas parciales y dinámicas (Lewontin 2011). En la biología postgenómica esta distinción se apoya comúnmente en el uso de modelos sencillos de GPM (véase, por ejemplo, Soyer 2012). En consecuencia, hay más de un tipo de genotipos y fenotipos. Un modelo GPM puede ser especificado de diferentes maneras. Para los propósitos de este ensayo establecemos que el genotipo estará representado por una red reguladora de genes (GRN gene regulatory network), y los fenotipos por un perfil o configuración de expresión de genes.

No obstante, vale la pena notar que en la época actual de secuenciación de la próxima generación (NGS next-generation-sequencing) y la biología de la célula individual, la caracterización empírica de genotipos completos de células individuales se está volviendo factible. Por desgracia, tanto por razones conceptuales como técnicas, no puede decirse lo mismo de los fenotipos, aunque se estén desarrollando algunas estrategias fenotípicas sistemáticas específicas (véase, por ejemplo, Houle *et al.* 2010; Hancock 2014).

El mapeo genotipo-fenotipo uno a uno y el dogma central

Crick declaró por primera vez el “dogma central de la biología molecular” en 1958, que luego fue reiterado una vez más en 1970 (Crick 1958, 1970). En términos sencillos el *dogma* asevera que la información fluye al interior de las células del ADN al ARN y de ahí a las proteínas; como resultado se determina el fenotipo celular (Shapiro 2009).

Las simplificaciones involucradas en este modelo ya han sido cuestionadas desde un punto de vista de la información, concluyendo que descubrimientos realizados en las últimas décadas han hecho que el dogma sea insostenible (Shapiro 2009). En lugar de esto, nos enfocaremos sobre el papel consolidado del dogma en relación con lo que se refiere al esquema de causación de fenotipos moleculares, lineal y unidireccional, que se ha adoptado de manera implícita. Según una interpretación explícita del dogma, un gen codifica para una proteína, que de alguna manera determina un rasgo observable (es decir, un fenotipo). Esta visión simplista puede enmarcarse de manera efectiva en un modelo de GPM uno a uno (véase figura 1a). ¿Cómo definimos un fenotipo? Aquí se asigna un fenotipo a una molécula, una proteína, porque se dice que tiene una *función*. Esta función debería ser entonces una característica observable de la célula (organismo). Por lo tanto, el primer GPM uno a uno por discutir sería: un gen (es decir, el genotipo) codifica para una proteína que lleva a cabo una función específica que determina una característica observable (es decir, el fenotipo).

¿Es empíricamente sostenible en biología postgenómica este modelo uno a uno (gen-función)?

La primera dificultad que se nos ocurre es de naturaleza conceptual. ¿Qué significa *función*? Definir una función en biología no es un asunto menor (Huang 2000; Huneman 2013; Brunet y Doolittle 2014; Doolittle *et al.* 2014). En primer lugar, la asignación de funciones se le puede encargar a entidades a niveles diversos de la organización molecular, tales como gen, proteína, dominio de proteína, complejo de proteínas o vía (Huang 2000). En años recientes, investigadores en las áreas de genómica y epigenómica hasta abogan por el mapeo de funciones a nivel de genoma y con resolución de un solo nucleótido (Kellis *et al.* 2013). Para ser concretos, vamos a enfocarnos en la función a nivel de proteína. Aunque lo que definimos como función de la proteína suele ser condicionado por el contexto —es decir, el entorno celular— (Huang 2000), para propósitos de nuestra discusión vamos a asumir que una función de proteína puede ser asignada invariablemente. Así, en el modelo simple uno a uno, un gen está vinculado invariablemente a una función específica a través de la acción de una proteína.

