



FACULTADE DE MATEMÁTICAS

Traballo Fin de Grao

# La topología de lo posible

Laura Fernández López

2019/2020

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



GRAO DE MATEMÁTICAS

Traballo Fin de Grao

# La topología de lo posible

Laura Fernández López

13/07/2020

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



# Trabajo propuesto

|   |
|---|
| <b>Área de Coñecemento: Xeometría e Topoloxía</b>   |
| <b>Título: La topología de lo posible</b>   |
| <b>Breve descripción do contido</b>   |
| En el artículo:<br>The Topology of the Possible: Formal Spaces Underlying Patterns of Evolutionary Change J. theor. Biol. (2001) 213, 241-274<br>los autores proponen, en el marco de la teoría de la evolución, dotar al espacio de fenotipos de una estructura topológica en lugar de la habitual estructura de espacio con una noción de distancia.<br>El trabajo consiste en la lectura crítica del artículo así como la revisión de los conceptos topológicos incluidos en el mismo. |
| <b>Recomendacións</b>   |
| Facilidad de lectura de textos científicos en inglés  |
| <b>Outras observacións</b>  |
|   |



# Índice general

|   |            |
|---|------------|
| <b>Resumen</b>  | <b>VII</b> |
| <b>Introducción</b>   | <b>IX</b>  |
| <b>1. Biología evolutiva</b>  | <b>1</b>   |
| 1.1. Contexto . . . . .   | 1          |
| 1.2. Accesibilidad . . . . .  | 2          |
| 1.3. Patrones evolutivos y accesibilidad fenotípica . . . . .             | 8          |
| <b>2. El concepto de pretopología</b>                                     | <b>11</b>  |
| 2.1. Contexto histórico . . . . .   | 11         |
| 2.2. Estructuras sobre un conjunto . . . . .                              | 12         |
| 2.2.1. Proximidad . . . . .   | 13         |
| 2.2.2. Los entornos . . . . .   | 14         |
| 2.2.3. Relaciones entre proximidad y entornos . . . . .                   | 15         |
| 2.3. Paso de pretopología a topología . . . . .                           | 16         |
| 2.4. Continuidad . . . . .  | 21         |
| 2.5. Comportamiento en conjuntos finitos . . . . .                        | 23         |
| <b>3. La combinación de las pretopologías y el mapa genotipo-fenotipo</b> | <b>25</b>  |
| 3.1. La pretopología de la accesibilidad . . . . .                        | 25         |
| 3.2. Sistema de entornos estadístico . . . . .                            | 29         |
| 3.3. La continuidad de las trayectorias evolutivas . . . . .              | 30         |
| 3.4. El caso del ARN . . . . .  | 32         |
| <b>Bibliografía</b>   | <b>37</b>  |



## Resumen

El objetivo de este trabajo consiste en, tras la lectura del artículo “The Topology of the Possible: Formal Spaces Underlying Patterns of Evolutionary Change J. theor. Biol. (2001)”, presentar las bases matemáticas necesarias para la construcción de una teoría genética evolutiva consistente, a través de la introducción del concepto de pretopología y sus propiedades.

Finalmente presentamos un ejemplo sencillo de esta implementación a través del comportamiento del ARN.

## Abstract

The aim of this dissertation is to establish a mathematic basis which is required to build a solid genetic theory through the introduction of the concept of pretopology and its properties, according to the article “The Topology of the Possible: Formal Spaces Underlying Patterns of Evolutionary Change J. theor. Biol. (2001)”.

Finally, RNA behaviour will be used as an easy example of our implementation.



# Introducción

La actual implementación del modelo neodarwiniano asume que el conjunto de posibles fenotipos está organizado de forma muy simétrica como un espacio regular equipado con una noción de distancia, como por ejemplo, un espacio vectorial euclidiano. Muchos estudios computacionales sobre el comportamiento del plegado de las secuencias de ARN en sus estructuras secundarias, caso que tomaremos como referencia a la hora de hablar del comportamiento entre genotipo y el fenotipo, sugieren otra cosa. Si los fenotipos están organizados en función de la accesibilidad genética, al espacio resultante carece de métrica y será formalizado por una estructura con la que no estamos familiarizados, la pretopología. Los patrones de evolución fenotípica surgen de forma natural en este espacio. Las consideraciones anteriores asociaban estos patrones exclusivamente a la selección natural. Proponemos como solución extender los fundamentos de accesibilidad hasta incluir la estructura topológica del espacio de fenotipos inducida por el esquema genotipo-fenotipo. Introducimos así el concepto matemático y las herramientas para formalizar la noción de pretopología de accesibilidad relativa con la que podemos hablar de continuidad en el esquema genotipo-fenotipo y en trayectorias evolutivas.

La intención de este texto es ofrecer una perspectiva evolutiva, un marco de trabajo utilizando los elementos matemáticos adecuados que implementen esta perspectiva, y un par de resultados que muestren como este entorno de trabajo se puede poner en práctica. El caso del ARN se usará como ejemplo a lo largo de todo el texto.



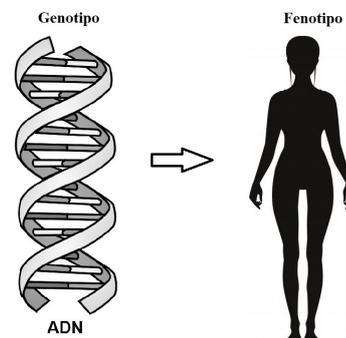
# Capítulo 1

## Biología evolutiva

### 1.1. Contexto

A lo largo de la historia hemos presenciado como el desarrollo progresivo de la teoría evolutiva iba parejo a descubrimientos en distintos campos, como la genética o las matemáticas. No será hasta mediados del siglo XIX cuando Mendel, mediante el análisis muy detallado de resultados experimentales, propondrá sus ya famosas leyes. A partir de ellas concluye un resultado totalmente contrario a las creencias de la época y defiende como los caracteres hereditarios tienen un carácter discreto y no continuo, como se pensaba hasta entonces. Será Fisher quien a principios del siglo XX combinará las Leyes de Mendel con la teoría de la selección natural de Charles Darwin, proporcionándole a este último una base matemática estable en la que sustentarse. Actualmente, uno de los modelos más importantes es el denominado neodarwinismo, que concibe la evolución como la combinación de mutaciones exitosas y los cambios producidos por la selección natural y la deriva genética en una población. Dos conceptos muy importantes en este contexto serán el genotipo y el fenotipo. El conjunto de todos los genes presentes en un organismo o aquellos que son heredables de padres a hijos es el denominado genotipo. Por otro lado, el fenotipo es la manifestación física de las características codificadas por algún genotipo. Uno de los problemas más importantes en el contexto evolutivo consiste en definir la relación entre estos conceptos.

Para ello, en primer lugar debemos situar el entorno que las relaciona, el denominado esquema genotipo-fenotipo. El desconocimiento durante gran parte de la historia de una teoría genética consistente trajo como conse-



cuencia una gran simplicidad a la hora de definir estas relaciones. Como consecuencia, habitualmente se supone que la accesibilidad de fenotipos a partir del genotipo es modelizable mediante un espacio métrico, en ocasiones hasta estructuras más fuertes, como puede ser un espacio vectorial euclidiano o grafos bastante complejos. Esto se ha convertido en un clásico a la hora de hablar de una estructura matemática dentro de la teoría evolutiva genética, suponiendo una clara fluidez en la organización fenotípica y el nivel genético. Muchos estudios acreditan cómo esta supuesta fluidez es básicamente ficticia y sugieren una acentuada asimetría en la accesibilidad entre los fenotipos y los genotipos.

Todo esto deja de manifiesto la necesidad de localizar una base matemática aplicable a la teoría evolutiva basada en espacios más simples. Uno de los estudios que ha propiciado esta necesidad de cambio, es el del plegado de moléculas de ARN, que nos servirá como ejemplo práctico a lo largo de todo este trabajo. De forma esquemática, el estudio del ARN refleja como afecta un cambio sobre una secuencia de ARN, conocida como estructura primaria, a la forma en la que se pliega esa secuencia, construcción que denotaremos como estructura secundaria. La forma en la que se pliega una proteína puede llegar hasta a afectar a su función, por lo que es un problema de gran relevancia. Un ejemplo de ello es el conocido “Mal de las vacas locas”, enfermedad causada por la mutación de una proteína la cual modificando su estructura secundaria y terciaria acababa convirtiéndose en un patógeno, es decir, un microorganismo capaz de producir daños al individuo que lo contiene. En ese caso, el contacto con esa proteína modificada inducía un cambio en cadena entre las proteínas restantes del individuo. La acumulación de esta mutación entre las células neuronales provocaría muerte celular y las ya conocidas consecuencias.

El problema al que nos enfrentamos entonces es cómo poder predecir el comportamiento de tales mutaciones, ya que un número pequeño de mutaciones en la secuencia genética puede tanto afectar a la estructura secundaria como puede pasar que un número mayor de mutaciones deje el plegamiento inalterado. Por lo tanto, a partir de una secuencia no podemos predecir con un modelo matemático cuantas mutaciones necesitamos introducir para cambiar su comportamiento. Debido a todo esto el concepto de entorno basado en una métrica es inútil en este contexto, lo que nos obligará a buscar con conceptos topológicos más débiles con los que trabajar como las pretopologías o las topologías.

Veremos estas ideas de forma más detallada a lo largo del capítulo.

## 1.2. Accesibilidad

Un concepto muy importante dentro de la biología evolutiva son las llamadas estructuras accesibles, también denominados espacios de configuración. El conjunto en el que se

sitúan es el compuesto por todas las posibles configuraciones junto con “operadores variacionales” que representan los procesos necesarios para trasladarnos de una configuración a otra. En el contexto genético, estos operadores serán las conocidas como mutaciones. Las mutaciones nos aportan una noción intuitiva de entornos ya que establecen qué elementos son accesibles a partir de otros. Por ejemplo, en el caso de ADN, diremos que los elementos más cercanos a una secuencia mediante mutaciones puntuales serán aquellas que difieran en un solo error de la secuencia original. En muchos casos, esos operadores variacionales nos proporcionan una noción intuitiva de “distancia” la cual nos permite llegar a trabajar hasta con estructuras de un espacio métrico.

**Definición 1.1.** Una métrica es una aplicación  $d$  que relaciona los elementos de un conjunto  $X$  con números reales positivos,  $d : X \times X \rightarrow \mathbb{R}_0^+$  que satisface los siguientes axiomas para todo  $x, y, z \in X$ :

1.  $d(x, x) = 0$ .
2.  $d(x, z) \leq d(x, y) + d(y, z)$ .
3.  $d(x, y) = d(y, x)$ .
4. Si  $d(x, y) = 0$  entonces  $x = y$ .

El conjunto  $(X, d)$  será lo conocido como espacio métrico.

**Ejemplo 1.2.** El conjunto de cadenas binarias de longitud  $n$  y aquellas que se pueden obtener entre ellas mediante mutaciones puntuales es un claro ejemplo del espacio métrico que comentábamos. Conectando cada secuencia con sus  $n$  posibles mutaciones obtenemos un grafo asociado al hipercubo. El hipercubo es un elemento topológico  $n$ -dimensional muy regular donde la distancia se corresponde con el número de posiciones en las que dos secuencias difieren (también conocida como distancia de Hamming).

