

## Las matemáticas de la evolución\*

por

José A. Cuesta

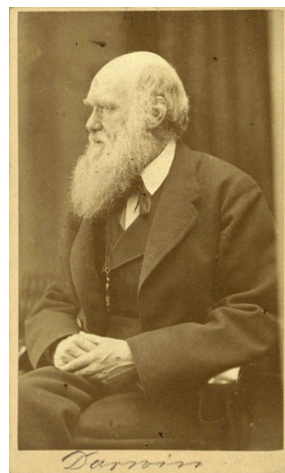
«Un matemático es un ciego en una habitación oscura buscando un gato negro que no está allí.»

—Charles Darwin

Seguramente resulta chocante encontrar un artículo sobre Darwin y la evolución en una revista de matemáticas. Tanto uno como la otra se asocian siempre, lógicamente, a la Biología. Sin embargo la teoría de la evolución, como toda teoría científica que merezca ese nombre, admite formulaciones cuantitativas en un buen número de aspectos. El propio Darwin, no muy ducho en matemáticas pero hombre culto como era, conocedor seguramente de la mecánica newtoniana, admitía que la única forma de llegar a comprender los descubrimientos científicos era a través de su formulación matemática. Darwin consideraba a los matemáticos personas con un sexto sentido que les permite «ver» en lugares del pensamiento donde el resto del mundo es «ciego». De ahí la cita que encabeza este artículo y que, pese a lo que pueda parecer, expresa la admiración que Darwin sentía hacia los matemáticos.

La teoría de la evolución por selección natural [3], la gran obra de Darwin, recibió una aportación fundamental a comienzos de siglo cuando se le agregó el conocimiento de las leyes de la herencia que descubrió Gregor Mendel en 1865 [14], tras su redescubrimiento por de Vries, Correns y von Tschermak en 1900. A partir de ese momento, y en una forma en que posteriormente se ha dado en denominar «genética de poblaciones», un puñado de matemáticos, entre los que destacan Fisher, Haldane, Wright y más tarde Kimura, sentaron las bases de la teoría matemática de la evolución. Hoy día esta teoría ha adquirido un estatus propio dentro de la Matemática Aplicada, y se ha desarrollado y diversificado permitiéndonos comprender los mecanismos tan sutiles con que la evolución opera, no solamente en la Biología, sino en muchas otras disciplinas que comparten principios semejantes, como pueden ser la Lingüística, la Economía, la Sociología o la Informática.

Este año, cuando se cumple el segundo centenario del nacimiento de Darwin, ciento cincuenta años después de la publicación de *El Origen de las Especies*, parece



\*Con este artículo el autor desea contribuir a la celebración durante 2009 del Año Darwin.

un momento adecuado para hacer una revisión de la teoría de la evolución desde un punto de vista matemático. Ése es el motivo de este artículo, y me apresuro a aclarar que las aportaciones de la Matemática a esta teoría son tantas y tan variadas que sólo puedo aspirar a mencionar una pequeña parte de ellas. El tono divulgativo del artículo tampoco permite profundizar en exceso en ellas, de manera que remito al lector interesado a investigar las referencias de las que me he ayudado para escribir este artículo [5, 7, 15]. Y antes de entrar en materia quisiera hacer una aclaración. Existen dos tipos de reproducción en los organismos vivos: la asexual, en la que un solo organismo genera copias de sí mismo, y la sexual, en la que se requiere la intervención de más de un organismo (casi siempre dos) en la reproducción. La primera es propia (aunque no exclusiva) de organismos simples, como los virus, y por ello es el objeto de análisis de la mayoría de los modelos. La segunda implica combinación de material genético procedente de al menos dos progenitores, y ello da lugar a complicaciones específicas que también son objeto de modelos más sofisticados. Este artículo se centrará casi exclusivamente en la primera (ya de por sí complicada de entender), aunque aludiré a la segunda en determinadas ocasiones.

## 1. MECANISMOS FUNDAMENTALES DE LA EVOLUCIÓN

Cuando uno reflexiona acerca de qué es necesario, a un nivel abstracto, para que exista un proceso evolutivo, no importa el contexto de que se trate, se da cuenta de que son tres los mecanismos fundamentales que tienen que concurrir: un mecanismo de *replicación* que permita a las entidades en juego crear copias de sí mismas; un mecanismo de *mutación* que dé lugar a la aparición de ligeras variantes en esas copias, y un mecanismo de *selección* que haga que algunas de esas copias, las «mejores» en un cierto sentido, desplacen a las demás generación tras generación. Vamos a analizar por separado estos tres mecanismos.

### 1.1. REPLICACIÓN

Una bacteria típica se divide cada 20 minutos dando lugar a dos copias de sí misma; 20 minutos más tarde tendremos cuatro; al cabo de una hora serán ocho. . . La población de bacterias en una generación,  $n_t$ , está relacionada con la de la generación anterior por la sencilla ecuación  $n_t = 2n_{t-1}$ , cuya solución, tomando  $n_0 = 1$ , es  $n_t = 2^t$ . Un proceso replicativo como el que acabamos de describir conduce a un crecimiento exponencial. Fue Malthus quien primero propuso esta ley de crecimiento en su *Ensayo sobre el principio de la población*. De acuerdo con ella, y en una situación en la que las generaciones se entremezclan, si una población  $n(t)$  crece con una tasa de replicación por individuo constante  $r$ , esto es,  $\dot{n} = rn$ , entonces  $n(t) = n(0)e^{rt}$ . Éste es el modelo malthusiano de crecimiento de la población humana y la versión continua de la ley de reproducción bacteriana que hemos obtenido.

Malthus supuso una de las grandes influencias en el pensamiento de Darwin debido a lo que se vino a denominar «catástrofe malthusiana», que se ilustra muy bien con el ejemplo de las bacterias. De acuerdo a la ley que hemos deducido, al cabo de tan solo dos días (144 divisiones) tendremos  $2^{144} \approx 2 \times 10^{43}$  bacterias. Para

hacernos una idea: si el diámetro de una bacteria es  $\sim 1 \mu\text{m}$  y su densidad es la del agua ( $1 \text{ g cm}^{-3}$ ), entonces  $2 \times 10^{43}$  bacterias tienen una masa de aproximadamente  $2 \times 10^{28} \text{ kg}$ , o sea, unas 3000 veces la masa de la tierra. ¡En tan solo dos días! Es evidente que un crecimiento así no es sostenible y que debe llevar a la muerte a la inmensa mayoría de los individuos. De hecho, la escasez de recursos se puede incluir como una modificación de la ley de Malthus haciendo que la tasa de replicación por individuo sea una función decreciente de la población, anulándose estrictamente cuando la población alcanza el nivel máximo sostenible. En la forma más sencilla esto da lugar a la ley *logística*

$$\dot{n} = r \left(1 - \frac{n}{K}\right) n, \tag{1}$$

cuya solución,

$$n(t) = \frac{K n_0 e^{rt}}{K + n_0 (e^{rt} - 1)}, \quad \lim_{t \rightarrow \infty} n(t) = K, \tag{2}$$

es una curva monótona creciente en el tiempo (cuando  $n_0 < K$ ) que satura al valor máximo  $K$ , conocido como «capacidad de carga» del entorno. Podemos considerar, por tanto, que las poblaciones con un crecimiento de este tipo se encuentran en un equilibrio con su entorno que mantiene la población constante.<sup>1</sup> En la práctica la población fluctuará en torno a ese valor debido a efectos estocásticos que las leyes que estamos considerando no contemplan.

### 1.2. SELECCIÓN

En una situación de escasez de recursos como la que acabamos de describir existe una lucha a muerte entre los individuos por conseguirlos. La lucha con los semejantes da lugar a la saturación de población que predice la ley logística, pero cuando hay individuos de distinto tipo (distintas especies), las diferencias desempeñan un papel clave en decidir quién sobrevive y quién muere. El parámetro evolutivo clave en este aspecto es el *grado de adaptación* (en inglés *fitness*), que se define como el número medio de individuos en edad reproductiva que cada individuo aporta a la siguiente generación. En el caso de que la población siga la ley de Malthus, el grado de adaptación se mide mediante el parámetro  $r$ . Si sigue la ley logística el grado de adaptación,  $f = r(1 - n/K)$ , dependerá de la población. En general,  $f$  será una función de la población.