Según la más reciente versión del ensamblaje del genoma humano que

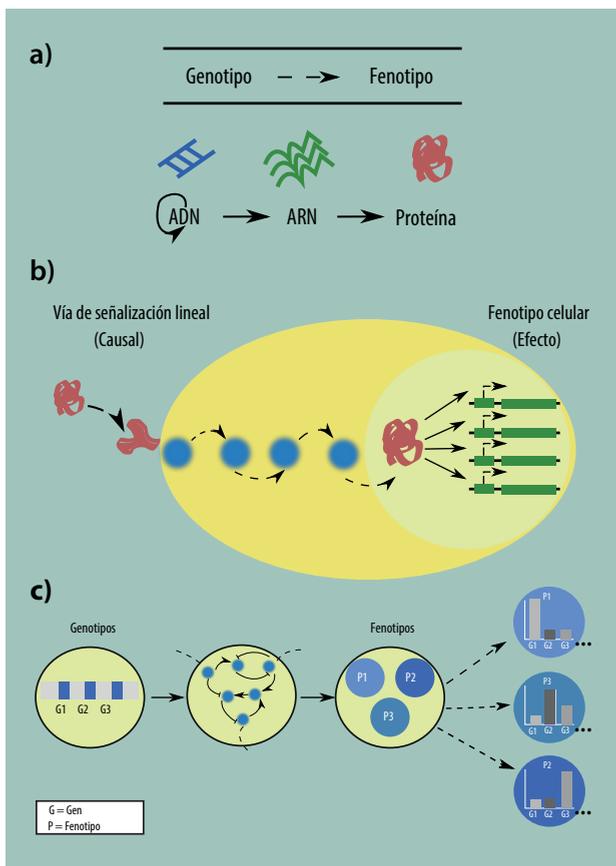


Figura 1. Representación esquemática de los GPMs expuestos en el texto principal. a) GPM uno a uno que representa el dogma central de biología molecular: un gen (p. ej., el genotipo) codifica para una proteína, la cual lleva a cabo una función específica que determina una característica observable (p. ej., el fenotipo). b) GPM uno a uno que representa el paradigma de la vía de señalización “causal” lineal: genes codifican para las proteínas involucradas en una vía de señalización (genotipo), y esta última convierte una señal molecular específica en un fenotipo celular específico. c) GPM no lineal que representa la especificación de los fenotipos celulares mediante la dinámica de GRN: los genes de un genoma individual (genotipo) interactúan en GRNs complejas cuyas interacciones regulatorias determinan los fenotipos celulares observables.

aparece en la base de datos Ensembl (<http://www.ensembl.org/>) el ser humano tiene 20,389 genes codificadores, 9,656 pequeños genes no codificadores, y 14,470 genes largos no codificadores. La primera observación obvia es que no todos los genes codifican para proteínas. Hay dos hechos postgenómicos: 1) la

mayor parte del genoma humano no codifica para proteínas (Alexander *et al.* 2010), y 2) la transcripción ocurre con mucha más frecuencia de la esperada (Carninci *et al.* 2005; Cheng *et al.* 2005). ¿Será posible que los genes que no codifican para proteínas también definen el fenotipo? Bueno, probablemente, de alguna manera; pero con toda seguridad no lo hacen por medio de un GPM uno a uno, dado el punto de vista que va emergiendo en el sentido de que la transcripción no codificadora está estrechamente ligada a la regulación génica y a la especificación del tipo celular (Natoli y Andrau 2012). Por ejemplo, se mostró recientemente que el ARN transcrito de potenciadores llamado eRNA, es capaz de regular la transcripción (Plosky 2014). Como veremos más abajo, la regulación génica en sí misma es el mecanismo central detrás de la definición de las redes regulatorias génicas; también es fundamental para entender la conducta colectiva de las redes. Conceptualizar la conducta celular en términos de redes moleculares, a su vez, representa una desviación total del GPM uno a uno.

Además de genes (no) codificadores, la cantidad de proteínas codificadas en el genoma humano, y representadas por modificaciones de transcripción ha sido estimada entre 50,000 y 500,000 (Uhlen y Ponten 2005). Considerando la ahora conocida cifra de genes y (estimada) de proteínas en otros organismos, varios autores han señalado que la complejidad genómica (y proteómica) no está correlacionada con la complejidad fenotípica (véase por ejemplo Huang 2002). Este hecho empírico, nuevamente, no es consistente con lo que esperaríamos de la extensión del dogma.