Recordemos que una mutación puntual no es más que una mutación de una secuencia del genotipo a otra a partir de un cambio en una sola entrada de la secuencia.

En ese caso la distancia sería una métrica adecuada para definir la accesibilidad entre secuencias. De esta forma, podemos obtener cada elemento a partir de otro mediante una serie de mutaciones y no se genera ningún tipo de sesgo, es decir, será igual de sencillo obtener el elemento  $y$  a partir de  $x$  como obtener el  $x$  a partir de  $y$ . Esta simetría se asumen habitualmente para justificar ciertos comportamientos en el espacio, por lo que la única vía restante para la evolución sería la selección. El problema surge cuando las mutaciones que se producen en ese espacio contradicen estas suposiciones, haciendo imposible el modelarlas en este entorno.

En el caso del ARN, tanto el genotipo como el fenotipo son propiedades de una misma molécula. El plegado del ARN en su estructura secundaria (que llamaremos también forma) sugiere un esquema genotipo-fenotipo simple basado en su comportamiento, que es experimentalmente y computacionalmente factible. En poblaciones simuladas, el comportamiento del ARN muestra muchos de los fenómenos conocidos sobre la evolución de organismos. Por tanto, el ARN nos puede mostrar cuales de esos patrones de la evolución fenotípica se deben a las regularidades estadísticas del esquema genotipo-fenotipo.

Un aspecto importante del plegado del ARN es que las formas que pueden adoptar las secuencias de longitud  $n$  no aparecen con la misma frecuencia. Sólo una pequeña parte serían las consideradas como formas “típicas”, que presentan una frecuencia significativamente mayor. Como consecuencia, las referencias evolutivas (simuladas) presentan unas regularidades que podemos expresar en términos de propiedades estadísticas de las formas típicas.

Una de estas características estadísticas es de especial interés en nuestro contexto. Sea  $\alpha$  una posible forma de plegado, entonces por construcción,  $\alpha$  puede ser la forma típica de energía mínima correspondiente a secuencias. A la hora de organizar el espacio de formas se omiten los elementos que componen la secuencia y simplemente nos restringimos al esquema asociado a cada estructura, por lo que no es difícil que secuencias diferentes presenten un mismo esquema. En ese caso, el conjunto de secuencias que se pliegan de forma  $\alpha$  se denominan secuencias “neutras” con respecto a la estructura  $\alpha$ . Otra forma de verlo es considerar el espacio de formas como un espacio cociente, por lo que las secuencias neutras no son más que secuencias que pertenecen a una misma clase de equivalencia. En el caso de suceder mutaciones puntuales, si una mutación y su secuencia original presentan la misma forma, esa secuencia se conoce como un “vecino neutro”. Por “neutralidad” de una secuencia entendemos que la tercera parte de sus mutaciones puntuales son neutras. Cualquier secuencia que se pliegue en una forma típica tendrá una gran cantidad de vecinos neutros, y lo mismo les ocurrirá a esos vecinos. De forma más gráfica, podemos representar las distintas secuencias relacionadas por una mutación puntual que se pliegan en una misma forma como una cadena, que denominaremos “cadena neutra”. El hecho de modificar una secuencia y conservar el fenotipo original es un factor evolutivo clave. Por todo esto, la neutralidad determina la accesibilidad en el espacio de los fenotipos. Podemos observar gráficamente estas construcciones en la figura 1.1.

Para observar todo esto, veamos primero que entendemos por espacio de fenotipos en el caso del ARN. En primer lugar destacar que no estamos en un espacio, sino en el conjunto de las posibles formas de una secuencia de longitud  $n$ . Por lo tanto, para poder hablar de espacio debemos definir relaciones, en este caso de proximidad, entre

los distintos elementos. Un método posible sería establecer una comparación morfológica entre las distintas formas y a partir de ella obtener una noción de distancia. Así podríamos asumir un espacio métrico con el que trabajar. El problema de este razonamiento es que no representa la accesibilidad entre las formas, ya que los operadores variacionales bajo esta definición no se corresponden con los procesos físicos reales. En la práctica evolutiva, las formas se modifican a través de mutaciones en la secuencia original y la modificación viene dada por la aplicación que genera el plegado, el ya conocido esquema genotipo-fenotipo. Una relación de proximidad significativa en términos evolutivos entre las formas debe estar definida en base a la aplicación de plegado y no ser independiente a ella. En el caso del ARN, la aplicación del plegado traslada una gran cantidad de secuencias diferentes a una misma forma, por lo que para obtener una noción robusta de proximidad entre dos formas, esta debe reflejar la adyacencia de sus correspondientes cadenas neutras en el contexto de mutaciones en el espacio de genotipos.

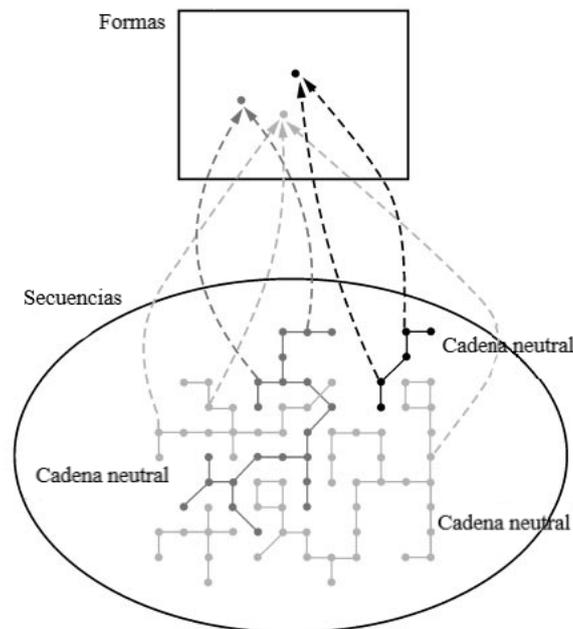


Figura 1.1: *Representación de la accesibilidad en el espacio de las formas. En esta representación del esquema que relaciona el genotipo (secuencias) con el fenotipo (formas), la distancia entre dos fenotipos viene determinada por la cantidad de frontera compartida entre sus cadenas neutras, situadas en el espacio de las secuencias. En este contexto, un paso desde la cadena más clara es propenso a llevarnos a la cadena de un tono gris intermedio. Independientemente, el proceso contrario es muy poco probable que suceda.*

*Notación 1.3.* Denotaremos por  $S(\alpha)$  al conjunto de las secuencias que se pliegan en forma  $\alpha$ . En ese caso,  $\partial S(\alpha)$  será el conjunto conocido como frontera, compuesto por las las secuencias que obtenemos a través una mutación puntual sobre las secuencias de  $S(\alpha)$ .

Más en particular, sean  $\alpha$  y  $\beta$  dos posibles formas, la proximidad de  $\beta$  a  $\alpha$  debe estar correlacionada con la probabilidad de una transición de  $\alpha$  a  $\beta$  a través de una mutación puntual. En el más sencillo de los casos, esta probabilidad viene dada por el cociente entre la frontera compartida entre los conjuntos neutros del genotipo de  $\beta$  y  $\alpha$ , y el tamaño total de la frontera del conjunto neutro de  $\alpha$ .

En todo caso,  $S(\beta) \cap \partial S(\alpha)$  denota las secuencias que se pliegan de forma  $\beta$  y las cuales son secuencias próximas a  $\alpha$ . La accesibilidad de llegar a  $\beta$  a partir de  $\alpha$ , que denotaremos por  $A(\beta \curvearrowright \alpha)$ , se entiende entonces como la frecuencia:

$$A(\beta \curvearrowright \alpha) = \frac{|S(\beta) \cap \partial S(\alpha)|}{|\partial S(\alpha)|}$$

donde  $|X|$  se corresponde con el cardinal del conjunto  $X$ .

$A(\beta \curvearrowright \alpha)$  no es una distancia y por lo tanto que el espacio que induce no será un espacio métrico. En general, este espacio también carece de simetría:  $A(\beta \curvearrowright \alpha) \neq A(\alpha \curvearrowright \beta)$ , ya que los conjuntos de elementos neutros, así como las fronteras, pueden diferir mucho en cuanto a tamaño.

**Ejemplo 1.4.** Para ilustrar este concepto pensemos por un momento en el mapa de comunidades de España. Supondremos la accesibilidad entre las comunidades autónomas en función del porcentaje de frontera que compartan. En esta topología, la comunidad de Castilla y León estaría próxima a Cantabria, ya que un paseo aleatorio desde Cantabria tiene muchas probabilidades de llevarnos hasta Castilla y León. Por otro lado, Cantabria no se consideraría próximo a Castilla y León, ya que en proporción a su frontera, la probabilidad de llegar a Cantabria desde Castilla y León es mucho menor.

Seleccionemos, por ejemplo, una forma  $\beta$  de entre las posibles estructuras del ARN que difiere de la forma  $\alpha$  en una pequeña acumulación situada en una región de la secuencia. Esta nueva región aporta un coste energético añadido al equilibrio anterior. Desencadenar una transformación desde  $\alpha$  hasta  $\beta$  requiere unas secuencias especialmente colocadas. Tales secuencias constituyen solo una pequeña parte del genotipo neutro de  $\alpha$ , sin embargo constituyen toda la frontera que  $\alpha$  comparte con  $\beta$ , por lo que es difícil llegar a  $\beta$  a partir de  $\alpha$ , o lo que es lo mismo en este lenguaje topológico,  $\beta$  no está próximo a  $\alpha$ .

Tomemos ahora una secuencia al azar del genotipo neutro de  $\beta$ . Mientras siga presente la acumulación anterior, es poco probable que esa secuencia esté energéticamente equilibrada. Una mutación puntual que destruya un par de bases estable provocará que la acumulación

se deshaga, obteniendo entonces la forma  $\alpha$ . Por lo tanto, muchas secuencias de la cadena de  $\beta$  limitan la cadena de  $\alpha$ , es decir,  $\alpha$  está cerca de  $\beta$  aunque como ya hemos visto,  $\beta$  no sea próximo a  $\alpha$ .

**Definición 1.5.** Decimos que una secuencia es compatible con una estructura si es capaz de plegarse de esa forma, aunque esta no sea (y habitualmente no lo es) la estructura de coste energético mínimo.

Las secuencias que se pliegan en forma  $\beta$  y que a la vez están en la frontera de la cadena neutra de  $\alpha$ , es decir las  $s \in S(\beta) \cap \partial S(\alpha)$ , son a menudo compatibles con las dos estructuras.

En términos de topología, una trayectoria evolutiva es “continua” en un punto del tiempo en el cual sucede un cambio en el fenotipo, si la mutación que lo desencadena es entre genotipos vecinos (en la topología del espacio genotipo) y el nuevo fenotipo está próximo al antiguo (en la topología del espacio fenotípico).

Como ya hemos visto, no tenemos una noción de distancia real en términos del fenotipo, por lo que debemos introducir una estructura nueva. A lo largo de este trabajo debatiremos como el sistema apropiado se corresponde con el inducido por el esquema genotipo-fenotipo el cual establece las probabilidades de transformar un fenotipo en otro mediante una operación genética. Esto debe ser diferenciado del enfoque en el que la continuidad está definidas por la presencia o ausencia de algún carácter discreto o incluso por un simple “salto” entre las posibilidades. En nuestro marco, un cambio en el fenotipo es discontinuo si constituye un “salto” desde una perspectiva evolutiva, es decir, si se puede realizar como un cambio genético continuo solo para una pequeña parte de los genotipos.