Supongamos que hay  $n$  especies distintas en la población con grados de adaptación  $f_i$ ,  $i = 1, \dots, n$ . Por definición, la ley de crecimiento de sus respectivas poblaciones será  $\dot{n}_i = f_i n_i$ . La población total,  $N = \sum_{i=1}^n n_i$ , crecerá entonces de acuerdo a la ecuación

$$\dot{N} = \sum_{i=1}^n \dot{n}_i = \sum_{i=1}^n f_i n_i = N \phi, \quad \phi = \sum_{i=1}^n f_i x_i, \tag{3}$$

---

<sup>1</sup>El ser humano aún no ha alcanzado ese equilibrio con su entorno, dado que su entorno es todo el planeta, si bien esta ley se cumple en poblaciones aisladas en entornos muy escasos de recursos. La situación de hambruna que viven muchos países africanos nos muestra cómo será nuestra situación cuando alcancemos ese estado de equilibrio con los recursos del planeta.

siendo  $x_i = n_i/N$  la fracción de población de la especie  $i$ , y  $\phi$  el *grado de adaptación medio* de la población. La población total sigue, pues, una ley de Malthus con tasa de replicación  $\phi$  (recuérdese que las  $f_i$  pueden ser funciones de  $N$ ). Ahora podemos obtener una ley de crecimiento para  $x_i$ ,

$$\dot{x}_i = \frac{\dot{n}_i}{N} - x_i \frac{\dot{N}}{N} = x_i(f_i - \phi). \quad (4)$$

Esta ley de evolución recibe el nombre de *ecuación del replicador* [15], y se aplica no sólo a la evolución de sistemas biológicos, sino que también tiene un papel destacado en Teoría de Juegos [8]. La ley que siguen las fracciones de población recuerda mucho la ley de Malthus, solo que ahora la tasa de replicación está referida a la media de la población. Eso significa que la fracción de población de una especie crecerá si está por encima de esa media, pero decrecerá si está por debajo. Ya se vislumbra el principio de la «supervivencia del más apto». De hecho es fácil deducir este principio de las ecuaciones (3). Supongamos que la especie  $k$  es más apta que las demás, esto es,  $f_k > f_i \forall i \neq k$ ; la ley para  $x_k$  se puede reescribir como

$$\dot{x}_k = x_k \left( f_k - \sum_{i=1}^n f_i x_i \right) = x_k \sum_{i=1}^n (f_k - f_i) x_i = x_k \sum_{i \neq k, i=1}^n (f_k - f_i) x_i, \quad (5)$$

donde hemos usado que  $\sum_{i=1}^n x_i = 1$ . Es evidente de (5) que el sumatorio de la derecha será positivo en tanto que al menos una de las especies  $i \neq k$  mantenga  $x_i > 0$ , y mientras eso ocurra  $x_k$  crecerá en el tiempo a costa de las poblaciones de las restantes especies. En definitiva,

$$\lim_{t \rightarrow \infty} x_k(t) = 1, \quad \lim_{t \rightarrow \infty} x_i(t) = 0, \quad \forall i \neq k, \quad (6)$$

que expresa, de una forma matemática, el principio de supervivencia del más apto.

En el caso en que los grados de adaptación sean constantes tenemos un resultado, debido a Fisher, denominado «teorema fundamental de la selección natural» [5]. Su derivación pasa por calcular

$$\dot{\phi} = \sum_{i=1}^n f_i \dot{x}_i = \sum_{i=1}^n f_i x_i (f_i - \phi) = \sum_{i=1}^n x_i (f_i - \phi)^2 = \sigma_f^2 \geq 0. \quad (7)$$

En otras palabras: el grado de adaptación medio nunca decrece en el tiempo, y su tasa de crecimiento es la varianza del grado de adaptación. Así que crecerá siempre que haya variabilidad en la población, y lo hará a costa de aumentar la población de los más aptos.

### 1.3. MUTACIÓN

El proceso de replicación puede tener errores. En general, estos errores hacen que el nuevo individuo no sea viable y perezca, con lo que se puede dar cuenta de ellos ajustando la tasa de replicación. Sin embargo, ocasionalmente una mutación puede

dar lugar a un descendiente de un tipo distinto pero viable. Así que podemos ver la mutación como un proceso estocástico por el cual individuos de la especie  $i$  pueden engendrar individuos de la especie  $j$  con una probabilidad  $q_{ij}$  ( $\ll 1$ ). Este mecanismo permite que aparezca variabilidad en poblaciones inicialmente homogéneas. La ecuación del replicador (3) debe ser modificada para incluir este nuevo proceso:

$$\dot{x}_i = x_i \left( f_i - f_i \sum_{j \neq i} q_{ij} + \sum_{j \neq i} f_j q_{ji} - \phi \right), \tag{8}$$

donde el primer término nuevo da cuenta de las mutaciones que convierten un individuo de tipo  $i$  en otro de un tipo distinto, y el segundo término da cuenta de las mutaciones que convierten individuos de otro tipo en individuos de tipo  $i$ . Definiendo  $q_{ii} = 1 - \sum_{j \neq i} q_{ij}$  ( $\geq 0$ ) e introduciendo la matriz estocástica  $Q = (q_{ij})$  (que denominaremos *matriz de mutaciones*), la ecuación anterior se puede escribir en forma vectorial

$$\dot{\mathbf{x}} = \mathbf{x}FQ - \phi\mathbf{x}, \tag{9}$$

siendo  $\mathbf{x} = (x_1, \dots, x_n)$  y  $F = \text{diag}(f_1, \dots, f_n)$ . Esta ecuación recibe el nombre de *ecuación del replicador-mutador* [15]. El carácter estocástico de  $Q$  se puede expresar en forma matricial como  $Q\mathbf{1}^T = \mathbf{1}^T$ , siendo  $\mathbf{1} = (1, \dots, 1)$ . Gracias a ello, como  $\phi = \mathbf{x}F\mathbf{1}^T = \mathbf{x}FQ\mathbf{1}^T$ , es inmediato ver que  $\mathbf{x}\mathbf{1}^T = 1$  es una ligadura que (9) preserva.

Como  $FQ$  es una matriz no negativa, la teoría de Perron-Frobenius [17] nos dice que los equilibrios del sistema dinámico (9) están dados por los autovectores por la izquierda de la matriz  $FQ$  correspondientes al mayor autovalor (de componentes no negativas). Si la matriz  $Q$  es, además, irreducible, entonces sólo hay un autovector de ese tipo. Toda matriz irreducible tiene en cada fila al menos un elemento no diagonal, lo que hace que el vector de equilibrio tenga necesariamente más de una componente no nula. Así pues, cuando todas las especies pueden mutar no es posible tener una población homogénea en el equilibrio. Ésta es una observación interesante porque implica que  $\phi$  no se maximiza en el equilibrio, como ocurría en ausencia de mutación: hay una competición entre la selección, que empuja  $\phi$  hacia su máximo, y la mutación, que la hace disminuir. Podemos investigar qué le ha pasado al teorema fundamental en este caso. Un razonamiento similar al que usamos en (7) nos lleva a que

$$\dot{\phi} = \mathbf{x}(FQ - \phi I)^2\mathbf{1}^T, \tag{10}$$

que se parece mucho a la expresión de  $\sigma_f^2$ , salvo por el hecho de que ya no hay garantías de que ese término no sea negativo.

## 2. EL PROBLEMA DE LA REVERSIÓN

Pese a que Mendel publicó su trabajo casi al mismo tiempo que Darwin, parece que éste nunca fue consciente de su existencia. Esto le causó a Darwin un serio problema, y basta leer *El Origen de las Especies* para ver cómo vuelve a él una y otra vez: la teoría de la herencia en poblaciones con reproducción sexual que se tenía por

válida en la época de Darwin (que Galton formuló en 1875 [2]) conducía al fenómeno de la reversión. De acuerdo a esta teoría, la contribución de cada progenitor a un determinado carácter, por ejemplo la altura, es al 50 %. Eso nos permite formular una sencilla ley estocástica para la altura de un descendiente en función de la de los padres:  $X_{n+1} = \frac{1}{2}(X_n^{(1)} + X_n^{(2)}) + Z_n$ , donde  $X_n^{(1)}$  y  $X_n^{(2)}$  son dos variables estocásticas con la misma distribución,  $P_n(x) = \Pr(X_n \leq x)$ , de alturas de la población en la generación  $n$ , y donde  $Z_n$  es un ruido, que podemos suponer, para todo  $n$ , normal, de media 0 y desviación estándar  $\sigma$ . Si  $F_n(q) = \int_{-\infty}^{\infty} e^{iqx} dP_n(x)$  denota la función característica de la distribución  $P_n$ , entonces  $F_{n+1}(q) = F_n(q/2)^2 e^{-\sigma^2 q^2/2}$ . La solución de esta ecuación es

$$\log F_n(q) = 2^n \log F_0(2^{-n}q) - \frac{\sigma^2}{2} q^2 \sum_{k=0}^{n-1} 2^{-k}, \quad (11)$$

y por lo tanto,

$$\lim_{n \rightarrow \infty} F_n(q) = e^{qF'_0(0) - \sigma^2 q^2}. \quad (12)$$

Como  $F'_0(0) = i\mu$ , siendo  $\mu$  el valor medio de la distribución inicial  $P_0(x)$ , de aquí se sigue que  $P_n(x)$  tiende a una normal  $N(\mu, \sqrt{2}\sigma)$  cuando  $n \rightarrow \infty$ . Por lo tanto, independientemente de la distribución inicial (siempre que tenga media finita), la distribución de alturas tiende a estabilizarse en una distribución normal con el mismo valor medio que tenía la distribución inicial. Eso quiere decir que, incluso si ésta es bimodal, al final toda la población acaba siendo de un mismo tipo homogéneo. La dificultad que esto pone a la teoría de Darwin es la de suprimir la variabilidad que introducen las mutaciones. Por ello Darwin recurre frecuentemente al argumento de que las poblaciones mutantes deben permanecer aisladas mucho tiempo para que aparezca una nueva especie.