Más allá de la curiosidad despertada por datos genómicos generados recientemente, una desventaja más sería del GPM uno a uno asociado con el *dogma central* reside en que pasa completamente por alto las interacciones entre genes (Tyler *et al.* 2009). La *epistasis* se refiere al fenómeno en el cual el efecto funcional de un gen está condicionado por otros genes (Phillips 2008), mientras la *pleiotropía* se refiere a una función que es afectada por muchos genes (Stearns

Según la más reciente versión del ensamblaje del genoma humano que aparece en la base de datos Ensembl, el ser humano tiene 20,389 genes codificadores, 9,656 pequeños genes no codificadores, y 14,470 genes largos no codificadores. La primera observación obvia es que no todos los genes codifican para proteínas. Hay dos hechos postgenómicos: 1) la mayor parte del genoma humano no codifica para proteínas, y 2) la transcripción ocurre con mucha más frecuencia de la esperada. ¿Será posible que los genes que no codifican para proteínas también definen el fenotipo?

2010); estos dos fenómenos constituyen hechos (y conceptos) bien establecidos en genética moderna y clásica (Lehner 2011; Wagner y Zhang 2011). En la actualidad esta clase de interacciones genéticas están siendo estudiadas sistemáticamente a escala genómica. Por ejemplo, ahora es posible probar millones de combinaciones diferentes de mutantes dobles y evaluar sus efectos sobre una función cuantificable, tal como hicieron Costanzo y sus colegas usando la levadura de cerveza *Saccharomyces cerevisiae* (Costanzo *et al.* 2010). Estudios como éste han mostrado con claridad que el efecto de un gen sobre un fenotipo específico depende de la actividad (o ausencia de la misma) de muchos genes más. En este sentido una interacción *génica* se define sobre la base de este efecto funcional condicionado. Si bien una discusión a fondo de la epistasis y la pleiotropía se halla fuera de los alcances de este trabajo, vale la pena hacer notar que tales mecanismos están estrechamente vinculados con dos tipos innegables de evidencia experimental: 1) se pueden generar resultados muy diferentes de juegos de genes casi idénticos, o el mismo genotipo puede producir fenotipos contrastantes, y 2) es posible arribar a puntos fenotípicos virtualmente idénticos utilizando genotipos muy diferentes. Evidentemente, estos hechos no cuadran con un GPM uno a uno. Aunque son aparentemente paradójicos, ambas declaraciones pueden reconciliarse perfectamente usando un modelo de GPM muchos a muchos en el cual las interacciones entre componentes genéticos y no genéticos son considerados de manera explícita; un punto de vista mucho más consistente con la manera en que los sistemas vivientes y adaptables se comportan y evolucionan.

El mapeo uno a uno y las vías de señalización

Extendiendo el concepto de uno a uno hasta un nivel más alto, los biólogos moleculares lo aplican para asociar una vía de señalización alterada con una particular condición fenotípica. Las señales extracelulares son transmitidas por proteínas intermediarias a proteínas efectoras, que eventualmente activan los conjuntos de genes responsables del establecimiento de fenotipos “*apropiados*”. Nótese que el término vía por sí solo hace referencia a un grupo de eventos que se producen ordenadamente a lo largo de una *línea*. Por lo tanto, en un sentido, este modelo multimolecular continúa la idea *dogmática* de transferencia de información lineal y unidireccional. Por lo tanto, desde nuestro punto de vista, también constituye efectivamente un modelo de GPM uno a uno (véase figura 1b). Los genes codifican para las proteínas involucradas en la vía (genotipo), y éstas mapean hacia una señal molecular específica (instrucción) que conduce a un fenotipo celular específico. La propiedad lineal de las vías de señalización también implica causa y efecto unidireccional: se cree que una señal de

instrucción dada causa directamente una manifestación fenotípica. Tradicionalmente los biólogos han adoptado esta imagen simple de la vía como explicación válida a nivel molecular para muchos fenotipos celulares. Ni siquiera es válido un enfoque de uno a uno para asociar una red con un fenotipo.

¿Es sostenible empíricamente este modelo uno a uno (señal a fenotipo) en la biología postgenómica?