De nuevo, en el caso del ARN, las estructuras secundarias son elementos discretos y también los son sus mutaciones. Lo que determina la continuidad no es el grado en el que una modificación se obtiene de manera progresiva, sino el grado en el que una modificación es factible a partir de las relaciones establecidas entre el genotipo y el fenotipo. Del mismo modo, la discontinuidad y la magnitud de cambio no son siempre congruentes. El caso de las acumulaciones ilustra como los cambios morfológicamente grandes pueden ser continuos.

Por último, la asimetría del espacio de fenotipos puede provocar cambios evolutivos direccionales aún en ausencia de selección direccional. La diferencia de tamaño entre las cadenas neutrales grandes y pequeñas actúa como un mecanismo para la transición discontinua en el fenotipo que mantiene fijas las posibilidades. Si un fenotipo  $\alpha$  está próximo a  $\beta$ , pero  $\beta$  no está próximo a  $\alpha$  y que además,  $\alpha$  es considerablemente más grande que  $\beta$ , una transición desde  $\beta$  hasta  $\alpha$  es mucho más complicada de revertir ya que el punto desencadenante se perderá en la larga cadena de  $\alpha$ .

Siguiendo este razonamiento organizaremos estadísticamente el espacio de formas del ARN en términos de la variabilidad a través de mutaciones, observando como este no se corresponde con un espacio métrico.

### 1.3. Patrones evolutivos y accesibilidad fenotípica

La población recorre a lo largo de un proceso evolutivo una cadena neutra en el espacio de genotipos conservando así el mismo fenotipo  $\alpha$  hasta que se encuentra con una nueva cadena neutral asociada a un fenotipo  $\beta$  ventajoso. Si  $\beta$  no está próximo a  $\alpha$  en el sentido topológico descrito anteriormente, una población finita pasará un largo tiempo en la cadena asociada a  $\alpha$ . Recalquemos, de todas formas, que nuestra definición de “discontinuidad” no depende de algo repentino, el fenómeno de “conservar la estabilidad durante un largo periodo de tiempo que finalmente termina con un cambio en el fenotipo” es una manifestación de la intervención de la estructura topológica del espacio de fenotipos inducida por el esquema genotipo-fenotipo, y deja claro que no se necesitan sucesos externos para el cambio. Es por lo tanto tentador suponer que alguno de estos eventos vistos en la evolución fenotípica son consecuencia de transiciones discontinuas en el sentido del desarrollo. Al mismo tiempo, las transiciones graduales típicas del modelo neodarwiniano se corresponden con trayectorias evolutivas continuas conectando fenotipos próximos.

La accesibilidad se relaciona de forma directa con la idea de que las restricciones en el desarrollo acentúan las limitaciones de las variaciones fenotípicas factibles en los entornos del genotipo. Transformar un caracol en un caballo en un solo paso es, no solo un proceso discontinuo, sino que también imposible en el caso de que “paso” signifique un cambio genético continuo, es decir, uno que permanezca en el entorno del genotipo dado. De todas formas, tienen que existir caminos continuos en el espacio de fenotipos que conecten a un caracol con un caballo. La existencia o ausencia de estos caminos es una declaración sobre la accesibilidad de las estructuras en el espacio de fenotipos. La cuestión es importante, ya que esas rutas tienden a percibirse como trayectorias evolutivas decisivas.

Estos desarrollos nos aportan patrones bajo los que comparar individuos, en nuestro caso individuos dentro del espacio de genotipos. Diremos que dos características diferentes en dos especies son homologas si ambas han desarrollado a partir del mismo carácter de un ancestro común. La noción de homología implica la existencia de resultados alternativos en el fenotipo para un mismo problema evolutivo. Si solo tuviésemos una alternativa, entonces todos los organismos con la misma función serían estructuralmente idénticos, lo cual es evidente que no son. Por ejemplo, hay una diferencia clara entre las alas de un insecto y las alas de un pájaro, aunque ambas tengan la misma función. Al mismo tiempo,

las característica homologas pueden tener funciones radicalmente diferentes y conservar la misma estructura básica. La homología es un concepto sobre lo que permanece igual bajo la influencia de la selección natural la cual actúa para adaptar un carácter a las posibles funciones o situaciones. La homología es por tanto las dos cosas, una hipótesis sobre la existencia de un carácter ancestral común y la accesibilidad a los distintos caracteres a partir de mutaciones y selección. En nuestro contexto, existen diferentes homólogos en distintos dominios de accesibilidad y dos materializaciones del mismo homólogo son elementos del mismo dominio de accesibilidad.

La idea más informal de innovación evolutiva describe el hecho de que ciertos cambios en el fenotipos son difíciles de conseguir y parecen más importantes para la evolución de subsecuencias de un carácter que otros. El concepto está muy próximo a la homología en ese sentido, ya que reconocer “innovación” nos aportará una idea de lo que sería ser “lo mismo”. Una innovación se caracteriza como una transformación en el carácter del fenotipo que cambia radicalmente el conjunto de los posteriores fenotipos accesibles. Es tentador suponer que esta idea está relacionada con la noción de cambio discontinuo.

Por otra banda, existen muchos ejemplos de reversibilidad evolutiva, es decir, procesos evolutivos bidireccionales, aunque no todas las transiciones evolutivas son fácilmente reversibles. En la mayoría de los casos los cambios evolutivos irreversibles suponen una perdida de información genética, ya que es más fácil perder una parte funcional del genoma, y por lo tanto de su correspondiente fenotipo, que reconstruirlo mediante mutación. Incluso si la evolución como tal no es direccional, existen evidencias suficientes que sugieren que algunas transformaciones son extremadamente sesgadas. Un sentido de la transformación es muy fácil de conseguir, como puede ser la perdida de un elemento potenciador, pero el proceso inverso es muy poco probable que suceda. Por ello, alguna de las tendencias evolutivas direccionales se explican de forma más natural a partir de las asimetrías en las probabilidades de transición en lugar que a partir de una selección direccional. Este sesgo combinatorio se refleja en la asimetría presente en la accesibilidad entre fenotipos.

La definición de accesibilidad depende entonces de los operadores genéticos disponibles y del tipo de mutación. Anteriormente, ayudándonos del caso del ARN, definimos accesibilidad en términos de una mutación puntual. Esta elección viene dada por su simplicidad y aporta la precisión suficiente de la replicación en donde la reproducción de un genoma es poco probable que contenga dos o más mutaciones con respecto al original. En general, la definición de accesibilidad debe reflejar el tipo de mutación y los cambios genéticos que son relevantes en el contexto dado.

Debemos destacar la diferencia entre la topología estática del espacio de fenotipos inducida por el esquema genotipo-fenotipo bajo operadores genéticos y las consecuencias de

esa topología en las dinámicas evolutivas. Las dimensiones de estas consecuencias dependen de diferentes factores, como por ejemplo el tamaño poblacional. En una población infinita, la evolución es esencialmente un problema cinemático en el cual la cuestión de la innovación no figura. Dos cadenas neutrales que solo comparten una pequeña fracción de su frontera sufren las transiciones fenotípicas de formas diferentes en poblaciones finitas e infinitas. Esta contribución solo interviene en el desarrollo de un lenguaje topológico para los aspectos estáticos del espacio de fenotipos, motivado por nuestra creencia de que esta topología afecta profundamente a la evolución de poblaciones finitas.

Los tipos de mutaciones, los operadores mutacionales, el entorno y el esquema genotipos-fenotipos en sí mismo son materia evolutiva y del mismo modo lo es también la topología del espacio de fenotipos.

Habiendo expuesto los argumentos por los cuales es deseable, si no necesario, introducir nuevas estructuras accesibles no métricas en términos de la teoría evolutiva matemática, procedemos a adentrarnos en un contexto matemático en el concepto principal de este trabajo, como es el de la pretopología.

## Capítulo 2

# El concepto de pretopología

### 2.1. Contexto histórico

Como se desarrolló a lo largo del capítulo anterior, en el espacio genotipo-fenotipo la topología no resulta la estructura más adecuada con respecto a las propiedades de este espacio. Surge así la necesidad de una nueva estructura menos rica que se adapte mejor a sus condiciones. Otros problemas similares surgen en campos como la economía o la biología, lo que nos llevará a la búsqueda desde diferentes ramas del concepto que denominaremos como pretopología.

Para entender como surge el concepto de pretopología debemos remontarnos al origen y proceso de desarrollo de lo que actualmente conocemos como topología. Nos situamos a finales del siglo XIX principios del XX, aunque surgen indicios y conceptos predecesores mucho antes. Uno de los primeros indicios es la resolución del conocido como “El problema de los siete puentes de Königsberg” de Euler en 1735, el cual prioriza la relación y posición entre los puentes antes que la distancia real, y que se considera el inicio de la teoría de grafos. Una de las primeras apariciones del concepto de topología sería la dada por el alemán Johann Listing a través de la publicación de “Vorstudien zur Topologie,” en 1847.

Centrándonos en la topología general del momento, lo que se buscaba principalmente era el traslado de conceptos ya conocidos del análisis, como la continuidad, a situaciones más generales. Esta es la idea intuitiva en la que se basa la topología, la definición de una estructura en la que no intervengan relaciones u operaciones entre sus elementos sino los propios subconjuntos de nuestro conjunto. En línea con el estudio de la convergencia e implementando una mejora de la sintaxis matemáticas usada hasta el momento, Fréchet presenta en 1905 una serie de axiomas para modelar la convergencia en un conjunto abstracto y también algunos axiomas para establecer un espacio métrico. Finalmente será Hausdorff quien en 1914 relaciona mediante una serie de axiomas el comportamiento de

entornos, la métrica y los límites de sucesiones sobre espacios generales sentando las bases para el concepto de topología. Su definición se corresponderá con lo que actualmente denotamos como “espacios Hausdorff”. La concepción actual de la topología viene dada por Kuratowski quien en 1922 formaliza la idea de clausura la cual nos permite expresar la proximidad entre un punto y un conjunto.

Aunque entraremos más en detalle a lo largo del capítulo, el éxito de esta nueva construcción en los diferentes campos se basa en que nos permite el estudio de conjuntos finitos sin la necesidad de esa proximidad explícita.

## 2.2. Estructuras sobre un conjunto

Entendemos espacio como un conjunto al que dotamos de cualquier tipo de estructura. En términos de espacios uno de los más habituales e intuitivos sería el espacio vectorial  $\mathbb{R}^n$  o el  $n$ -espacio Euclidiano.  $\mathbb{R}^n$  es el producto de  $n$  copias de la recta real y está compuesto de forma que:

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in \mathbb{R}\}$$

En este caso podemos entender cada  $x \in \mathbb{R}^n$  como un punto del  $n$ -espacio, pero lo podemos ver también como un vector con respecto al origen  $\mathcal{O} = (0, 0, \dots, 0)$ . En cualquier caso la **norma** de  $x$  se define como  $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$ . Podemos hablar de ella como la longitud de  $x$  como vector o como la distancia entre  $\mathcal{O}$  y el punto  $x$ . Esta última concepción se puede trasladar a cualesquiera dos puntos del espacio de la siguiente forma:

Sean  $p = (p_1, p_2, \dots, p_n)$  y  $q = (q_1, q_2, \dots, q_n)$  dos puntos de nuestro espacio, la distancia entre  $p$  y  $q$  viene dada de la forma:

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + \dots + (p_n - q_n)^2}$$

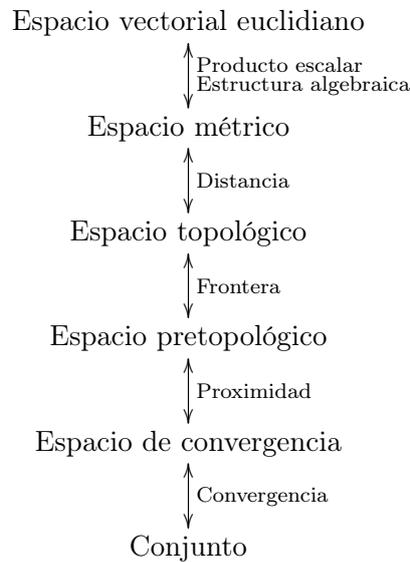
Esta aplicación verifica los axiomas necesarios para ser una métrica, concepto que ya introducimos en el capítulo anterior a través de la definición 1.1.