### 3. MENDEL O LA TEORÍA «CUÁNTICA» DE LA HERENCIA

La aportación crucial de Mendel consistió en descubrir que los caracteres se transmiten en «cuantos» de herencia (lo que hoy denominamos *genes*) y no admiten gradaciones.<sup>2</sup> Caracteres como la altura, que parecen violar este principio, no son sino caracteres complejos resultantes del efecto integrado de muchos caracteres simples, todos los cuales son «cuánticos» (están presentes o ausentes). Cada individuo sexuado posee dos de tales cuantos por cada carácter, uno procedente del padre y otro de la madre, y a su vez transmite uno de los dos (al azar) a cada descendiente.

Imaginemos un carácter (p. ejp., el color rojo de una rosa) determinado por una variante A del correspondiente gen (lo que en Genética se denomina un *alelo*). Imaginemos que este alelo tiene una variedad mutada, a, que no produce el color. De acuerdo al modelo hereditario de Galton, los descendientes híbridos deberían ir mostrando variedades tonales del rosa hasta que, transcurridas suficientes generaciones,

<sup>2</sup>No deja de ser una notable coincidencia que de Vries redescubriera para el mundo científico la genética mendeliana en 1900, el mismo año en que Planck propone por primera vez la hipótesis cuántica.

la especie revertera al tipo rojo original. De acuerdo con las leyes descubiertas por Mendel, si en la población inicial de padres hay  $p$  alelos de tipo A y  $q$  de tipo a; si los individuos se emparejan totalmente al azar, y si el tamaño de la población es suficientemente grande, entonces en la segunda generación habrá  $p^2$  individuos de tipo AA,  $2pq$  individuos de tipo Aa y  $q^2$  individuos de tipo aa, y a partir de ese momento ésa es la distribución estable de tipos en la población. Este resultado es conocido en Genética como la ley de Hardy-Weinberg [5]. Esta ley implica que no hay reversión: los tipos mutantes puros aa permanecen, generación tras generación, en la población en una proporción que depende de cuántos mutantes hay inicialmente. La genética mendeliana es, pues, la responsable de que la mutación pueda introducir y mantener variabilidad en las poblaciones.

#### 4. EL CUARTO ELEMENTO: LA DERIVA GENÉTICA

Hemos comentado al principio de la sección 1 que replicación, selección y mutación forman la terna básica de una dinámica evolutiva. Aunque estrictamente esto es así, hay un cuarto ingrediente que, aunque innecesario en principio, es sin embargo crucial para entender los mecanismos de adaptación y especiación de las entidades sometidas a un proceso evolutivo: la *deriva genética*. Todo el análisis que hemos hecho hasta ahora presupone que las poblaciones sometidas a evolución son infinitamente grandes; por eso las dinámicas que hemos estado discutiendo son deterministas, y son válidas leyes como la de Hardy-Weinberg, por ejemplo. Cuando hablamos de poblaciones finitas aparece un ruido debido al simple muestreo estadístico que hay implícito en los procesos de replicación. Ese ruido es lo que se conoce como deriva genética, y sus efectos pueden llegar a ser dramáticos en poblaciones pequeñas. Por eso resulta un mecanismo fundamental cuando las poblaciones pasan por un «cuello de botella» evolutivo, esto es, una situación en la que la población se queda muy reducida (por efecto de epidemias, cambios climáticos, aislamiento de poblaciones, etc.).

Hay dos modelos matemáticos básicos de la deriva genética: el de Fisher-Wright y el de Moran [5]. Ambos presuponen que el tamaño de la población permanece constante, limitado por el entorno. El primero es adecuado en situaciones en que cada generación reemplaza a la anterior (como es el caso de las plantas estacionales); el segundo describe mejor una situación en la que las generaciones se solapan.

##### 4.1. MODELO DE FISHER-WRIGHT

Imaginemos una población de  $N$  individuos y fijémonos en un determinado carácter (el color rojo de una rosa, de nuevo). Supongamos que ese carácter viene determinado por un alelo A y que hay un mutante a que muestra otra variante del carácter (color blanco, p. ejp.). Imaginemos que la población se reproduce de forma asexual (es decir, A engendra A y a engendra a) y que hay una probabilidad  $\mu$  de que ocurra la mutación  $A \rightarrow a$  y una probabilidad  $\nu$  de que ocurra la mutación contraria  $a \rightarrow A$ . Si inicialmente hay  $k$  individuos con el alelo A y no hay diferencias

selectivas entre los dos alelos (ambos se reproducen con la misma tasa de replicación), entonces la probabilidad de que un individuo de la generación siguiente sea de tipo A será  $\psi_k = (k/N)(1 - \mu) + (1 - k/N)\nu$ . Imaginemos que cada individuo engendra un gran número de descendientes y que de ellos se extraen, de forma aleatoria, los  $N$  individuos de la siguiente generación. Si  $X_t \in \{0, 1, \dots, N\}$  es una variable aleatoria que denota el número de alelos A de la generación  $t$ , entonces el modelo de Fisher-Wright está definido como una cadena de Markov en la que

$$P_{kj} = \Pr(X_{t+1} = j \mid X_t = k) = \binom{N}{j} \psi_k^j (1 - \psi_k)^{N-j}. \quad (13)$$

La cadena es ergódica para todo  $0 < \mu, \nu < 1$  ya que  $P_{kj} > 0 \forall k, j \in \{0, 1, \dots, N\}$ , por tanto tiene una distribución de probabilidad estacionaria,  $\mathbf{w} = (w_0, w_1, \dots, w_N)$ , que se obtiene mediante la ecuación  $\mathbf{w} = \mathbf{w}P$  [10, 17]. No se conoce una expresión exacta para  $\mathbf{w}$  excepto para un caso particular:  $\mu + \nu = 1$  (sin interés biológico, ya que supone unas tasas de mutación enormes), porque entonces  $\psi_k = \nu = 1 - \mu$  y

$$(\mathbf{w}P)_j = \binom{N}{j} \nu^j (1 - \nu)^{N-j} \quad (14)$$

para cualquier  $\mathbf{w}$ , con lo que el lado derecho de la ecuación anterior determina la distribución estacionaria.

Lo que sí se puede calcular de manera general es el valor medio de  $\mathbf{w}$ . Para ello denotamos  $X_\infty = \lim_{t \rightarrow \infty} X_t$  y definimos el vector  $\boldsymbol{\xi} = (0, 1, \dots, N)$ , con el cual podemos escribir  $E(X_\infty) = \mathbf{w}\boldsymbol{\xi}^T = \mathbf{w}P\boldsymbol{\xi}^T$ . Ahora bien,

$$(P\boldsymbol{\xi}^T)_k = \sum_{j=0}^N j \binom{N}{j} \psi_k^j (1 - \psi_k)^{N-j} = N\psi_k = k(1 - \mu) + (N - k)\nu, \quad (15)$$

por lo que  $E(X_\infty) = (1 - \mu)E(X_\infty) + [N - E(X_\infty)]\nu$  y de ahí  $E(X_\infty) = N\nu/(\mu + \nu)$ . Por un procedimiento análogo (aunque más tedioso) se puede obtener la varianza  $\sigma^2 = N^2\mu\nu/[(\mu + \nu)^2(1 + 2N\mu + 2N\nu)] + \epsilon$ , donde  $\epsilon$  contiene términos de menor orden (suponiendo, p. ejp., que  $\mu, \nu = O(N^{-1})$ , entonces  $\epsilon = O(N)$ ).