Aquí podemos plantear preguntas similares a las que se enuncian más arriba. Por ejemplo, ¿hay suficientes vías de señalización celular para la cantidad de posibles señales extracelulares? ¿Hay una relación directa, de uno a uno, entre señales y fenotipos? Y si esto es así, ¿por qué los fenotipos celulares (es decir, los tipos de célula) parecen ser entes discretos mientras, por ejemplo, las señales portadas por los factores solubles de crecimiento exhiben concentraciones sujetas a variaciones continuas? Y, más importante, ¿cómo y por qué los fenotipos celulares se mantienen después de que ha cesado la señal? Como explicaremos más adelante, reexaminar el comportamiento de las células como resultado de las restricciones impuestas por las interacciones regulatorias de las redes moleculares complejas es útil para abordar estas cuestiones.

La explosión genómica ha llevado a la caracterización por fuerza bruta (de procesamiento de datos) de los componentes moleculares y sus interacciones, que ahora están siendo integrados en grandes bases de datos (Chattrayamontri *et al.* 2013). Como se esperaba, se han hecho esfuerzos para tratar de clasificar tales componentes en colecciones de vías de señalización a nivel genómico en múltiples organismos (Schaeffer *et al.* 2009; Croft *et al.* 2010). ¿Qué se aprendió? ¿Facilita la caracterización exhaustiva de vías de señalización la comprensión de fenotipos celulares y su plasticidad? En analogía con el fracaso de la predicción de que la caracterización de todos los genes de un organismo permitiría entender las reglas codificadas por el genoma que rigen su comportamiento, el enlistado de vías de los componentes moleculares y sus interacciones sólo ha revelado un cuadro que es mucho más complejo que lo anticipado. Pero falta mucho para explicar las manifestaciones fenotípicas por medio de cadenas lineales de causación molecular (Huang 2011) o, en otras palabras, de asociaciones lineales con preferencia sobre modelos explicativos.

Décadas de experimentación han mostrado que hay extenso intercambio de información entre las vías de señalización caracterizadas individualmente. De manera correspondiente, los fenómenos de epistasis y pleiotropía explicados en párrafos anteriores se extienden naturalmente en el nivel de las vías. Mientras varias vías diferentes pueden converger hacia fenotipos específicos, una vía y señal molecular específicas pueden también generar fenotipos diferentes

dependiendo de cada contexto (Huang 2000). Estas observaciones sugieren interacciones más allá de las cascadas lineales. Por otra parte, se puede producir un efecto similar al “causado” por una señal molecular específica mediante estímulos no específicos o incluso de manera independiente de cualquier estímulo. Por ejemplo, estímulos mecánicos como los inducidos por alteraciones de la forma celular pueden inducir fenotipos celulares sin ninguna instrucción extra celular específica ni cambio genético (Huang 2000). Por el otro lado, dada la estocasticidad intrínseca de las reacciones bioquímicas tanto extra como intracelulares, las células de un tipo específico a linaje pueden asumir diferentes fenotipos heredables, ya sea en ausencia de una diferencia genética o ambiental asociada, o mediante el procesamiento de señales ambientales estocásticas no específicas (Perkins y Swain 2009; Balazsi *et al.* 2011). Estos hechos también hacen que una explicación mecanicista por medio de un GPM uno a uno a nivel de vías de señalización sea insostenible. La inevitable plasticidad del comportamiento celular y la robustez de las manifestaciones fenotípicas observadas exigen un modelo explicativo alternativo. Razonamos más adelante que la perspectiva formal de la conducta celular como propiedad emergente de las restricciones impuestas por las redes regulatorias de los genes brinda un punto de vista alternativo acerca de cómo los genotipos mapean para generar fenotipos, ofreciendo un punto de partida para abordar procesos que, de otra manera, son sumamente complejos.