**Ejemplo 2.1.** Las formas de las moléculas, como pueden ser las proteínas o el propio ARN, o las secuencias de genes son ejemplos de elementos que en la práctica constituyen un espacio métrico: Las formas o secuencias no pueden ser sumadas pero su distancia (o similitud) puede ser cuantificada.

No en todos los casos es interesante o útil definir una métrica o distancia. En ese caso nos restringiremos a conceptos más básicos que nos permitan relacionar dos elementos, como pueden ser los abiertos en el caso de las topologías, los entornos o las relaciones de proximidad. En el caso de no poder definir una métrica, el concepto de proximidad no

será cuantificable: dos elementos pueden estar próximos a un tercero pero no podremos identificar cual de ellos estaría más próximo.

Nuestro interés se centrará a lo largo de este capítulo en los casos en los que no es posible obtener esa noción de distancia. En esos casos, los subconjuntos nos permitirán organizar el espacio, como es el caso de las topologías con sus correspondientes abiertos. Una noción menos fuerte, sería restringirnos únicamente a la proximidad entre elementos, obteniendo la conocida como pretopología. Podemos por último eliminar el concepto de entorno y conservar una estructura suficientemente estable como para poder hablar de convergencia. Este orden jerárquico queda resumido en el siguiente esquema:



En el contexto biológico en el que se sitúa este trabajo, la estructura en la que nos centraremos será la pretopología y los elementos básicos de esta.

### 2.2.1. Proximidad

En el momento en que no disponemos de una medida de distancia clara, podemos establecer una relación de proximidad entre nuestros elementos.

**Definición 2.2.** Una relación de proximidad  $U$  en  $X$  no es más que una relación binaria  $U \subset X \times X$  tal que  $\Delta \subset U$ . Diremos que  $y$  está próxima a  $x$  si y sólo si  $(x, y) \in U$ .

$$U = \{(x, y) \mid \text{“} y \text{ es } U\text{-próxima a } x\text{”}\}$$

De forma general,  $U$  no tiene por qué ser simétrico, es decir:  $(x, y) \in U \not\Rightarrow (y, x) \in U$ .

Por otro lado, es trivial que  $U \neq \emptyset$ , ya que  $(x, x) \in U$  para todo  $x \in X$  (un punto  $x$  siempre estará próximo a sí mismo). En el caso de tener dos relaciones de proximidad  $U$

y  $U'$  sobre el mismo conjunto  $X$ , diremos que es  $U$  más restrictivo que  $U'$  si  $U \subset U'$ . Así,  $U$  expresa en una escala más fina la proximidad en comparación con  $U'$  y de esta forma conseguimos hablar de grados de proximidad, a pesar de que esta no venga dada de forma cuantificable.

Sean entonces  $x, y \in X$  dos elementos tales que  $y$  es  $U$ -próximo a  $x$  y también es  $U'$ -próximo a  $x$ , entonces  $y$  es  $(U \cap U')$ -próximo a  $x$ .

*Observación 2.3.* Nótese que  $(U \cap U')$  es un conjunto no vacío ya que  $(x, x)$  está contenido en cada uno de ellos para todo  $x \in X$  y por tanto estará en la intersección.

Sean  $x, y \in X$  tales que  $y$  es  $U$ -próximo a  $x$ , entonces  $y$  será  $U'$ -próximo a  $x$  para toda relación  $U'$  menos restrictiva.

**Definición 2.4.** Una colección  $\mathcal{U}$  de relaciones  $U$  en  $X$  que verifican los siguientes axiomas:

(U1) Sea  $\Delta = \{(x, x) | x \in X\}$  la conocida como **diagonal**, entonces  $\Delta \subseteq U$  para todo  $U \in \mathcal{U}$ .

(U2) Sean  $U, U' \in \mathcal{U}$ , entonces  $(U \cap U') \in \mathcal{U}$ .

(U3) Sea  $U \in \mathcal{U}$  y  $U \subset U'$ , entonces  $U' \in \mathcal{U}$ .

se denomina **preuniformidad**.

### 2.2.2. Los entornos

Otra forma de establecer la proximidad de los elementos en un conjunto  $X$  es a través de los conocidos como entornos. La idea consiste en definir un conjunto de elementos como próximos, en lugar de trabajar con pares que verifican alguna condición.

**Definición 2.5.** Una aplicación  $\mathcal{N} : X \rightarrow \mathcal{P}(\mathcal{P}(X))$  tal que:

(N1)  $x \in N, \forall N \in \mathcal{N}(x)$ .

(N2) Sean  $N_1, N_2 \in \mathcal{N}(x)$ , entonces  $N_1 \cap N_2 \in \mathcal{N}(x)$ .

(N3) Sea  $N_1 \in \mathcal{N}(x)$  y  $N_1 \subset N$ , entonces  $N \in \mathcal{N}(x)$ .

se denominará **conjunto o sistema de entornos**.  $\mathcal{N}(x)$  será el conjunto de entornos del punto  $x$ . El par  $(X, \mathcal{N})$  es el denominado **espacio pretopológico**, concepto en el que profundizaremos más adelante.

Por construcción podríamos suponer que este sistema de entornos define una topología, lo cual no es cierto ya que, al contrario que en una topología, aquí para todo punto  $x$  de un conjunto  $N$ ,  $N$  no tiene por qué ser un entorno de  $x$ . Esta será una de las condiciones que usaremos más adelante para relacionar el concepto de pretopología con las ya conocidas topologías.

Hablaremos de una **base de entornos** si y sólo si se verifican (N1) y una versión más débil de la propiedad (N2), que denotaremos como (N2') y que establece:

- Sean  $N_1, N_2 \in \mathcal{N}(x)$ , entonces existe un  $N_3 \in \mathcal{N}(x)$  tal que  $N_3 \subseteq N_1 \cap N_2$

Esta condición es análoga a las que les exigimos a una familia de subconjuntos para ser base de topología. De nuevo, la diferencia radica en el tipo de subconjuntos que seleccionamos.

### 2.2.3. Relaciones entre proximidad y entornos

Debido a estas similitudes, es sencillo establecer una relación entre las preuniformidades y los sistemas de entornos  $\mathcal{N}$  de una forma intuitiva.

Dado  $X$  con una preuniformidad  $\mathcal{U}$ , para cada  $x \in X$  definimos su sistema de entornos  $\mathcal{N}_{\mathcal{U}}(x)$  como aquel compuesto por los conjuntos de la forma:

$$U[x] = \{y \in X \mid (x, y) \in U\} \quad \text{para todo } U \in \mathcal{U}.$$

Es sencillo comprobar como los conjuntos  $U[x]$  satisfacen las condiciones asociadas a los entornos, definiendo entonces un sistema de entornos.

- (N1)  $(x, x) \in U$  por definición de preuniformidad, por lo tanto  $x \in U[x]$ .
- (N2) Sean  $U, U' \in \mathcal{U}$ , en virtud de (U2) tenemos  $U \cap U' \in \mathcal{U}$ . Por lo tanto, siendo  $U[x]$  y  $U'[x]$  dos entornos del punto  $x$ ,  $U[x] \cap U'[x] = (U \cap U')[x]$  es también un entorno de  $x$ .
- (N3) Sea  $U_1 \in \mathcal{U}$  y  $U$  tal que  $U_1 \subset U$  entonces gracias a (U3) sabemos que  $U \in \mathcal{U}$ . Luego si tenemos  $U_1[x]$  entorno de  $x$  y tal que  $U_1[x] \subset U[x]$  entonces  $U[x]$  es un entorno de  $x$ .

Por tanto, el sistema de entornos inducido por la preuniformidad  $\mathcal{U}$  será de la forma  $\mathcal{N}_{\mathcal{U}}(x) = \{U[x] \mid U \in \mathcal{U}\}$ . De manera análoga, dado un sistema de entornos  $\mathcal{N}$  podemos construir la correspondiente preuniformidad como una colección  $\mathcal{U}_{\mathcal{N}}$  cuyos elementos  $U$  se construyen, dado cierto  $N_x \in \mathcal{N}(x)$ , como:

$$U_{N_x} = \{(x, y) \mid x \in X, y \in N_x\}.$$

La construcción de  $U$  puede entenderse de manera más informal de la siguiente forma: para cada  $x \in X$  seleccionamos alguno de sus entornos. En este entorno establecemos todos sus elementos como próximos a  $x$ . Los entornos ya seleccionados se eliminan de  $\mathcal{N}$  y el proceso se repite hasta que todos los entornos hayan sido asociados.

Una preuniformidad  $\mathcal{U}$  y la pretopología que induce son formas muy parecidas de organizar el conjunto  $X$ , veámoslo a través del siguiente resultado:

**Lema 2.6.** *La pretopología  $\mathcal{N}_{\mathcal{U}, \mathcal{N}}$  inducida por la preuniformidad  $\mathcal{U}_{\mathcal{N}}$  de  $\mathcal{N}$  coincide con la pretopología  $\mathcal{N}$ .*

*Demostración.* Los elementos de  $\mathcal{U}_{\mathcal{N}}$  son de la forma  $U' = \{(x, y) | x \in X, y \in N_x\}$ . Cualquier elemento  $N_x \in \mathcal{N}(x)$  lo podemos escribir de la forma  $N_x = \{(x, y) | y \in N_x\}[x]$ , construcción que coincide con  $U'[x]$ . En ese caso, si un conjunto  $U \in \mathcal{U}_{\mathcal{N}}$  contiene a  $U'$ , entonces  $U[x]$  contiene a algún  $N_x \in \mathcal{N}(x)$  y por lo tanto es un entorno de  $x$ . □

El caso contrario no se cumple de forma general ya que la relación entre preuniformidades y pretopologías de  $X$  no es una correspondencia directa. En general, podemos obtener una misma pretopología a partir de diferentes preuniformidades.

### 2.3. Paso de pretopología a topología

**Definición 2.7.** Sean  $U'$  y  $U''$  dos relaciones de proximidad sobre el conjunto  $X$ , definimos la concatenación de relaciones como:

$$U' \circ U'' = \{(x, y) | \exists z \in X : (x, z) \in U' \quad y \quad (z, y) \in U''\}$$

**Proposición 2.8.** *La concatenación de dos relaciones de proximidad establece una relación de proximidad.*

$U' \circ U''$  contiene tanto a  $U'$  como a  $U''$  ya que la diagonal  $\Delta = \{(x, x) | x \in X\}$  está contenida en virtud del axioma (U1). La concatenación nos permite introducir un grado de proximidad más débil, ya que los elementos de  $U'$  y  $U''$  están próximos en una escala más fina que los elementos de  $U' \circ U''$ . Podemos entender que el elemento  $z$  de la definición es un elemento “entre”  $x$  e  $y$ .