El caso más interesante a analizar se da cuando sólo hay deriva genética ( $\mu = \nu = 0$ , con lo que  $\psi_k = k/N$ ). La cadena de Markov deja entonces de ser ergódica y pasa a tener dos estados absorbentes:  $k = N$  (todos los individuos de tipo A) y  $k = 0$  (todos los individuos de tipo a). Eso significa que, independientemente del estado inicial de la población y pese a no haber factores selectivos que favorezcan uno u otro tipo, con el transcurrir del tiempo la población termina siendo de un tipo homogéneo. La magnitud de interés entonces es la probabilidad  $\pi_j = \Pr(X_\infty = N \mid X_0 = j)$  de que la población acabe siendo de tipo A si inicialmente hay  $j$  individuos de ese tipo. Es evidente que  $\boldsymbol{\pi}^T = P\boldsymbol{\pi}^T$ , denotando  $\boldsymbol{\pi} = (\pi_0, \pi_1, \dots, \pi_N)$ , y que  $\pi_0 = 0$  y  $\pi_N = 1$ ; sin embargo, la forma más directa de hallar  $\boldsymbol{\pi}$  es probando que la cadena de Markov es una martingala, esto es,  $E(X_t \mid X_{t-1}) = X_{t-1}$  (la prueba es inmediata: se trata del valor medio de la distribución binomial). Eso significa que  $E(X_\infty) = j$ . Pero este valor medio se puede obtener también como  $E(X_\infty) = N\pi_j + 0(1 - \pi_j)$ , de donde



$\pi_j = j/N$ . Una interesante consecuencia de este resultado es la probabilidad de que un mutante acabe invadiendo la población:  $\pi_1 = 1/N$ . Esta probabilidad es pequeña en poblaciones grandes, pero no es despreciable en los cuellos de botella evolutivos, así que la fijación de un nuevo alelo es algo que seguramente ha ocurrido más de una vez en el pasado.

#### 4.2. MODELO DE MORAN

El modelo de Moran es más interesante desde un punto de vista teórico porque permite hacer cálculos analíticos con más facilidad que el de Fisher-Wright. La situación de partida es la misma:  $k$  alelos de tipo A y  $N - k$  de tipo a en una población de  $N$  individuos. La diferencia es que ahora los individuos se reproducen en el tiempo a una tasa constante  $\tau$ . El vástago será un mutante con las mismas probabilidades  $\mu$  y  $\nu$  que hemos considerado en el proceso de Fisher-Wright. Tras el evento de reproducción el nuevo individuo reemplazará a cualquier individuo de la población (¡incluso a su progenitor!) con la misma probabilidad. El proceso de Markov queda especificado dando las probabilidades condicionadas

$$\Pr(X(t + dt) = j \mid X(t) = k) = \tau T_{kj} dt, \quad \forall j \neq k, \tag{16}$$

donde  $X(t)$  representa la población de tipo A en el instante  $t$ . La magnitud de interés en este proceso estacionario es  $P_{ij}(t) = \Pr(X(t + s) = j \mid X(s) = i)$ , la solución de la ecuación que se obtiene multiplicando (16) por  $P_{ik}(t)$  y sumando en  $0 \leq k \leq N$ ,

$$\dot{P}_{ij}(t) = \tau \sum_{k=0}^N [P_{ik}(t)T_{kj} - P_{ij}(t)T_{jk}], \tag{17}$$

o bien multiplicando (16) por  $P_{jl}(t)$  y sumando en  $0 \leq j \leq N$  (con el adecuado cambio de índices),

$$\dot{P}_{ij}(t) = \tau \sum_{k=0}^N [T_{ik}P_{kj}(t) - T_{ik}P_{ij}(t)]. \tag{18}$$

La condición inicial en ambos casos es, por supuesto,  $P_{ij}(0) = \delta_{ij}$ . Las ecuaciones (17) y (18) son, respectivamente, las formas avanzada y retardada de la *ecuación maestra* del proceso.

En el proceso de Moran, en un intervalo dado  $(t, t + dt)$  sólo puede haber un cambio desde el estado  $k$  a los estados  $k, k + 1$  ó  $k - 1$ , dado que a lo sumo puede tener lugar un único proceso de reproducción. Denotando  $\tau T_{kk+1} = \lambda_k$  y  $\tau T_{kk-1} = \mu_k$ , la ecuación (17) se convierte en la *ecuación de Kolmogorov prospectiva*

$$\dot{P}_{ij}(t) = \mu_{j+1}P_{i,j+1}(t) - (\mu_j + \lambda_j)P_{ij}(t) + \lambda_{j-1}P_{i,j-1}(t), \quad 0 \leq j \leq N, \tag{19}$$

y la ecuación (18) en la *ecuación de Kolmogorov retrospectiva*

$$\dot{P}_{ij}(t) = \lambda_i P_{i+1,j}(t) - (\mu_i + \lambda_i)P_{ij}(t) + \mu_i P_{i-1,j}(t), \quad 0 \leq j \leq N, \tag{20}$$

siendo  $P_{i-1}(t) = P_{-1i}(t) = P_{iN+1}(t) = P_{N+1i}(t) = \mu_0 = \lambda_N = 0$ . Los procesos de Markov de este tipo reciben el nombre de *procesos de nacimiento-muerte* [10]. En el proceso de Moran  $\lambda_k = \tau\psi_k(1 - k/N)$  (la probabilidad de que aparezca un alelo A y reemplace a uno a) y  $\mu_k = \tau(1 - \psi_k)k/N$  (la probabilidad de que aparezca un alelo a y reemplace a uno A), siendo  $\psi_k$  como en el proceso de Fisher-Wright.

El estado estacionario  $w_j = \lim_{t \rightarrow \infty} P_{ij}(t)$  se puede obtener anulando la derivada en (19), lo que nos da  $\mu_{j+1}w_{j+1} - \lambda_jw_j = \mu_jw_j - \lambda_{j-1}w_{j-1}$  para todo  $1 < j < N$ . Para  $j = 0$  tenemos  $\mu_1w_1 - \lambda_0w_0 = 0$ , por lo tanto  $\mu_jw_j - \lambda_{j-1}w_{j-1} = 0$ , una sencilla ecuación en diferencias cuya solución es

$$w_j = w_0 \frac{\lambda_0 \lambda_1 \cdots \lambda_{j-1}}{\mu_1 \mu_2 \cdots \mu_j}, \quad 0 < j \leq N, \tag{21}$$

o alternativamente

$$w_j = w_N \frac{\mu_{j+1} \mu_{j+2} \cdots \mu_N}{\lambda_j \lambda_{j+1} \cdots \lambda_{N-1}}, \quad 0 \leq j < N. \tag{22}$$

El valor de  $w_0$  o el de  $w_N$  se determinan mediante la normalización  $\sum_{j=0}^N w_j = 1$ .

Para el modelo de Moran con mutaciones se puede obtener una expresión cerrada para  $w$  en el límite  $N \rightarrow \infty, k \rightarrow \infty$  con  $k/N = x \in [0, 1], N\mu \rightarrow \gamma$  y  $N\nu \rightarrow \kappa$  [10]. Para ello escribimos  $\lambda_j = A(N - j)j[1 + a/j]$  y  $\mu_j = A(N - j)j[1 + b/(N - j)]$ , donde  $A = \tau(1 - \mu - \nu)/N^2, a = N\nu/(1 - \mu - \nu)$  y  $b = N\mu/(1 - \mu - \nu)$ . Entonces

$$w_k = \frac{w_0 N a}{k(N - k)[1 + b/(N - k)]} \prod_{j=1}^{k-1} \frac{1 + a/j}{1 + b/(N - j)}. \tag{23}$$

Ahora bien, usando la serie de Taylor de  $\log(1 + x)$  y que  $\sum_{j=1}^k j^{-1} \sim \log k$  cuando  $k \rightarrow \infty$  en tanto que  $\sum_{j=1}^{\infty} j^{-p} < \infty$  para todo entero  $p > 1$ ,

$$\begin{aligned} \sum_{j=1}^{k-1} \log \left( 1 + \frac{a}{j} \right) &\sim \log(k^a) + c, \\ \sum_{j=1}^{k-1} \log \left( 1 + \frac{b}{N - j} \right) &\sim \log \left( \frac{N^b}{(N - k)^b} \right) + d \end{aligned} \tag{24}$$

cuando  $k \rightarrow \infty$ , para ciertas constantes  $c$  y  $d$ . En consecuencia (teniendo en cuenta que en el límite que estamos considerando  $b/(N - k) \rightarrow 0$ ),

$$w_k \sim C k^{\kappa-1} \left( 1 - \frac{k}{N} \right)^{\gamma-1}, \tag{25}$$

siendo  $C$  una constante de normalización. Para fijarla tenemos en cuenta que

$$\sum_{k=0}^N w_k = C N^{\kappa} \sum_{k=0}^N \frac{1}{N} \left( \frac{k}{N} \right)^{\kappa-1} \left( 1 - \frac{k}{N} \right)^{\gamma-1} \sim C N^{\kappa} \int_0^1 x^{\kappa-1} (1 - x)^{\gamma-1} dx, \tag{26}$$

por lo tanto, la distribución estacionaria  $w_k \sim x^{\kappa-1}(1-x)^{\gamma-1} dx/B(\kappa, \gamma)$  tiene la forma de una distribución beta de parámetros  $\kappa$  y  $\gamma$ .