Más allá del GPM uno a uno: perspectiva de la dinámica de redes

¿Cómo quedan parados los dos puntos de vista expuestos anteriormente (el mapeo uno a uno de gen y vía de señalización a función) en la biología de sistemas postgenómica? Los genes, las proteínas codificadas y las vías de señalización molecular están incrustados en redes complejas de componentes genéticos y no genéticos que generalmente presentan varias asas de retroalimentación positiva y negativa y un comportamiento dinámico. Nos enfocamos ahora en la regulación génica, que es la base para conceptualizar las interacciones génicas, la propiedad fundamental que subyace en las redes regulatorias génicas no lineales. El concepto mismo de la regulación de los genes, que no es nada nuevo, no es consistente con un GPM uno a uno porque implica que el efecto fenotípico de una función génica dependerá de la actividad de otros genes que la regulan. Aunque la conciencia explícita del hecho de que los genes que codifican para todas las proteínas de la célula son necesariamente regulados por algunas otras proteínas regulatorias, que a su vez son reguladas, parece abrumadora, tal conciencia puede ser representada de manera sucinta en modelos cualitativos de redes de regulación génica (GRN en inglés). Estos modelos se están volviendo

muy útiles para seguir y comprender la acción concertada de múltiples componentes interactuantes.

Un modelo de trabajo que se usa comúnmente en biología de sistemas es aquel en que el genoma se mapea directamente a una GRN, y el fenotipo celular se representa por la actividad de cada uno de sus genes, su patrón de expresión. De esta manera, en una distinción genotipo-fenotipo basada en la dinámica de la GRN, una red representa efectivamente el genotipo de la célula, mientras su perfil asociado de expresión representa su fenotipo (Davila-Velderrain y Álvarez-Buylla 2014). La estructura del genoma (y la red) permanece virtualmente constante a través del desarrollo mientras el fenotipo celular va cambiando. ¿Por qué se observan cambios fenotípicos a lo largo del desarrollo bajo patrones tan robustos y reproducibles?

La naturaleza genómica de la GRN implica una estructura codificada físicamente, por medio de la cual la red restringe naturalmente el comportamiento temporal permisible de la actividad de cada gen. Por ejemplo, un gen específico "A" es regulado por un juego específico de genes. Dado el estado de la actividad de estos reguladores y la forma funcional de la regulación, dinámicamente el gen "A" será canalizado hacia estados futuros específicos. Esta simple regla regulatoria se aplica simultáneamente a todos los genes, generando un proceso auto organizativo que lleva inevitablemente al establecimiento de solamente aquellos estados celulares (fenotipos) que son consistentes con la lógica regulatoria subyacente. Por lo tanto, la GRN impone restricciones al comportamiento de la célula. La robustez y reproducibilidad observadas del comportamiento celular surgen naturalmente como un proceso auto organizado. Cualquier fuente de estímulo inductivo extracelular (no) específico convergiría inevitablemente hacia uno de los estados fenotípicos que son lógicamente consistentes con la función regulatoria subyacente de la red que se encuentra bajo consideración.

La exposición razonada expuesta brevemente en párrafos anteriores ha sido explotada para proponer GRNs fundadas en datos experimentales para entender cómo ocurre la especificación del destino celular durante, por ejemplo, el desarrollo temprano de la flor (véase Mendoza y Álvarez-Buylla 1998; Espinosa-Soto *et al.* 2004; y actualización en Sánchez-Corrales *et al.* 2010), y el establecimiento de patrones de células madre (Azpeitia *et al.* 2010); en la actualidad se encuentra apoyada por una gran cantidad de trabajo teórico y experimental consolidado (véase por ejemplo, Huang *et al.* 2005; Azpeitia *et al.* 2014).

En contraste con las suposiciones implícitas en el GPM uno a uno, las interacciones al interior de la red son fundamentales para el establecimiento del fenotipo y, por lo tanto, el efecto de una mutación sobre el fenotipo manifestado estará condicionado por el contexto de la red del gen bajo consideración (Davila-Velderrain *et al.* 2014). Dado que la multitud de estados celulares robustos

observados dependería de las restricciones de la red debidas a interacciones regulatorias, el papel orquestador de las GRNs constituye efectivamente un GPM (no lineal) de muchos a muchos en el cual la mayoría de los componentes pueden, al mismo tiempo, constituir tanto causas como efectos (véase figura 1c).