**Definición 2.9.** Sea  $\mathcal{U}$  una preuniformidad tal que para cada  $U \in \mathcal{U}$  existe un  $V \in \mathcal{U}$  con  $V \circ V \subseteq U$ , entonces  $\mathcal{U}$  se denomina **quasiuniformidad**.

La diferencia fundamental entre ellas consiste en que dados dos elementos  $x$  e  $y$ , en el caso de una quasiuniformidad podemos suponer que existe siempre un elemento  $z$  entre ellos. A partir de las relaciones que ya conocemos podemos extraer el siguiente resultado en materia de entornos:

**Proposición 2.10.** *Sea  $\mathcal{N}$  un sistema de entornos asociado a una quasiuniformidad, entonces para cada  $N \in \mathcal{N}(x)$ , existe un  $N' \in \mathcal{N}(x)$  tal que  $N \in \mathcal{N}(y)$  para todo  $y \in N'$ .*

*Demostración.* De acuerdo con la definición de quasiuniformidad, para cada  $U[x] \in \mathcal{U}[x]$  existe un  $V[x] \in \mathcal{U}[x]$  tal que  $(V \circ V)[x] \subseteq U[x]$ . Como consecuencia de la construcción tenemos que  $V \subseteq V \circ V$  y por tanto  $V[x] \subseteq (V \circ V)[x] \subseteq U[x]$ . De manera más explícita:

$$(V \circ V)[x] = \{y | \exists z \in V[x] \text{ e } y \in V[z]\} = \bigcup_{z \in V[x]} V[z]$$

Si ahora denotamos como  $N = U[x]$ ,  $N' = V[x]$ ,  $N'' = (V \circ V)[x]$  en la igualdad anterior,  $N''$  coincide con la unión de los entornos de los elementos de  $N'$ . Como  $N''$  contiene entornos de cada uno de los elementos de  $N'$ , entonces en base al axioma (N3), este constituye un entorno para cada  $y \in N'$ , y por lo tanto  $N'' \in \mathcal{N}(y)$ . Por otro lado sabemos que  $N'' \subseteq N$  y entonces  $N \in \mathcal{N}(y)$  para todo  $y \in N'$ , como queríamos demostrar. □

Debido a esto, podemos definir un nuevo axioma en materia de entornos

(N4) Para cada  $N \in \mathcal{N}(x)$  existe un  $N' \in \mathcal{N}(x)$  tal que  $N \in \mathcal{N}(y)$  .

**Proposición 2.11.** *Una pretopología que verifica (N4) es una **topología**.*

*Demostración.* Un resultado equivalente a nuestra proposición, es comprobar como la colección

$$\mathcal{T} = \{U \subset X | U \text{ es un entorno de cada uno de sus puntos}\} = \{U \subset X | U \in \mathcal{N}(x), \forall x \in U\}$$

es una topología.

En primer lugar,  $\emptyset \in \mathcal{T}$  de manera trivial .

$X \in \mathcal{T}$ , ya que para cualquier  $x \in X$  existe algún  $N$  tal que  $N \in \mathcal{N}(x)$  y en particular  $N \subset X$ , por lo que en base al axioma (N3) se tiene que  $X \in \mathcal{T}$ .

En el caso de la unión arbitraria, supongamos una familia de entornos  $\{U_i\}_{i \in \Omega}$  tal que  $U_i \in \mathcal{T}, \forall i \in \Omega$ . ¿  $\bigcup_{i \in \Omega} U_i \in \mathcal{T}$ ?

Tomando un  $i \in \Omega$  arbitrario y aplicando de nuevo el axioma (N3) tenemos que dado  $U_i \subset \bigcup_{i \in \Omega} U_i$  implica que  $\bigcup_{i \in \Omega} U_i \in \mathcal{N}(x)$  para todo  $x \in U_i$ . Por tanto  $\bigcup_{i \in \Omega} U_i$  será un entorno de cada uno de los puntos de todos sus elemento, es decir,  $\bigcup_{i \in \Omega} U_i \in \mathcal{T}$ .

Por último, comprobaremos la intersección finita en el caso de dos elementos. Sean  $U_1$  y  $U_2$  dos elementos de  $\mathcal{T}$ , ¿ $U_1 \cap U_2 \in \mathcal{T}$ ?

Como  $U_1, U_2 \in \mathcal{T}$ , por definición sabemos que  $U_1 \in \mathcal{N}(x)$  para todo  $x \in U_1$  y  $U_2 \in \mathcal{N}(y)$  para todo  $y \in U_2$ . Tenemos entonces dos posibilidades:

- Si  $U_1 \cap U_2 = \emptyset$  entonces en base a lo que ya hemos demostrado, podemos decir que  $U_1 \cap U_2 \in \mathcal{T}$
- En caso contrario, ¿ $U_1 \cap U_2 \in \mathcal{N}(z)$  para todo  $z \in U_1 \cap U_2$ ?

Sea  $z \in U_1 \cap U_2$ , entonces  $z \in U_1$  y  $z \in U_2$ , por lo que  $U_1, U_2 \in \mathcal{N}(z)$  y en base al axioma (N2) tenemos que  $U_1 \cap U_2 \in \mathcal{N}(z)$  para todo  $z \in U_1 \cap U_2$ .

Por lo que podemos concluir que  $\mathcal{T}$  es una topología. □

La diferencia entre una pretopología y una topología radica en la construcción de la frontera y por tanto en el comportamiento del operador clausura. Aunque en este caso hemos optado por definir el concepto de pretopología a partir del concepto de entorno, existen otras formas para definirlo a partir de la frontera o del operador clausura. Veamos a continuación como podemos llegar a los anteriores a partir de los entornos.

**Definición 2.12.** Dado un sistema de entornos sobre un conjunto  $X$  y un subconjunto  $A \subseteq X$ , decimos que  $x \in X$  es un **elemento de frontera** de  $A$  si todos los entornos de  $x$  se intersecan de forma no vacía tanto con  $A$  como con su complementario  $X \setminus A$ .

La **frontera** de  $A$ , que denotamos por  $\partial A$ , es la colección de todos los elementos de frontera de  $A$ :

$$\partial A = \{x \in X \mid \forall N \in \mathcal{N}(x) : N \cap A \neq \emptyset \text{ y } N \cap (X \setminus A) \neq \emptyset\} \quad (2.1)$$

A partir de ella podemos definir tanto el **interior** como la **clausura** de un conjunto.

$$\overset{\circ}{A} = A \setminus \partial A \quad \bar{A} = A \cup \partial A$$

**Definición 2.13.** Diremos que un conjunto  $A$  es **abierto** si no contiene a ningún elemento de su frontera. Otra forma,  $A$  es abierto si contiene un entorno para cada uno de sus elementos, es decir, para cada  $x \in A$  existe un entorno  $N \in \mathcal{N}(x)$  tal que  $N \subseteq A$ .

Los abiertos juegan un papel importante dentro de la topología, ya que la colección de todos los conjuntos abiertos que contienen a  $x$ ,  $\mathcal{T}(x) = \{A \mid A \text{ es abierto y } x \in A\}$

constituye una base de entornos abiertos de  $x$ . Por otro lado, si un sistema de entornos  $\mathcal{N}(x)$  satisface el axioma (N4) entonces  $\mathcal{T}(x)$  es una base de  $\mathcal{N}(x)$  y viceversa.

Si nos centramos en las diferencias entre topología y pretopología, el aspecto principal es el comportamiento del operador clausura, que definimos sobre un conjunto  $A$  como  $\bar{A} = A \cup \partial A$ . Para simplificar la notación asumiremos  $\mathcal{T}(x)$  como la base de abiertos de la topología. En este espacio, obtenemos la clausura de  $A$  añadiéndole todos los elementos de su frontera. ¿Ocurrirá algo más si repetimos este procedimiento sobre el conjunto obtenido? Intuitivamente, una vez añadimos los puntos de frontera, no existen más elementos y entonces no ocurrirá nada.

Por el contrario, si aplicamos el operador clausura sobre alguno de los entornos de una pretopología no sucederá lo mismo. Sea  $A$  uno de los entornos de la pretopología  $\mathcal{N}$  y sea  $x \in \partial A$ , un elemento de la frontera de  $A$ . Supongamos entonces que existe un elemento  $y$  fuera de  $A$  tal que cualquier de entorno  $N_y$  contiene al punto  $x$ . Si aplicamos el operador clausura sobre un conjunto  $A$  obtenemos un nuevo conjunto de la forma  $\bar{A} = A \cup \partial A$ . Debido a que ahora  $x \in \bar{A}$ ,  $y$  pasa a ser un elemento de frontera de  $\bar{A}$ . Esto no puede ocurrir en el caso de la topología, ya que si  $x$  está contenido en todos los entornos  $N_y$  entonces cada  $N_y$  contiene al menos a un entorno de  $x$ , ya que los subconjuntos  $N_y$  son abiertos. Debido a esta construcción, todo  $N_y$  tendrá intersección no vacía con  $A$  y por lo tanto  $y$  sería también elemento de frontera de  $A$ . De esta manera todos los elementos de la frontera se habrían añadido a  $A$  y no se generarían nuevos elementos de frontera.

En el caso de la topología, el operador frontera de un conjunto verifica que  $\partial\partial A \subseteq \partial A$  propiedad que es equivalente a la idempotencia  $\overline{\bar{A}} = \bar{A}$  en términos del operador clausura.

Podemos establecer la definición (2.1) en el términos de la pretopología, pero su comportamiento no es tan intuitivo como puede serlo en una topología. Los espacios pretopológicos pueden definirse de muchas formas aparentemente no relacionadas, pero equivalentes entre ellas. Anteriormente nos hemos centrado en los entornos como concepto base para nuestra definición. En la siguiente tabla, resumimos las equivalencias entre los conceptos de entorno y los operadores principales tomando en cada caso uno como base para las demás definiciones.

|                      | Entornos  | Clausura                                  | Interior   | Frontera                         |
|----------------------|---|---|--|----------------------------------|
| $\mathcal{N}(x)$     | $=$   | $\{A x \notin \overline{X \setminus A}\}$ | $=$  | $\{A x \in \overset{\circ}{A}\}$ |
| $\bar{A}$            | $= \{x \forall N \in \mathcal{N}(x) : A \cap N \neq \emptyset\}$  | $=$                                       | $X \setminus \overline{(X \setminus A)}$                             | $= A \cup \partial A$            |
| $\overset{\circ}{A}$ | $= \{x \exists N \in \mathcal{N}(x) : N \subseteq A\}$  | $= X \setminus \overline{X \setminus A}$  | $=$  | $A \setminus \partial A$         |
| $\partial A$         | $= \{x \forall N \in \mathcal{N}(x) : A \cap N \neq \emptyset, (X \setminus A) \cap N \neq \emptyset\}$ | $= \bar{A} \cap \overline{X \setminus A}$ | $= X \setminus (\overset{\circ}{A} \cup \overline{(X \setminus A)})$ |                                  |

A partir de estas equivalencias, podemos reescribir las condiciones necesarias para cons-

tituir una pretopología en función de cualquier operador. Las condiciones necesarias están recogidas en la siguiente tabla, siendo la cuarta fila la condición necesaria para poder constituir una topología.