En ausencia de mutaciones ( $\mu = \nu = 0$  y  $\psi_k = k/N$ ),  $\lambda_0 = \mu_N = 0$ . Eso hace que (21) y (22) den lugar, respectivamente, a dos vectores:  $\mathbf{w} = (1, 0, \dots, 0)$  y  $\mathbf{w} = (0, \dots, 0, 1)$ , que expresan el hecho de que, con probabilidad 1, el sistema acaba absorbido en  $k = 0$  o en  $k = N$ . La magnitud que caracteriza esta absorción es  $\pi_k = \lim_{t \rightarrow \infty} P_{kN}(t)$ , es decir, la probabilidad de que empezando el proceso con  $k$  alelos A acabe con  $N$  alelos de ese tipo. La ecuación para  $\pi$  la deducimos anulando del miembro de la izquierda de (20). La ecuación resultante se puede escribir  $\lambda_k(\pi_{k+1} - \pi_k) = \mu_k(\pi_k - \pi_{k-1})$ , donde  $0 < k < N$ . Teniendo en cuenta que  $\pi_0 = 0$ , la ecuación implica que  $\pi_k - \pi_{k-1} = q_{k-1}\pi_1$ , para  $0 < k \leq N$ , donde

$$q_0 = 1, \quad q_k = \frac{\mu_1 \mu_2 \cdots \mu_k}{\lambda_1 \lambda_2 \cdots \lambda_k}. \tag{27}$$

Entonces  $\pi_k = \pi_1 \sum_{j=0}^{k-1} q_j$  para  $0 < k \leq N$ , y  $\pi_1$  se deduce de la condición  $\pi_N = 1$ . La forma de  $\pi_k$  para el proceso de Moran es muy simple, ya que  $\mu_j = \lambda_j$  para todo  $0 \leq j \leq N$ . Entonces  $\pi_k = k/N$ , igual que en el proceso de Fisher-Wright.

Karlin y McGregor demostraron [9] que la solución de (20) puede expresarse en la forma

$$P_{ij}(t) = \frac{w_j}{w_0} \int_0^\infty e^{-xt} R_i(x) R_j(x) d\varphi(x), \tag{28}$$

donde  $w_j$  está dado por (21); los  $R_j(x)$  son el sistema (finito porque  $\lambda_N = 0$ ) de polinomios definido por la recurrencia a tres términos

$$-xR_j(x) = \lambda_j R_{j+1}(x) - (\lambda_j + \mu_j) R_j(x) + \mu_j R_{j-1}(x), \quad 0 \leq j < N, \tag{29}$$

con  $R_{-1}(x) = 0$  y  $R_0(x) = 1$ , y  $\varphi(x)$  es una medida única, de masa unidad, con incrementos en  $N + 1$  puntos, respecto a la que la familia de polinomios es ortogonal.

Una ventaja añadida del modelo de Moran es que permite incluir selección. Si  $f_A$  y  $f_a$  denotan los grados de adaptación de cada tipo, la probabilidad de que, dado un evento de reproducción, se reproduzca un individuo de tipo  $x$  será proporcional a  $f_x$ . Eso se traduce en una nueva expresión para  $\psi_k$ ,

$$\psi_k = \frac{k f_A (1 - \mu) + (N - k) f_a \nu}{k f_A + (N - k) f_a}. \tag{30}$$

Argumentos análogos a los utilizados para hallar, en el límite de poblaciones grandes y mutaciones pequeñas, la distribución estacionaria en ausencia de selección, conducen ahora a  $w_k \sim (N \log r^{-1})^{\kappa/r} x^{\kappa/r-1} r^{Nx} dx/\Gamma(\kappa/r)$  si  $r < 1$ , y  $w_k \sim (N \log r)^{\gamma r} (1-x)^{\gamma r-1} r^{N(x-1)} dx/\Gamma(\gamma r)$  si  $r > 1$ , donde  $r = f_A/f_a$  es el grado de adaptación relativo del alelo A frente al a.

Cuando no hay mutaciones  $\mu_j/\lambda_j = r^{-1}$ , por lo tanto  $\pi_k = (1 - r^{-k})/(1 - r^{-N})$ . Las probabilidades de fijación de un alelo mutante ( $\rho_A = \pi_1$  ó  $\rho_a = 1 - \pi_{N-1}$ ) serán, pues,

$$\rho_A = \frac{1 - r^{-1}}{1 - r^{-N}}, \quad \rho_a = \frac{1 - r}{1 - r^N}. \tag{31}$$

Como cabe esperar, si  $r > 1$  crece la probabilidad de fijar un alelo A y decrece la de fijar un alelo a, y viceversa si  $r < 1$ .

### 4.3. LA APROXIMACIÓN DE DIFUSIÓN

En el límite en que la población  $N \rightarrow \infty$ ,  $i, j \rightarrow \infty$  con  $i/N \rightarrow y$ ,  $j/N \rightarrow x$ , si  $NP_{ij}(t) \rightarrow f(x, t | y, 0)$ ,  $\lambda_j - \mu_j \rightarrow m(x)$  y  $\lambda_j + \mu_j \rightarrow s(x)$ , entonces las ecuaciones (19) y (20) se convierten, respectivamente, en

$$\frac{\partial}{\partial t} f(x, t | y, 0) = -\frac{\partial}{\partial x} [m(x)f(x, t | y, 0)] + \frac{1}{2N} \frac{\partial^2}{\partial x^2} [s(x)f(x, t | y, 0)], \quad (32)$$

$$\frac{\partial}{\partial t} f(x, t | y, 0) = -m(y) \frac{\partial}{\partial y} f(x, t | y, 0) + \frac{s(y)}{2N} \frac{\partial^2}{\partial y^2} f(x, t | y, 0). \quad (33)$$

En ambos casos la condición inicial es  $f(x, 0 | y, 0) = \delta(x - y)$ . Son dos ecuaciones de difusión que describen el proceso de Moran en este régimen especial. Es de destacar el hecho de que el término de difusión es proporcional a  $N^{-1}$ , lo que significa que si las poblaciones son muy grandes la aproximación determinista es válida. Este término también justifica la visión de que las poblaciones finitas actúan como un «ruido» sobre la aproximación determinista.

La aproximación de difusión puede realizarse también sobre el proceso de Fisher-Wright; el resultado es similar, pero con una escala de tiempo  $N$  veces más rápida que en el proceso de Moran. Esta aproximación permite hacer cálculos que serían inviables en la forma discreta de los procesos, y puede generalizarse con facilidad a situaciones en que hay múltiples alelos o en las que los individuos poseen más de un gen. Explorar sus posibilidades nos llevaría muy lejos del objetivo de este artículo, por lo que remito al lector interesado a las referencias [1, 5, 11].

## 5. EL PAPEL DE LA DERIVA GENÉTICA EN LA EVOLUCIÓN

La deriva genética aparece en los estudios iniciales de genética de poblaciones como un mecanismo adicional que permite dar cuenta del «ruido» que introducen en la dinámica evolutiva las poblaciones pequeñas, ruido que adquiere una especial importancia en los momentos singulares en que las especies atraviesan cuellos de botella evolutivos. Sin embargo, los avances producidos en el campo de la biología molecular han ido destacando cada vez más la relevancia que tiene este mecanismo incluso en situaciones normales, hasta el punto de que nos encontramos ante un incipiente cambio de paradigma: la evolución neutra, por simple deriva genética, parece no sólo la forma más común de evolución de los organismos, sino también el mecanismo que permite *de facto* su adaptación y especiación. La deriva genética es, pues, una de las claves del origen de las especies.

Pero para poder indagar un poco en el significado de la evolución neutra necesitamos hacer antes una pequeña excursión por la Biología.

### 5.1. LA BIOLOGÍA SE CONSTRUYE SOBRE SECUENCIAS Y REDES

Al nivel más fundamental, la vida está descrita por una molécula de ADN, que consiste en una larga secuencia de cuatro tipos de bases: adenina (A), timina (T), guanina (G) y citosina (C). Podríamos decir que esa secuencia de bases es una

codificación de toda la información sobre la construcción de una célula (y en última instancia de un ser vivo).<sup>3</sup>

El ADN se divide en intrones, o partes de código que no codifican proteínas, y exones, los tramos codificantes. En los primeros estadios de la transcripción, el ADN transfiere su información a moléculas de ARN (reemplazando la timina por una nueva base, U: el uracilo), en las cuales, mediante un mecanismo denominado *splicing*, se aíslan los exones y se combinan para formar los genes. Éstos son los que se traducen a proteínas en los ribosomas celulares (una especie de «máquinas» lectoras-traductoras). A este nivel podríamos decir que las cadenas de ADN que forman los cromosomas se pueden ver realmente como secuencias de genes.