¿Investigación biomédica y biotecnológica ciega, indiferente u orientada hacia el mercado?

A pesar de toda la evidencia generada por casi dos décadas de investigación postgenómica, la presencia subliminal del GPM uno a uno, demasiado simplificado —aunque la mayor parte del tiempo no se le reconozca— no puede ser negada. Se supone implícitamente como una meta principal que impulsa la investigación biomédica convencional que los genes causan, por ejemplo, cáncer, porque generan fenotipos mediante la codificación de proteínas (Huang 2013). Éste también es el caso para la investigación biotecnológica, en la que se reconoce que un gen particular de una especie en que se produce una “función” particular puede introducirse en otra especie con la esperanza de producir la misma “función” (Vaeck *et al.* 1987). Tomando en cuenta que hay una gran cantidad de estudios buscando mutaciones “causales”, esta suposición centrada en el gen sólo se nota en unos pocos casos o, alternativamente, se pasa por alto. A pesar de la enorme cantidad de recursos invertidos en proyectos de secuenciación de genomas, no se ha podido identificar exitosamente algo como una mutación (causal) universal para una enfermedad degenerativa (Huang 2013). No obstante, el contar con moléculas específicas como candidatas a factores causales de enfermedades particulares permite a las empresas desarrollar nuevos fármacos destinados al mercado. Dada la naturaleza limitada del simplista GPM uno a uno subyacente, es probable que este enfoque fracase. Se puede reproducir exclusivamente sobre la base de su limitada efectividad —y, en gran parte, merced a estrategias de mercadeo— en lugar de explicaciones profundas o soluciones agudamente necesarias. Hay que hacer notar que esta clase de investigación sostenida en la búsqueda de blancos moleculares potenciales en terapéutica, o soluciones tipo “bala mágica” de un solo gen para amenazas agrícolas complejas, pone en evidencia la prevalencia del GPM uno a uno, es decir, asume que hay una proteína para cada enfermedad o para cualquiera amenaza ambiental en agricultura.

El potencial para la terapia también complica las cosas, porque podría ser una meta de investigación perfectamente aceptable, sin importar su impacto sobre la mejora de la comprensión o si prueba efectivamente causación. Así, podría ser el caso que la investigación biomédica no ha evolucionado naturalmente hacia una situación tan ingenua; podría ser, en lugar de eso, que el carácter

tecnocrático impulsado por el mercado de la “ciencia” moderna se ha dedicado a estimular la herencia de ideas viejas que siguen siendo convenientes; desgraciadamente para la ciencia, sin embargo, la tasa de crecimiento de la comprensión conceptual no parece poder seguirle el paso a la rapidez de la evolución tecnológica.

Resumiendo, el paradigma dominante supone implícitamente que los genes determinan la conducta celular por medio de GPM uno a uno. Específicamente, los genes codifican para proteínas que determinan directamente los fenotipos y, en consecuencia, mutaciones en los genes deberían por sí solas alterar los fenotipos. Por lo tanto, tomar como blancos proteínas alteradas producidas a partir de genes mutados parece ser la mejor estrategia para “corregir” un fenotipo patológico; lo mismo podría decirse de las alteraciones epigenéticas, de las vías de señalización alteradas o incluso de las redes. Sin embargo, un gran volumen de evidencia postgenómica hace que el GPM uno a uno sea insostenible. En contraste, un GPM que tome en cuenta el papel orquestador de las redes regulatorias moleculares, que constituye un GPM de muchos a muchos, explica naturalmente muchas observaciones paradójicas y brinda un marco formal para la interpretación de la masa creciente de datos moleculares postgenómicos. ■

Reconocimientos

Este trabajo recibió el apoyo de becas ERAB: Conacyt (México) 180098 y 180380; y UNAM-DGAPA-PAPIIT: IN 203113.