|   | Entornos  | Clausura   | Interior   | Frontera   |
|---|---|--|--|--|
| 1 | $\forall N \in \mathcal{N}(x) : x \in N$  | $\overline{\emptyset} = \emptyset$                     | $\overset{\circ}{\emptyset} = \emptyset$                                   | $\partial\emptyset = \emptyset$  |
| 2 | $N \in \mathcal{N}(x), N \subset N'$<br>$\Rightarrow N' \in \mathcal{N}(x)$   | $A \subseteq \overline{A}$                             | $\overset{\circ}{A} \subset A$   | $\partial A = \partial(X \setminus A)$   |
| 3 | $N, N' \in \mathcal{N}(x)$<br>$\Rightarrow N \cap N' \in \mathcal{N}(x)$  | $\overline{A \cap B} = \overline{A} \cap \overline{B}$ | $\overset{\circ}{(A \cap B)} = \overset{\circ}{A} \cap \overset{\circ}{B}$ | $A \cap B \cap \partial(A \cap B) =$<br>$A \cap B \cap (\partial A \cap \partial B)$ |
| T | $N \in \mathcal{N}(x) \Rightarrow \exists N' \in \mathcal{N}(x) :$<br>$\forall y \in N' : \exists N_y \in \mathcal{N}(y) : N_y \subseteq N$ | $\overline{\overline{A}} = \overline{A}$               | $\overset{\circ}{\overset{\circ}{A}} = \overset{\circ}{A}$                 | $\partial\partial A \subseteq \partial A$  |

Por último, obtendremos la simetría tras exigir el siguiente requisito:

(US) Sea  $U \in \mathcal{U}$  entonces se tiene que  $U^{-1} = \{(x, y) | (y, x) \in U\} \in \mathcal{U}$ .

Como consecuencia las relaciones simétricas ( $U \cap U^{-1}$ ) serán relaciones de proximidad en términos de preuniformidades. Cuando una preuniformidad cumpla (US), hablaremos de una **semiuniformidad**. Si una quasiuniformidad verifica la condición de simetría entonces pasa a denominarse **uniformidad**.

Trasladando esta condición de simetría al contexto de los entornos obtenemos dos condiciones equivalentes entre si, que modelan la simetría en el caso de pretopologías.

**Teorema 2.14.** *Sea  $(X, \mathcal{N})$  es un espacio pretopológico, entonces*

(1)  $x \in \overline{\{y\}}$  implica que  $y \in \overline{\{x\}}$  para todo  $x, y \in X$ .

(2)  $x \in \bigcap \mathcal{N}(y)$  implica que  $y \in \bigcap \mathcal{N}(x)$ .

son equivalentes.

*Demostración.* Para empezar, observamos que:

$$\overline{\{y\}} = \{z | \forall N \in \mathcal{N}(z) : \{y\} \cap N \neq \emptyset\} = \{z | \forall N \in \mathcal{N}(z) : y \in N\} = \{z | y \in \bigcap \mathcal{N}(z)\}$$

Por lo tanto, si  $x \in \overline{\{y\}}$  se tiene que  $y \in \bigcap \mathcal{N}(x)$  y la equivalencia que buscamos es inmediata. □

## 2.4. Continuidad

El problema con respecto a la continuidad en términos evolutivos que ya habíamos comentado en el capítulo anterior nos lleva a buscar una definición formal de esta. En nuestro caso, la noción de proximidad será la base de la que partir. Antes de relacionar la continuidad con la proximidad, debemos introducir una definición más general con la cual podamos trabajar en espacios con estructuras más débiles y que sea aplicable al caso de las pretopologías. Debemos definir entonces un contexto en el que introducir la continuidad conserve la estructura ya definida en materia de pretopologías. Aparecen así los filtros.

**Definición 2.15.** Dado un conjunto  $X$ , un **filtro sobre  $X$**  consiste en una familia  $\mathcal{F}$  de subconjuntos de  $X$  que verifica las siguientes propiedades:

(F1)  $\emptyset \notin \mathcal{F}$ .

(F2) Sean  $F_1, F_2 \in \mathcal{F}$  entonces existe un conjunto  $F_3 \in \mathcal{F}$  tal que  $F_3 \subseteq F_1 \cap F_2$ .

(F3) Si  $F_1 \in \mathcal{F}$  y  $F_1 \subseteq F_2$ , entonces  $F_2 \in \mathcal{F}$ .

En el caso de satisfacer solo las dos primeras condiciones, hablaremos de  $\mathcal{F}$  como una **base de filtros** (que únicamente define un filtro). Es sencillo comprobar que el conjunto de entornos de un punto  $\mathcal{N}(x)$  constituye un filtro sobre  $X$ , y que una preuniformidad  $\mathcal{U}$  es un filtro sobre  $X \times X$ .

Sea entonces  $\mathcal{G}$  otro filtro definido sobre  $X$ , diremos que  $\mathcal{F}$  es más grueso que  $\mathcal{G}$  (o que  $\mathcal{G}$  es más fino que  $\mathcal{F}$ ) si  $\mathcal{F} \subseteq \mathcal{G}$ . De forma equivalente,  $\mathcal{F}$  es más grueso que  $\mathcal{G}$  si para cada  $F \in \mathcal{F}$  existe un  $G \subseteq F$  tal que  $G \in \mathcal{G}$ .

Un filtro  $\mathcal{F}$  es el más fino o un filtro maximal si no está contenido en ningún otro filtro, lo cual es equivalente a decir que para cualquier  $A \subset X$ , se tiene que  $A \in \mathcal{F}$  o  $(X \setminus A) \in \mathcal{F}$ , lo que justifica el nombre de “filtro”.

Los filtros nos serán útiles en este caso para introducir el concepto de convergencia en el espacio. Podemos pensar en ellos como generalizaciones del concepto de secuencias. Sea  $(x_n) = \{x_1, x_2, \dots\}$  una sucesión cualquiera, definimos el conjunto  $F_k$  de los elementos finales de la sucesión como  $F_k = \{x_k, x_{k+1}, \dots\}$ . Este conjunto verifica de manera inmediata los dos primeros axiomas de la definición de filtro y por tanto será una base de filtros para  $\mathcal{F}$ . En el caso de las secuencias en un espacio normado, afirmamos que  $(x_n)$  converge a un punto límite  $x$ ,  $x_n \rightarrow x$ , si para todo  $\epsilon > 0$  existe un entero  $n_\epsilon$  tal que  $\|x_k - x\| < \epsilon$  para todo  $k \geq n_\epsilon$ . La noción de filtro nos permite hablar de convergencia sin involucrar una noción de distancia como puede ser la norma  $\|x_k - x\|$ . En términos de entornos, la convergencia  $x_n \rightarrow x$  implica que para cada entorno  $N$  de  $x$ , hay un entero  $n_N$  tal que

$x_k \in N$  para todo  $k > n_N$ , lo cual significa que  $F_{n_N} \subseteq N$ . Recordemos que un sistema de entornos constituye un filtro. Así,  $(x_n)$  converge a  $x$  si y sólo si el filtro  $\mathcal{F}$  generado por los términos finales  $F_k$  de  $(x_n)$  es más fino que el filtro de entornos de  $x$ , es decir  $\mathcal{N}(x) \subseteq \mathcal{F}$ . Esto nos aporta una relación de proximidad más débil que una distancia e ilustra como se lleva a cabo la convergencia en un espacio pretopológico.

**Definición 2.16.** Sea  $(X, \mathcal{N})$  un espacio pretopológico y  $\mathcal{F}$  un filtro sobre  $X$ , entonces  $\mathcal{F}$  converge a  $x$ ,  $\mathcal{F} \rightarrow x$  o  $x \in \lim \mathcal{F}$ , si y sólo si  $\mathcal{N}(x) \subseteq \mathcal{F}$ .

La convergencia sienta las bases necesarias para introducir una definición de continuidad en este contexto.

**Definición 2.17.** Sea  $f : (X, \mathcal{N}) \rightarrow (Y, \mathcal{M})$  una función entre dos espacios pretopológicos y  $x$  un punto tal que  $x \in X$ . Decimos que  $f$  es **continua en  $x$**  si para todo filtro  $\mathcal{F}$  sobre  $X$ ,  $\mathcal{F} \rightarrow x$  implica que  $f(\mathcal{F}) \rightarrow f(x)$ .

Podemos trasladar esta definición de continuidad al lenguaje de los entornos, como vemos a continuación:

**Lema 2.18.** Sea  $f : (X, \mathcal{N}) \rightarrow (Y, \mathcal{M})$  una función arbitraria entre dos espacios pretopológicos. Los siguientes enunciados son equivalentes:

- (1)  $f$  es continua en  $x$ .
- (2) Para todo entorno  $M \in \mathcal{M}$  de  $f(x)$ , existe un entorno  $N \in \mathcal{N}$  de  $x$  tal que  $f(N) \subseteq M$ .
- (3)  $\mathcal{M}(f(x)) \subseteq f(\mathcal{N})$ .

El punto 3 se obtiene directamente como consecuencia de las definiciones de convergencia y continuidad. Nótese que “ $\mathcal{F} \rightarrow x \Rightarrow f(\mathcal{F}) \rightarrow f(x)$ ” da lugar a la implicación “ $\mathcal{N}(x) \subseteq \mathcal{F} \Rightarrow \mathcal{M}(f(x)) \subseteq f(\mathcal{F})$ ”. En el caso de los filtros, dada una función arbitraria  $f$ , esta conserva la relación “ser más fina”, es decir,  $\mathcal{F} \subseteq \mathcal{G}$  implica  $f(\mathcal{F}) \subseteq f(\mathcal{G})$ . Por lo que  $f(\mathcal{N}(x)) \subseteq f(\mathcal{F})$ , pero “ $f(\mathcal{N}(x)) \subseteq f(\mathcal{F})$  implica que  $\mathcal{M}(f(x)) \subseteq f(\mathcal{F})$ ” nos lleva a que “ $\mathcal{M}(f(x)) \subseteq f(\mathcal{N}(x))$ ”, que podemos interpretar como que el entorno de filtros de  $f(x)$  es más grueso que la imagen del entorno de filtros de  $x$ . Esto es equivalente al punto 2, el cual se corresponde con la definición de “ser más grueso” en términos de filtros. Por otro lado, esta es la definición de continuidad basada en entornos con la que estamos más familiarizados. Podemos entonces reescribir la continuidad en términos de relaciones de proximidad o preuniformidades.

**Definición 2.19.** Sea  $f : X \rightarrow Y$  una función arbitraria y  $U \subset X \times X$  una relación de proximidad tales que podemos introducir la notación  $(f \times f)(U) = \{(f(x), f(y)) | (x, y) \in U\}$ . Entonces dado  $f : (X, \mathcal{U}) \rightarrow (Y, \mathcal{V})$  donde  $\mathcal{U}$  y  $\mathcal{V}$  son preuniformidades sobre  $X$  e  $Y$  respectivamente, diremos que  $f$  es **uniformemente continua** si:

Para todo  $V \in \mathcal{V}$  tal que existe  $U \in \mathcal{U}$  se tiene  $(f \times f)(U) \subseteq V$ .