El ARN se transcribe a proteínas convirtiendo *codones* (secuencias de tres bases consecutivas) en aminoácidos. Esta traducción es un código universal<sup>4</sup> que se conoce como *código genético* y que da lugar a nuevas secuencias (las proteínas), esta vez formadas por 20 tipos distintos de aminoácidos. Como consecuencia de las interacciones entre sus aminoácidos, las proteínas se pliegan en estructuras tridimensionales, a veces rígidas, a veces con partes móviles, como si de pequeñas máquinas se tratara. Esa estructura tridimensional determina su función, hasta el punto de que un cambio en su conformación puede hacer que la proteína pierda o cambie la función biológica que desempeña. El conjunto de proteínas de una célula (el proteoma) forma un complejo entramado: las proteínas interaccionan entre sí y con el genoma en formas muy distintas, activando o inhibiendo la producción de otras proteínas, catalizando reacciones, etc. El resultado es una red metabólica de interacciones, una especie de ecosistema proteínico, cuyo resultado es la propia actividad celular.

Todavía se puede ascender más en la escala y considerar los organismos pluricelulares como otra compleja red de distintos tipos de células que interaccionan entre sí, y a su vez estos organismos (animales, plantas...) entrelazan su actividad vital compitiendo por recursos, comiéndose unos a otros, cooperando... para dar lugar al orden más alto en la escala biológica: los ecosistemas.

Todas estas organizaciones biológicas, sean secuencias o redes, comparten una propiedad: están formadas por un conjunto definido de elementos, de tal modo que su modificación (por cambio, eliminación o adición) puede cambiar completamente la estructura. Al nivel más básico, en las cadenas de ADN, estas modificaciones son las denominadas mutaciones. Éstas pueden consistir en el cambio de unas bases por otras, en la adición o eliminación de una o más bases, en la inversión de tramos completos de la secuencia, en su duplicación, etc. El efecto de algunas de estas transformaciones es tan drástico que la mutación resulta letal: el organismo resultante deja de ser viable (por ejemplo, imagínese el efecto de eliminar una sola

---

<sup>3</sup>Este dogma clásico no es, en realidad, totalmente cierto, ya que en la división celular cada célula hija no sólo hereda una copia literal del ADN de la célula progenitora, sino también el material contenido en el citoplasma. Eso hace que las dos células hijas no sean exactamente iguales y que su (si bien ligeramente) distinta composición dé lugar a distintas expresiones de los genes. Éste es el fundamento de la diferenciación celular, y su estudio forma parte de lo que hoy día se conoce como *epigenética*.

<sup>4</sup>En 1979 se descubrió que no es completamente universal: mitocondrias y algunas bacterias y levaduras utilizan códigos ligeramente distintos del «universal».

base, si tenemos en cuenta que la transcripción a proteínas se realiza leyendo secuencias consecutivas de tres bases). Sin embargo otras pueden resultar inocuas (y algunas incluso pueden dar lugar a modificaciones viables): el código genético contiene 64 codones distintos que codifican tan solo 20 aminoácidos más una señal de parada; la redundancia es tal que algunos aminoácidos están codificados hasta por 6 codones distintos. Esto significa que hay muchas sustituciones de bases que no producen absolutamente ningún cambio en la transcripción a proteínas y son, por tanto, inocuas.

Pensemos ahora en las proteínas. Ya hemos comentado que, como consecuencia de la interacción entre sus aminoácidos, éstas se pliegan formando una estructura tridimensional (estructura terciaria) que determina su función. Pero resulta que en ese plegamiento muchos aminoácidos apenas influyen, en realidad éste depende de una fracción de aminoácidos situados en lugares estratégicos. Cambiar alguno de estos aminoácidos cambia la estructura terciaria y en consecuencia la función de la proteína, pero cambiar alguno de los restantes, o bien no cambia esta estructura, o bien lo hace levemente. Eso significa que incluso mutaciones que sí cambian los aminoácidos de la proteína pueden no tener ningún efecto funcional apreciable, o éste ser muy leve. Así que la transcripción a proteínas aumenta (enormemente) la redundancia que ya existía en el código genético.

En el entramado metabólico que forman las proteínas de una célula también es frecuente que añadir o eliminar una proteína, o reemplazarla por otra, altere poco la dinámica global del sistema, así que también a este nivel ocurren cambios inocuos no sujetos, por tanto, a selección. Y no sólo a este nivel: lo mismo ocurre en el nivel ecológico con las especies que conforman un ecosistema. Hay, por tanto, cambios que pueden ocurrir a todas las escalas de un sistema biológico cuyo efecto es nulo o casi nulo y sobre los que la selección, por tanto, no puede actuar. Pero, ¿qué fracción representan frente a todos los cambios posibles?

## 5.2. LA MAYOR PARTE DE LA EVOLUCIÓN ES NEUTRA

En 1968 Kimura sorprendió a la comunidad científica con la propuesta de que la mayor parte de las mutaciones del genoma de los mamíferos son neutras [12]. Su argumento era el siguiente: los estudios comparativos de algunas proteínas dan como resultado que en cadenas de alrededor de 100 aminoácidos ocurra una sustitución cada 28 millones de años; la longitud de las cadenas de ADN de una de las dos dotaciones cromosómicas de los mamíferos es de alrededor de  $4 \times 10^9$  pares de bases; cada 3 pares de bases codifican un aminoácido y, debido a la redundancia, sólo el 80 % de las sustituciones de un par de bases provocan una sustitución de un aminoácido, de manera que hay 16 millones de sustituciones en toda la dotación genómica cada 28 millones de años, o lo que es lo mismo, aproximadamente ¡una sustitución cada 2 años! La conclusión de Kimura es que los organismos sólo pueden tolerar semejante carga mutacional si la inmensa mayoría de las mutaciones son neutras.

Estudios relativamente recientes de moléculas de ARN llegan a conclusiones similares [6]. Las moléculas de ARN se pliegan sobre sí mismas en función de las secuencias de bases que las componen. Ese plegamiento constituye el fenotipo de las

moléculas y, por tanto, sobre él actúa la selección natural. La cantidad de secuencias distintas que se pliegan en una misma estructura es enorme, lo que implica, una vez más, que una inmensa cantidad de mutaciones de la molécula dejan su fenotipo intacto escapando así de la selección.

La neutralidad parece ser la tónica, al menos al nivel molecular. Las consecuencias de ello son de muy profundo calado, pero para poder hacernos una idea cabal tenemos que introducir un concepto nuevo: los *paisajes adaptativos*, y recurrir otra vez a la Matemática.

### 6. PAISAJES ADAPTATIVOS

Quizá la contribución más relevante de Sewall Wright a la teoría evolutiva sea su metáfora del *paisaje adaptativo* [7]. Desde un punto de vista formal, un paisaje adaptativo es una aplicación  $f : \mathcal{X} \rightarrow \mathbb{R}$ , donde  $\mathcal{X}$  denota un cierto espacio de configuraciones equipado con alguna noción de adyacencia, proximidad, distancia o accesibilidad, y donde la imagen es el grado de adaptación que se asocia a la configuración concreta de  $\mathcal{X}$ . Dicha estructura y el hecho de que el conjunto de todos los paisajes adaptativos se pueda ver como el espacio vectorial  $\mathbb{R}^{|\mathcal{X}|}$ , permiten desarrollar una elaborada teoría de paisajes combinatorios que abarca no sólo los paisajes adaptativos de la Biología, sino también los paisajes de energía que aparecen en la Física o las funciones de optimización combinatoria que aparecen en problemas de computación [16].

En el caso que nos ocupa,  $\mathcal{X}$  está formado por el conjunto de las secuencias o redes subyacentes. Al referirnos a secuencias hablaremos de forma genérica de posiciones para referirnos a características genéticas concretas, y de alelos para referirnos a las variantes que pueden ocupar dichas posiciones. En función del contexto, pueden tratarse de distintas bases (ADN), aminoácidos (proteínas) o genes (cromosomas). Así pues, para una cadena de ADN de longitud  $L$ ,  $\mathcal{X} = \{A, T, C, G\}^L$ ; para una proteína de la misma longitud,  $\mathcal{X} = \{\text{Phe, Leu, Ile, } \dots, \text{Gly}\}^L$ , y para un cromosoma  $\mathcal{X} = \{A_1, \dots, A_{n_a}\} \times \{B_1, \dots, B_{n_b}\} \times \dots$ , donde las letras indican los distintos alelos de cada posición (*locus*). Sobre  $\mathcal{X}$  se puede definir la distancia Hamming,  $d_H(x, y)$ , como el número de posiciones distintas entre  $x, y \in \mathcal{X}$ .

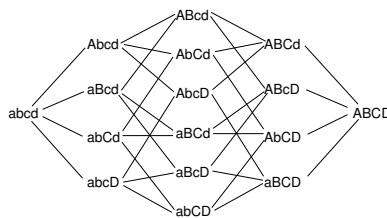


Figura 1: Grafo para una secuencia de longitud 4 y 2 alelos por sitio.