Referencias

- Alberch, P. «From genes to phenotype: dynamical systems and evolvability.» *Genetica* 84, n° 1 (1991): 5-11.
- Alexander, R. P., G. Fang, J. Rozowsky, M. Snyder y M. B. Gerstein. «Annotating non-coding regions of the genome.» *Nature Reviews Genetics* 11, n° 8 (2010): 559-571.
- Azpeitia, E., J. Davila-Velderrain, C. Villarreal y E. R. Álvarez-Buylla. «Gene regulatory network models for floral organ determination.» *Flower Development*, 2014: 441-469.
- ., M. Benítez, I. Vega, C. Villarreal y E. R. Álvarez-Buylla. «Single-cell and coupled GRN models of cell patterning in the *Arabidopsis thaliana* root stem cell niche.» *BMC systems biology* 4, n° 1 (2010): 134.
- Balázsi, G., Van Oudenaarden, A. y J. J. Collins. «Cellular decision making and biological noise: from microbes to mammals.» *Cell* 144, n° 6 (2011): 910-925.
- Brunet, T. D. y W. F. Doolittle. «Getting “function” right.» *Proceedings of the*

- National Academy of Sciences* 111, nº 33 (2014): E3365-E3365.
- Carninci, P. *et al.* «The transcriptional landscape of the mammalian genome.» *Science* 309, nº 5740 (2005): 1559-1563.
- Chatr-aryamontri, A. *et al.* «The BioGRID interaction database: 2013 update.» *Nucleic acids research* 41 (D1) (2013): D816-D823.
- Cheng, J. *et al.* «Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution.» *Science* 308, nº 5725 (2005): 1149-1154.
- Costanzo, M. *et al.* «The genetic landscape of a cell.» *Science* 327, nº 5964 (2010): 425-431.
- Crick, F. H. «Central dogma of molecular biology.» *Nature* 227, nº 5258 (1970): 561-563.
- . «On protein synthesis.» *Symposia of the Society for Experimental Biology* 12 (1958): 138.
- Croft, D. *et al.* «Reactome: a database of reactions, pathways and biological processes.» *Nucleic acids research*, 2010: gkq1018.
- Davila-Velderrain, José, A. Servin-Marquez y E. R. Álvarez-Buylla. «Molecular evolution constraints in the floral organ specification gene regulatory network module across 18 angiosperm genomes.» *Molecular biology and evolution* 31, nº 3 (2014): 560-573.
- y E. R. Álvarez-Buylla. «Bridging genotype and phenotype.» En *Frontiers in Ecology, Evolution and Complexity*, editado por Octavio Miramontes, Alfonso Valiente-Banuet y Mariana Benítez. CopIt ArXives, 2014.
- Doolittle, W. F., T. D. Brunet, S. Linquist y T. R. Gregory. «Distinguishing between “function” and “effect” in genome biology.» *Genome biology and evolution* 6, nº 5 (2014): 1234-1237.
- Espinosa-Soto, C., P. Padilla-Longoria y E. R. Álvarez-Buylla. «A gene regulatory network model for cell-fate determination during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles.» *The Plant Cell Online* 16, nº 1 (2004): 2923-2939.
- Hancock, J. M. (Ed.). *Phenomics*. CRC Press, 2014.
- Ho, W. C. y J. Zhang. «The Genotype-Phenotype Map of Yeast Complex Traits: Basic Parameters and the Role of Natural Selection.» *Molecular biology and evolution* 31, nº 6 (2014): 1568-1580.
- Houle, D., D. R. Govindaraju y S. Omholt. «Phenomics: the next challenge.» *Nature Reviews Genetics* 11, nº 12 (2010): 855-866.
- Huang, S., G. Eichler, Y. Bar-Yam y D. E. Ingber. «Cell fates as high-dimensional attractor states of a complex gene regulatory network.» *Physical Review Letters* 94, nº 12 (2005): 128701.
- . «Genetic and non-genetic instability in tumor progression: link between the fitness landscape and the epigenetic landscape of cancer cells.» *Cancer*