**Teorema 2.20.** Si  $f : (X, \mathcal{U}) \rightarrow (Y, \mathcal{V})$  es uniformemente continua, entonces  $f : (X, \mathcal{N}_{\mathcal{U}}) \rightarrow (Y, \mathcal{M}_{\mathcal{V}})$  es continua con respecto a las pretopologías inducidas sobre  $X$  e  $Y$ .

*Demostración.* Recordemos que  $N_x \in \mathcal{N}_{\mathcal{U}}(x)$  si y sólo si hay un  $U \in \mathcal{U}$  tal que  $U[x] \subseteq N_x$ . Entonces,  $f$  es continua con respecto a  $\mathcal{N}_{\mathcal{U}}$  y  $\mathcal{M}_{\mathcal{V}}$ , si para cada  $V \in \mathcal{V}$  hay un  $U \in \mathcal{U}$  tal que  $f(U[x]) \subset V[f(x)]$ . Supongamos entonces que  $f$  es uniformemente continua, es decir, para cada  $V \in \mathcal{V}$ , hay un  $U \in \mathcal{U}$  tal que  $(f \times f)(U) \subseteq V$ , por lo que:

$$\begin{aligned} (f \times f)(U)[f(x)] &= \{(f(x), f(y)) | (x, y) \in U\}[f(x)] = \{f(y) | (x, y) \in U\} \\ &= \{f(y) | y \in U[x]\} = f(U[x]) \subseteq V[f(x)] \end{aligned}$$

con lo que se concluye la demostración. □

## 2.5. Comportamiento en conjuntos finitos

El trabajo con pretopologías se simplifica considerablemente cuando estamos en espacios finitos, en ese caso existen entonces un número finito de filtros y todos son de la forma  $\mathcal{F} \equiv \dot{F} = \{F' | F \subseteq F'\}$ , donde  $F$  es un subconjunto de  $X$ . Denotaremos estos filtros como **filtros discretos**. Esta correspondencia entre los subconjuntos de  $X$  y sus filtros nos permite trasladar muchos resultados establecidos sobre filtros al caso de los subconjuntos. Un subconjunto particularmente útil es la vecindad asociada al entorno de filtros  $\mathcal{N}(x)$ :

$$N(x) = \bigcap \mathcal{N}(x) = \bigcap \{N | N \in \mathcal{N}(x)\}.$$

Esta noción de vecindad puede usarse para establecer una correspondencia entre un espacio pretopológico  $(X, \mathcal{N})$  y su correspondiente grafo orientado  $\Gamma(X, E)$ , donde  $X$  es el conjunto de los nodos y  $E$  el conjunto de aristas orientadas desde  $x$  hasta  $y$ ,  $E = \{(x, y) | x \in X, y \in N(x) \setminus \{x\}\}$ , de forma que se relaciona cada  $x$  con los elementos de su vecindad. De esta forma propiedades de la teoría de grafos pueden ser trasladadas al caso de los espacios pretopológicos. Algunos ejemplos de esto serían:

- Por construcción, la vecindad  $N(x)$  está compuesta por los principales elementos próximos a  $x$  y él mismo.
- Un espacio pretopológico  $(X, \mathcal{N})$  es más fino que otro sobre el mismo conjunto  $X$ , pongamos por ejemplo  $(X, \mathcal{M})$ , sí y sólo si  $\Gamma(X, E_{\mathcal{N}})$  es un subgrafo de  $\Gamma(X, E_{\mathcal{M}})$ .
- Una función  $f : (X, \mathcal{N}) \rightarrow (Y, \mathcal{M})$  es continua en  $x$  sí y sólo si  $f(N(x)) \subseteq M(f(x))$ , es decir, si  $f$  lleva la vecindad de  $x$  en la vecindad de  $f(x)$ .
- La segunda de las condiciones de simetría que definimos sobre los entornos hacen de  $(X, \mathcal{N})$  un grafo orientado simétrico, es decir, dos nodos están o conectados por una arista en ambos sentidos o no están conectados. Los grafos simétricos son isomorfos a los no orientados, los cuales son exactamente los asociados a espacios pretopológicos finitos que verifican la primera de las propiedades sobre simetría.

Como ya habíamos comentado una pretopología finita es una topología sí y sólo si, para cada  $x \in X$  y para cada  $y \in N(x)$  existe un  $N(y)$  tal que  $N(y) \subseteq N(x)$ . Esto implica que las vecindades antes definidas deben ser siempre conjuntos abiertos. Por lo que:

- La vecindad abierta  $T(x)$  es el conjunto abierto más pequeño que contiene a  $x$ . Este conjunto está compuesto por todos los elementos que pueden ser alcanzados a partir de  $x$  a través de cualquier arista directa.

Hasta aquí hemos introducido someramente los conceptos necesarios en materia de pretopologías, necesarios para estudiar la evolución de moléculas como el ARN. A lo largo del siguiente capítulo extenderemos su uso de manera explícita en un contexto biológico.

## Capítulo 3

# La combinación de las pretopologías y el mapa genotipo-fenotipo

A lo largo de este capítulo nos centraremos en la forma de implementar la estructura de las pretopologías sobre el espacio de los fenotipos, y como así podemos establecer distintos tipos de estructuras que presentan las condiciones que necesitábamos.

### 3.1. La pretopología de la accesibilidad

Consideraremos a partir de aquí la estructura del espacio de fenotipos inducida por la aplicación  $f$  desde los genotipos  $G$  hasta los fenotipos  $P$ . El plegado de las secuencias de ARN seguirá sirviéndonos de ejemplo.

El esquema de genotipos-fenotipos asigna a cada genotipo  $g$  un fenotipo  $\psi = f(g)$ . El problema principal en este momento es cómo organizar el conjunto de fenotipos, es decir, qué sistema de entornos es el *natural* para los distintos fenotipos. La pregunta equivalente en el caso de los genotipos no tiene ningún tipo de dificultad, ya que existen procesos físicos mediante los cuales los genotipos cambian directamente y así proporcionan una estructura de entornos en el conjunto de los posibles genotipos. Los fenotipos, de todas formas, no se modifican directamente. La innovación fenotípica es el resultado de la modificación genética a través del desarrollo (el esquema genotipo-fenotipo). Este razonamiento motivó a considerar la noción de entornos fenotípicos inducida por el esquema genotipo-fenotipo que básicamente se diferencia de la noción de proximidad entre fenotipos solamente en la comparación de sus rasgos morfológicos.

La estructura de entornos inducida en el conjunto de los fenotipos refleja la “accesibilidad” de un fenotipo a otro a través de mutaciones en el genotipo original. Un caso interesante es cuando el esquema genotipo-fenotipo va de muchos elementos distintos a

uno, lo cual es un caso habitual en la práctica. La noción de proximidad de un fenotipo  $\psi$  a otro debe ser una propiedad robusta, independiente de cualquier genotipo particular que de lugar a  $\psi$ . Debe, de alguna forma, reflejar una característica común a todos los genotipos cuyo fenotipo es  $\psi$ . En los esquemas que van de muchos elementos a un mismo resultado, los fenotipos denotarían clases de equivalencia sobre los genotipos (cada clase se correspondería con el conjunto de genotipos que comparten el mismo fenotipo). La proximidad entre fenotipos debe entonces reflejar la adyacencia natural de estas clases de equivalencia provocada por el sistema de entornos del espacio de genotipos.

Nosotros planteamos esta idea formalmente proponiendo primero una pregunta aparentemente no relacionada: ¿Qué tipo de sistema de entornos  $\mathcal{M}$  en el conjunto de los fenotipos hace el esquema genotipo-fenotipo sea continuo en todo punto?

A partir del Lema 2.20, sabemos que para que  $f$  sea continua, para todo fenotipo  $\psi$  y los genotipos  $g$  tales que  $f(g) = \psi$ , se debe tener que  $\mathcal{M}(\psi) \subseteq f(\mathcal{N}(g))$ . Cuando muchos genotipos  $g_i$  dan lugar a un mismo fenotipo  $\psi$ , el requisito para la continuidad se implementa de forma que

$$\begin{aligned}\mathcal{M}(\psi) &\subseteq f(\mathcal{N}(g_1)) \text{ y} \\ \mathcal{M}(\psi) &\subseteq f(\mathcal{N}(g_2)) \dots\end{aligned}$$

para todo  $g \in f^{-1}(\psi)$ . Definimos entonces el sistema de entornos

$$\mathcal{A}(\psi) := \bigcap_{g \in f^{-1}(\psi)} f(\mathcal{N}(g)) = \{S \mid S \in f(\mathcal{N}(g)), \forall g \in f^{-1}(\psi)\}.$$

De acuerdo con el axioma (N3),  $\mathcal{A}(\psi)$  debe incluir todos los conjuntos que contengan algún conjunto  $S$  de la forma dada en la definición de  $\mathcal{A}(\psi)$ , pero no tiene porque declararse explícitamente. El sistema de requisitos que establecimos antes podemos reescribirlo de manera equivalente como:

$$\mathcal{M}(\psi) \subseteq \mathcal{A}(\psi) \text{ para todo fenotipo } \psi.$$

Debemos entender que  $\mathcal{A}(\psi)$  tiene una interpretación muy sencilla. Por su definición anterior,  $\mathcal{A}(\psi)$  no es más que una colección de subconjuntos  $S$  que contienen la imagen de algún entorno que comparten todos los genotipos  $g$  con fenotipo  $\psi$ :

$$\mathcal{A}(\psi) = \{S \mid \exists N_g \in \mathcal{N}(g) \text{ tal que } f(N_g) \subseteq S, \forall g \in f^{-1}(\psi)\} = \dots,$$

la cual no es más que la colección de las imágenes de los entornos compartidos por todos los  $g$  con fenotipo  $\psi$  (incluyendo los conjuntos más grandes en base de la propiedad (N3)):

$$\dots = \{f(N) \mid N \in \mathcal{N}(g), \forall g \in f^{-1}(\psi)\} = f\left(\bigcap_{g \in f^{-1}(\psi)} \mathcal{N}(g)\right).$$

La colección  $\bigcap_{g \in f^{-1}(\psi)} \mathcal{N}(g)$  es el conjunto de los entornos compartidos por todos los genotipos  $g$  con un fenotipo  $\psi$ , es decir,  $\{N \mid N \in \mathcal{N}(g), \forall g \in f^{-1}(\psi)\}$ .

En el caso del ARN,  $f^{-1}(\psi)$  es lo que denominábamos como conjunto neutro (o cadena neutra cuando todas las secuencias que se pliegan en una misma estructura secundaria están conectadas mediante mutaciones), y  $\bigcap_{g \in f^{-1}(\psi)} \mathcal{N}(g)$  sería el sistema de entornos del conjunto neutro,  $\mathcal{N}(f^{-1}(\psi))$ . En resumen tendríamos:

$$\mathcal{M}(\psi) \subseteq \mathcal{A}(\psi) = \bigcap_{g \in f^{-1}(\psi)} f(\mathcal{N}(g)) = f\left(\bigcap_{g \in f^{-1}(\psi)} \mathcal{N}(g)\right) = f(\mathcal{N}(f^{-1}(\psi))).$$

En otras palabras, un fenotipo  $\vartheta$  está contenido en un entorno  $N_\psi$  del fenotipo  $\psi$  ( $N_\psi \in \mathcal{A}(\psi)$ ) si y sólo si hay un entorno de  $g \in f^{-1}(\psi)$  que contiene a un genotipo  $h$  que se pliega en forma  $\vartheta$ . Este es algo directo en el caso de trabajar con aplicaciones entre conjuntos finitos, donde la estructura de entornos viene definida por la vecindad (el entorno más pequeño, como ya comentamos en el capítulo anterior). En el espacio de genotipos, la vecindad del conjunto neutro de  $\psi$  incluye todas las secuencias obtenidas mediante una mutación puntual a partir de secuencias que se pliegan como  $\psi$ . Con respecto a los fenotipos, la vecindad de  $\psi$ ,  $A(\psi)$ , consiste en todas las estructuras  $\vartheta$  que podemos conseguir a partir de una única mutación puntual sobre secuencias que se pliegan en forma  $\psi$ .

$$A(\psi) = \bigcup_{g \in f^{-1}(\psi)} f(N(g)) = \{\vartheta \mid \exists g \in f^{-1}(\psi) \text{ y } h \in N(g) \text{ tal que } \vartheta = f(h)\}.$$

Como podemos ver, en este caso la definición presenta una unión en lugar de la intersección antes descrita. Esto es una propiedad que obtenemos gracias a trabajar con filtros. Veamos estas propiedades un poco más en profundidad.