La estructura de  $\mathcal{X}$  está determinada por las transiciones permitidas entre sus secuencias, que denominaremos, también genéricamente, mutaciones. Hablaremos de mutaciones puntuales para referirnos a cambios en una única posición de la cadena, esto es, sustituciones de una base en una cadena de ADN, de un aminoácido en una proteína, o de una variante de un gen (un alelo propiamente dicho) en un cromosoma. Si todas las mutaciones son puntuales podemos construir el grafo  $\mathcal{G} = \{\mathcal{X}, \mathcal{L}\}$ , donde el conjunto de enlaces  $\mathcal{L}$  está definido por los pares de secuencias  $x, y \in \mathcal{X}$  tales que  $d_H(x, y) = 1$  (la figura 1 muestra un ejemplo para  $L = 4$  y dos alelos por posición).

Podemos también considerar mutaciones que supongan la eliminación o adición de una posición de la secuencia. Si  $\mathcal{X}_L$  denota las configuraciones para una longitud dada  $L$ , entonces  $\mathcal{X} = \bigcup_L \mathcal{X}_L$ , y el grafo incluirá enlaces entre secuencias de distintas longitudes.

La probabilidad de que ocurra una mutación dada de entre las que definen los enlaces de  $\mathcal{G}$  no tiene por qué ser uniforme. La forma genérica de describir la evolución de una secuencia dada es definiendo una matriz de transiciones  $T$ , cuyo elemento  $T_{xy}$  sea la probabilidad de pasar de la secuencia  $x \in \mathcal{X}$  a la secuencia  $y \in \mathcal{X}$ . Los ceros de la matriz de adyacencia del grafo  $\mathcal{G}$  son también ceros de la matriz  $T$ . La evolución de la secuencia es, por tanto, el camino aleatorio sobre  $\mathcal{X}$  descrito por el proceso de Markov asociado a  $T$ .

Pero hablábamos al principio de esta sección de metáfora. Y en efecto, por encima de la definición formal y el tratamiento matemático que podamos hacer de los paisajes adaptativos está la imagen que éstos nos proporcionan y que guía nuestras intuiciones. De hecho, en el desarrollo de la genética de poblaciones tres han sido las metáforas que se han empleado: el paisaje Fujiyama (Fisher), el paisaje rugoso (Wright) y el paisaje plano (Kimura). Examinémoslos en detalle.

### 6.1. PAISAJE FUJIYAMA: CUASIESPECIES Y LA CATÁSTROFE DE ERROR

Fisher concebía que las especies se encontraban en una situación en que su adaptación era óptima. Por ello imaginaba que éstas se encontraban en una cima local del paisaje adaptativo. Según esta imagen, una secuencia tiene grado de adaptación máximo, y a partir de ella, según nos alejamos en distancia Hamming, el grado de adaptación decrece progresivamente. Fisher no tenía en mente secuencias cuando elaboró su metáfora, porque la biología molecular estaba aún en una fase incipiente; fue Eigen [4] quien con ella construyó su teoría de las *cuasiespecies* y dedujo la *catástrofe de error*. Si el proceso de Markov definido sobre el conjunto de secuencias  $\mathcal{X}$  es ergódico, habrá una distribución de probabilidad estacionaria. Cuando el paisaje es tipo Fujiyama, esta distribución estará localizada alrededor de la secuencia de grado de adaptación máximo o secuencia maestra. Si bien la secuencia maestra es la más probable, alrededor de ella (con algunas mutaciones) hay una «nube» de secuencias un poco menos adaptadas que coexisten con la secuencia maestra. Cuasiespecie es el término que Eigen acuñó para describir esta colectividad.

Para investigar el comportamiento de una cuasiespecie recurriremos a la ecuación (9). Por simplicidad supondremos que las secuencias son de longitud fija  $L \gg 1$ , que sólo tienen dos alelos por posición, que la secuencia maestra tiene una tasa de reproducción  $f_1 = f > 1$  y que todas las demás tienen  $f_2 = \dots = f_{2^L} = 1$ . También supondremos que la probabilidad de que ocurra una mutación puntual es  $\mu \ll 1$ . Denotaremos la fracción de población de la secuencia maestra  $x_1 = x$  y por lo tanto  $x_2 + \dots + x_{2^L} = 1 - x$  y  $\phi = fx + 1 - x$ . La ecuación (9) se convierte entonces en

$$\dot{x} = x[f(1 - \mu)^L - 1 - (f - 1)x] + O(\mu). \quad (34)$$

Los términos  $O(\mu)$  dan cuenta de las transiciones que desde las  $L$  secuencias vecinas de la secuencia maestra revierten de nuevo a ella. Despreciando estos términos y



aproximando  $(1 - \mu)^L \approx e^{-L\mu}$  vemos que si  $fe^{-L\mu} > 1$  entonces  $x$  tiende asintóticamente a  $x^* = (e^{-L\mu}f - 1)/(f - 1)$ , en tanto que si  $fe^{-L\mu} < 1$  el paréntesis de la ecuación (34) es negativo y por tanto  $x = O(\mu)$ . El valor umbral  $\mu_{\text{err}} = \log f/L$  define la catástrofe de error. Cuando  $\mu < \mu_{\text{err}}$  la cuasiespecie está bien definida porque la secuencia maestra es la más probable, pero cuando  $\mu > \mu_{\text{err}}$  se pierde la identidad de la secuencia maestra entre la nube de mutantes y la cuasiespecie desaparece como tal.

Estudios experimentales realizados en los años 90 parecen confirmar [15] que la longitud del genoma de diversas especies (que van desde virus al *Homo sapiens*) y la tasa de mutación por base mantienen una relación  $\mu L \leq O(1)$ ; por ello, el incremento de la tasa de mutación se ha propuesto como un mecanismo para combatir las infecciones virales. Volveremos a este punto más adelante.

## 6.2. PAISAJE RUGOSO

Si bien localmente el paisaje puede tener el aspecto del Fujiyama, la forma en que Wright visualizaba un paisaje adaptativo es como una agreste cordillera llena de picos elevados separados por profundos valles. La razón es que es un hecho experimental que mutaciones que alteran poco la secuencia pueden producir grandes variaciones en la tasa de replicación de los individuos. Además, existe el conocido fenómeno de la *epistasia*, mediante el cual algunos genes interactúan constructiva o destructivamente, potenciando este efecto de variación grande por cambios pequeños.

De acuerdo con la metáfora del paisaje rugoso, las especies evolucionan subiendo los picos e instalándose en las cimas. Distintas cimas corresponden a distintas especies con distintos grados de adaptación. Este esquema de cosas, que parece encajar bien con nuestra imagen de la evolución, tiene un serio problema cuantitativo: las especies que ya se encuentran en una cima sólo pueden mejorar pasando a otra cima superior cercana cruzando un valle de inadaptación. En el caso más favorable el valle contiene un solo estado intermedio. La fórmula (31) nos dice que si la población es pequeña no es imposible que un alelo con menor grado de adaptación desplace a otro con grado de adaptación más alto; sin embargo, la probabilidad con que eso ocurre es baja, lo que implica tiempos de adaptación grandes. Y esto es sólo el caso más favorable. La velocidad con la que los virus son capaces de adaptarse a entornos rápidamente cambiantes echa por tierra este modelo evolutivo. ¿Qué falla entonces en nuestra imagen del paisaje adaptativo?

## 6.3. PAISAJE AGUJERADO: REDES NEUTRAS

Consideremos, para su análisis, el caso extremo de paisaje rugoso: el paisaje aleatorio. Dicho paisaje se genera asignando a cada secuencia de  $\mathcal{X}$  un valor del grado de adaptación aleatorio e independiente. En general, los paisajes rugosos no son tan extremos y hay cierta correlación entre los valores del grado de adaptación de secuencias vecinas; sin embargo, más allá de una longitud de correlación, dichos valores devienen independientes. El paisaje rugoso es el caso extremo en que dicha longitud de correlación es menor que 1. Supongamos ahora que la longitud de la

secuencia es grande, y que en cada posición hay  $A$  alelos independientes. El grado del grafo  $\mathcal{G}$  será  $g = (A - 1)L$  y su tamaño  $|\mathcal{X}| = A^L$ . Con  $L = 100$  y  $A = 2$  (valores relativamente pequeños),  $g = 100$  y  $|\mathcal{X}| = 2^{100} \approx 10^{30}$ . A todos los efectos, un grafo como éste se aproxima localmente a un árbol, tanto más cuanto mayor es su grado (véase la figura 2). Imaginemos un caso extremo de grado de adaptación que toma sólo dos valores: 1 si la secuencia es viable y 0 si no lo es. La fracción de secuencias viables se denota  $p$ . La evolución sólo puede proceder saltando entre nodos viables consecutivos. De acuerdo a la figura 2, donde se representa localmente este paisaje adaptativo en un grafo particular, es evidente que cuando  $p$  es pequeño el número de nodos viables a distancia  $d$  del nodo viable inicial resulta bien aproximado por un proceso de ramificación donde (excepto en la primera generación) el número de descendientes (nodos viables) está dado por  $p_k = \binom{g-1}{k} p^k (1-p)^{g-1-k}$ , cuyo valor esperado es  $(g-1)p$ . La teoría de procesos de ramificación [10] nos dice que, con probabilidad 1, el proceso continúa hasta infinito si  $(g-1)p > 1$ . Traducido a nuestros grafos lo que esto implica es que cuando  $p \gtrsim 1/g$  (siendo  $g \gg 1$ ), hay un subgrafo conexo de nodos viables que contiene una fracción finita de todos los nodos de  $\mathcal{G}$ . A este tipo de grafos se les da el nombre de *redes neutras* [7].