- and Metastasis Reviews* 32, nº 3-4 (2013): 423-448.
- . «Rational drug discovery: what can we learn from regulatory networks?» *Drug discovery today* 7, nº 20 (2002): s163-s169.
- . «Systems biology of stem cells: three useful perspectives to help overcome the paradigm of linear pathways. Philosophical Transactions of the Royal Society B.» *Biological Sciences* 366, nº 1575 (2011): 2247-2259.
- . «The practical problems of post-genomic biology.» *Nature biotechnology* 18, nº 5 (2000): 471-472.
- Huneman, P. *Functions: selection and mechanisms*. Springer, 2013.
- Kellis, M. *et al.* «Defining functional DNA elements in the human genome.» *Proceedings of the National Academy of Sciences* 111, nº 17 (2014): 6131-6138.
- Kuhn, T. S. *The structure of scientific revolutions*. University of Chicago Press, 2012.
- Lehner, B. «Molecular mechanisms of epistasis within and between genes.» *Trends in Genetics* 27, nº 8 (2011): 323-331.
- Lewontin, R. «The genotype/phenotype distinction.» En *Stanford Encyclopedia of Philosophy*. 2011.
- Mendoza, L. y E. R. Álvarez-Buylla. «Dynamics of the genetic regulatory network for *Arabidopsis thaliana* flower morphogenesis.» *Journal of Theoretical Biology* 193, nº 2 (1998): 307-319.
- Natoli, G. y J. C. Andrau. «Noncoding transcription at enhancers: general principles and functional models.» *Annual review of genetics* 46 (2012): 1-19.
- O'Malley, M. A. y Y. Boucher. «Paradigm change in evolutionary microbiology.» *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 2005: 183-208.
- Perkins, T. J. y P. S. Swain. «Strategies for cellular decision-making.» *Molecular systems biology* 5, nº 1 (2009).
- Phillips, P. C. «Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems.» *Nature Reviews Genetics* 9, nº 11 (2008): 855-867.
- Pigliucci, M. «Genotype-phenotype mapping and the end of the 'genes as blueprint' metaphor.» *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, nº 1540 (2010): 557-566.
- Plosky, Brian S. *eRNAs Lure NELF from Paused Polymerases*. *Molecular Cell*. 2014.
- Rose, M. R. y T. H. Oakley. «The new biology: beyond the Modern Synthesis.» *Biology direct* 2, nº 1 (2007): 30.
- Sanchez-Corrales, Y. E., E. R. Álvarez-Buylla y L. Mendoza. «The *Arabidopsis thaliana* flower organ specification gene regulatory network determines a robust differentiation process.» *Journal of Theoretical Biology* 264, nº 3 (2010): 971-983.

- Schaefer, C. F. *et al.* «PID: the pathway interaction database.» *Nucleic acids research*, 2009: D674-D679.
- Shapiro, J. A. «Revisiting the central dogma in the 21st century.» *Annals of the New York Academy of Sciences* 1178, nº 1 (2009): 6-28.
- Soyer, O. S. (Ed.). *Evolutionary systems biology* 751, Springer 2012.
- Stearns, F. W. «One hundred years of pleiotropy: a retrospective.» *Genetics* 186, nº 3 (2010): 767-773.
- Strohman, R. C. «The coming Kuhnian revolution in biology.» *Nature biotechnology* 15, nº 3 (1997): 194-200.
- Stuart A. K. *The origins of order: Self-organization and selection in evolution*. Oxford, Gran Bretaña: Oxford University Press, 1993.
- Tyler, A. L., F. W. Asselbergs, S. M. Williams y J. H. Moore. «Shadows of complexity: what biological networks reveal about epistasis and pleiotropy.» *Bioessays* 31, nº 2 (2009): 220-227.
- Uhlen, M. y F. Ponten. «Antibody-based proteomics for human tissue profiling.» *Molecular & Cellular Proteomics* 4, nº 4 (2005): 384-393.
- Vaeck, M., *et al.* «Transgenic plants protected from insect attack.» *Nature* 328 (1987): 33-37.
- Wagner, G. P. y J. Zhang. «The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms.» *Nature Reviews Genetics* 12, nº 3 (2011): 204-213.
- Wilkins, A. S. «Are there 'Kuhnian' revolutions in biology?» *BioEssays*, 1996: 695-696.