Si consideramos un modelo más general en el que  $\mathcal{P}(f)$  sea la densidad de probabilidad de que el grado de adaptación de un nodo esté entre  $f$  y  $f+df$ , si  $\int_{f_1}^{f_2} \mathcal{P}(f) df \gtrsim 1/g$  habrá una red cuasineutra en la que los valores del grado de adaptación de sus nodos estén en el intervalo  $(f_1, f_2)$ . Como los valores de  $g$  suelen ser muy grandes (proporcionales a la longitud de las secuencias), la existencia de redes neutras se convierte en la regla, más que en la excepción. La metáfora del paisaje de valles y montañas de Wright se vuelve totalmente inapropiada para describir la existencia de estas redes neutras en los paisajes adaptativos rugosos. Una metáfora más apropiada sería la de un paisaje plano (*á la Kimura*) con agujeros. La evolución mueve las secuencias a lo largo y ancho de la red neutra, de manera que éstas pueden transformarse por completo sin que ello haga decrecer su grado de adaptación. Sin duda este mecanismo acelera enormemente no sólo la adaptación de las especies al entorno, sino incluso la especiación.

El hecho de que el grado de adaptación dependa del fenotipo y no del genotipo favorece la aparición de redes neutras. Así se ve que ocurre en el ARN [6]. Las propiedades de estas redes tienen una enorme influencia en la dinámica evolutiva, una influencia que sólo ahora estamos empezando a entender. Sólo por dar un ejemplo, si reconsideramos el modelo de Eigen a la luz de esta nueva metáfora nos daremos cuenta de que la hipótesis principal, esto es, que localmente un paisaje adaptativo es de tipo Fujiyama, es rotundamente falsa. No hay una cadena con grado de adaptación máximo, sino toda una red neutra a la que pertenecen un número ingente de cadenas.

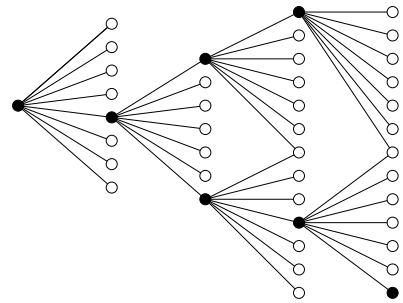


Figura 2: Sección local de un grafo correspondiente a  $A = 2$  y  $L = 8$ . Los nodos negros son viables; los blancos son inviables.

Esto hace que la probabilidad de recuperar el grado de adaptación máximo sea mucho mayor de la que la teoría de Eigen estima, porque se puede retornar a cualquier punto de la red neutra, no necesariamente a la cadena inicial. Cuando esa probabilidad es apreciable la catástrofe de error desaparece [13].

## 7. CONCLUSIÓN

En este artículo he intentado dar una panorámica de lo que la Matemática ha aportado a la teoría evolutiva. Desde la genética de poblaciones, pasando por los procesos estocásticos y acabando por la teoría de redes complejas, muchos resultados relevantes sobre los mecanismos evolutivos se han obtenido gracias a su descripción matemática. Aún no puede decirse que la Teoría de la Evolución sea una doctrina científica matemáticamente formulada en todos sus detalles, como a Darwin le habría gustado, pero es indudable que cada vez estamos más cerca de ello. Hoy en día podría decirse que los estudios teóricos de los procesos evolutivos son, al menos, tan importantes como los experimentales y que, como Darwin afirmaba, son ellos los que arrojan luz en la oscuridad.

Dado el carácter de este artículo y sus (y mis) limitaciones, en el camino me he dejado multitud de temas interesantes, tanto para dilucidar los mecanismos evolutivos como para ilustrar otras aportaciones matemáticas a la evolución. Tal es el caso del modelo de infinitos alelos [5] con el que hoy día se analizan las divergencias evolutivas de las cadenas de ADN o de las proteínas; o el proceso coalescente [5], una interesante y práctica reformulación retrospectiva de los procesos de deriva genética. Tampoco he hablado de las importantes aportaciones de la teoría de juegos a la evolución, con las cuales se está abordando actualmente, por ejemplo, el problema de la emergencia de la cooperación [15]. En lugar de todo ello he preferido destacar el tema de los paisajes adaptativos y las redes neutras porque, en mi opinión, es el área de trabajo que puede producir en los próximos años, sin cambiar su esencia, una reformulación más drástica del paradigma evolutivo. La comprensión de la evolución pasa, contra todo pronóstico, por comprender qué influencia tiene la deriva sobre redes neutras. Y debido a la complejidad del problema, es éste uno de los temas en que la Matemática puede ser nuestros ojos en la habitación oscura que nos ayuden a encontrar el gato que no está allí.

Quiero acabar agradeciendo a Susanna Manrubia su lectura crítica del manuscrito y sus valiosos comentarios, y a los proyectos MOSAICO y MOSSNOHO-CM el apoyo económico.

## REFERENCIAS

- [1] R. A. BLYTHE Y A. J. MCKANE, Stochastic Models of Evolution in Genetics, Ecology and Linguistics, *J. Stat. Mech.* **P07018** (2007), 1–58.
- [2] M. BULMER, Galton's law of ancestral heredity, *Heredity* **81** (1998), 579–585.
- [3] C. R. DARWIN, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, John Murray,

1859. Traducción al español: [http://es.wikisource.org/wiki/El\\_Origen\\_de\\_las\\_Especies](http://es.wikisource.org/wiki/El_Origen_de_las_Especies)
- [4] M. EIGEN, Selforganization of Matter and the Evolution of Biological Macromolecules, *Naturwissenschaften* **58** (1971), 465–523.
  - [5] W. J. EWENS, *Mathematical Population Genetics. I. Theoretical Introduction*, 2.<sup>a</sup> ed., Springer, 2004.
  - [6] W. FONTANA, Modelling ‘evo-devo’ with RNA, *BioEssays* **24** (2002), 1164–1177.
  - [7] S. GAVRILETS, *Fitness Landscapes and the Origin of Species*, Princeton University Press, 2004.
  - [8] J. HOFBAUER Y K. SIGMUND, *Evolutionary Games and Replicator Dynamics*, Cambridge University Press, 1998.
  - [9] S. KARLIN Y J. MCGREGOR, On a genetic model of Moran, *Proc. Cambridge Philos. Soc.* **58** (1962), 299–311.
  - [10] S. KARLIN Y H. M. TAYLOR, *A First Course in Stochastic Processes*, 2.<sup>a</sup> ed., Academic Press, 1975.
  - [11] S. KARLIN Y H. M. TAYLOR, *A Second Course in Stochastic Processes*, Academic Press, 1981.
  - [12] M. KIMURA, Evolutionary rate at the molecular level, *Nature* **217** (1968), 624–626.
  - [13] S. C. MANRUBIA, Biodiversidad, asexualidad y virus: tres esbozos sobre la complejidad de la evolución, *Apuntes de Ciencia y Tecnología* **32** (2009), 29–36.
  - [14] J. G. MENDEL, Versuche über Pflanzenhybriden, *Verhandlungen des naturforschenden Vereines in Brünn* **IV/1865** (1866), 3–47. Traducción al inglés: <http://www.esp.org/foundations/genetics/classical/gm-65.pdf>
  - [15] M. A. NOWAK, *Evolutionary Dynamics: Exploring the Equations of Life*, Belknap Press, 2006.
  - [16] C. M. REIDYS Y P. F. STADLER, Combinatorial landscapes, *SIAM Rev.* **44** (2002), 3–54.
  - [17] E. SENETA, *Non-negative Matrices and Markov Chains*, Springer, 2006.

JOSÉ A. CUESTA, GRUPO INTERDISCIPLINAR DE SISTEMAS COMPLEJOS (GISC), DPTO. DE MATEMÁTICAS, UNIVERSIDAD CARLOS III DE MADRID

Correo electrónico: [cuesta@math.uc3m.es](mailto:cuesta@math.uc3m.es)

Página web: <http://gisc.uc3m.es/~cuesta/>