Los filtros son conjuntos que pueden ser intersecados y unidos de la manera habitual. Debido al axioma (F3), la intersección de dos filtros puede expresarse como la unión de sus elementos. La unión de dos filtros por otra parte, no es la intersección de sus elementos y habitualmente no es ni un filtro. Además, la intersección de los elementos de dos filtros forma un filtro que denotaremos como  $\mathcal{F} \vee \mathcal{G} (\neq \mathcal{F} \cup \mathcal{G})$ . El siguiente lema formaliza la diferencia entre estos conceptos.

**Lema 3.1.** Sean  $\mathcal{F}$  y  $\mathcal{G}$  dos filtros, entonces

$$\mathcal{F} \cup \mathcal{G} = \{F \cup G \mid F \in \mathcal{F}, G \in \mathcal{G}\}.$$

*Demostración.* 1)  $F \cup G \in \mathcal{F} \cap \mathcal{G}$ , ya que  $F \in \mathcal{F}$  implica que  $F \cup G \in \mathcal{F}$  (debido a que  $F \subset F \cup G$ ), y de forma análoga,  $G \in \mathcal{G}$  implica que  $G \cup F \in \mathcal{G}$ .

2) Si  $H \in \mathcal{F} \cap \mathcal{G}$ , entonces  $H \in \mathcal{F}$  y  $H \in \mathcal{G}$ . Por tanto,  $H$  puede escribirse trivialmente como  $H = F \cup G$  con  $F = H \in \mathcal{F}$  y  $G = H \in \mathcal{G}$ .

□

Escribiremos entonces  $\mathcal{F} \vee \mathcal{G}$  para hablar de conjunto que obtenemos al intersecar los elementos de  $\mathcal{F}$  y  $\mathcal{G}$ ,

$$\mathcal{F} \vee \mathcal{G} := \{F \cap G \mid F \in \mathcal{F}, G \in \mathcal{G}\},$$

teniendo en cuenta que las intersecciones  $F \cap G$  son distintas del vacío para todo  $F \in \mathcal{F}$  y  $G \in \mathcal{G}$ . En otro caso,  $\mathcal{F} \vee \mathcal{G} = \emptyset$  y diríamos que  $\mathcal{F}$  y  $\mathcal{G}$  son disjuntos.

Observemos que  $F \cap G \subseteq F$  y  $F \cap G \subseteq G$ , lo que implica que  $F, G \in \mathcal{F} \vee \mathcal{G}$  para todo  $F \in \mathcal{F}$  y  $G \in \mathcal{G}$ , dados  $\mathcal{F}$  y  $\mathcal{G}$  no disjuntos. Así,  $F \cup G \subseteq \mathcal{F} \vee \mathcal{G}$ .

De todas formas, ya que nada nos asegura que la intersección  $F \cap G$  para algún  $F \in \mathcal{F}$  y  $G \in \mathcal{G}$  sea un elemento de  $\mathcal{F}$  o de  $\mathcal{G}$ ,  $\mathcal{F} \cup \mathcal{G}$  no es un filtro habitualmente. Por su parte,  $\mathcal{F} \vee \mathcal{G}$  es el filtro más grueso, que a su vez es más fino que  $\mathcal{F}$  y  $\mathcal{G}$ , ya que si hubiese un  $F \in \mathcal{F}$  y un  $G \in \mathcal{G}$  tal que  $F \cap G \notin \mathcal{F} \vee \mathcal{G}$ , no se cumpliría el axioma (F2) que habla sobre las intersecciones de filtros.

Volviendo entonces a nuestra definición anterior, la pretopología  $\mathcal{A}$  sobre el conjunto de fenotipos es la noción de accesibilidad fenotípica más débil. Débil en el sentido de que, de acuerdo con la definición de  $A(\psi)$ , para que el fenotipo  $\vartheta$  esté en un entorno de  $\psi$  simplemente necesitamos que  $\vartheta$  se obtenga a través de algún error puntual sobre una secuencia que se pliege de forma  $\psi$ .  $\mathcal{A}$  es la topología más fina en el conjunto de los fenotipos  $P$  que hace que  $f : (G, \mathcal{N}) \rightarrow P$  sea una función continua. Nos referiremos a  $\mathcal{A}$  como la **pretopología de la accesibilidad** del espacio de los fenotipos o la **pretopología final** generada por  $f$  sobre  $(G, \mathcal{N})$ .

El sentido más restrictivo de la accesibilidad surge tras exigir que  $\vartheta$  esté en el entorno de  $\psi$  si y sólo si  $\vartheta$  se obtiene en la vecinanza de toda secuencia con fenotipo  $\psi$ . En los casos finitos, esto se entiende como:

$$C(\psi) = \bigcap_{g \in f^{-1}(\psi)} f(N(g)) = \{\vartheta \mid \forall g \in f^{-1}(\psi) : \exists h \in N(g) \text{ tal que } \vartheta = f(h)\}.$$

En el caso de los conjuntos infinitos, no podemos simplemente reemplazar la intersección de filtros  $f(N(g))$  de la definición de  $\mathcal{A}(\psi)$  por su unión, ya que la unión de dos filtros no es un filtro de manera general, y como ya hemos visto. En su lugar, debemos usar el filtro que surge de las intersecciones entre los entornos de forma individual.

*Notación 3.2.* Usaremos la notación

$$\mathcal{F} \vee \mathcal{G} = \{F \cap G \mid F \in \mathcal{F}, G \in \mathcal{G}\}$$

para hablar del filtro más grueso que es más fino que  $\mathcal{F}$  y  $\mathcal{G}$ .

Debemos destacar que  $\mathcal{F} \vee \mathcal{G}$  sólo existe si  $F \cap G \neq \emptyset$  para todo  $F \in \mathcal{F}$  y  $G \in \mathcal{G}$ . Dado que  $f(g) \in N$  para todo  $N \in f(\mathcal{N}(g))$  y todo  $g \in f^{-1}(\psi)$ , ninguna intersección será vacía y el filtro de los entornos  $\mathcal{C}(\psi) = \bigvee_{g \in f^{-1}(\psi)} f(\mathcal{N}(g))$  existe.  $\mathcal{C}$  es la pretopología más gruesa que es más fina que  $f(\mathcal{N}(g))$  para todo  $g \in f^{-1}(\psi)$ . La denominamos **pretopología sombra** ya que el fenotipo  $\vartheta \in \mathcal{C}(\psi)$  “sigue” a  $\psi$  como una “sombra”, siendo la imagen de un entorno para cada  $g$  que se pliega en forma  $\psi$ .

### 3.2. Sistema de entornos estadístico

La topología de la accesibilidad  $\mathcal{A}$  fue construida a partir de la exigencia de que el esquema genotipo-fenotipo fuese continuo en todo punto. Esto parece un requisito demasiado fuerte, resultando en una estructura de entornos bastante más débil. Un sólo genotipo preparado para una transición de  $\alpha$  a  $\beta$  basta para hacer  $\beta$  accesible a partir de  $\alpha$ . La topología sombra  $\mathcal{C}$  peca en el otro extremo, ya que requiere que todo genotipo de  $\alpha$  sea mutable en un genotipo de  $\beta$ . En la simulación computacional del esquema genotipo-fenotipo del ARN, la pretopología  $\mathcal{C}$  coincide con la trivial, ya que los entornos de  $\mathcal{C}$  para cierto  $\alpha$  solo contienen a  $\alpha$ .

La idea de accesibilidad que desarrollamos en el capítulo anterior hacía incapié en la probabilidad de obtener una transición desde el fenotipo  $\alpha$  hasta el fenotipo  $\beta$  a través de mutaciones sobre los genotipos relacionados con  $\alpha$ . Esto nos aporta un método de interpolación entre los distintos extremos de la accesibilidad,  $\mathcal{A}$  y  $\mathcal{C}$ . La probabilidad de cambio en un fenotipo es proporcional al número de genotipos con un fenotipo  $\alpha$  que son adyacentes a genotipos con un fenotipo  $\beta$ .

Gráficamente podemos establecer la distribución de estos números en forma de esquema, como podemos ver en la figura 1.1. En el caso más simple, una versión de accesibilidad en términos de probabilidades introduce un punto de corte. Si  $A(\beta \curvearrowright \alpha)$  está por debajo de ese punto de corte,  $\beta$  no es accesible desde  $\alpha$ . Dependiendo del punto de corte, podemos construir un rango de estructuras accesibles sobre los fenotipos. El punto de corte apropiado debe estar determinado por factores biológicos, como la tasa de mutación, el tamaño de la población o el momento en el que sucede.

Las pretopologías que hemos ido definiendo a lo largo de todo este trabajo nos permiten justificar formalmente la idea de una topología del espacio de fenotipos definida en términos de la adyacencia mutacional de las clases de equivalencia del genotipo.

### 3.3. La continuidad de las trayectorias evolutivas

Una trayectoria evolutiva se puede ver como una aplicación entre el margen temporal y el espacio de los fenotipos. Cuando analizamos las series de muestras paleontológicas o las series de transiciones sobre las formas obtenidas en la simulación de la evolución del ARN, el margen temporal es discreto por naturaleza con una clara pretopología de forma natural. Simplemente numeramos las subsecuencias de la muestra y definimos las vecinanzas en el margen temporal para ser  $N(t) = \{t, t + 1\}$ . El espacio pretopológico correspondiente lo denotaremos como  $(\mathbb{N}, \mathcal{T})$ . Su correspondiente grafo es la cadera orientada presente en la parte izquierda de la figura 3.1.

Una trayectoria evolutiva consiste en la composición de dos funciones. En primer lugar, una función  $g : (\mathbb{N}, \mathcal{T}) \rightarrow (X, \mathcal{G})$  que asigna un genotipo  $g(t)$  a cada punto temporal  $t$ , considerando el tiempo de manera discreta. El espacio de genotipos  $(X, \mathcal{G})$  es un espacio pretopológico inducido por los operadores genéticos, como la mutación puntual de la figura 3.1. Esta primera función es entonces compuesta con el esquema genotipo-fenotipo definido como  $f : (X, \mathcal{G}) \rightarrow (Y, \mathcal{N})$ . La estructura del espacio de fenotipos  $Y$  será la pretopología de la accesibilidad, es decir,  $\mathcal{N} = \mathcal{A}, \mathcal{C}$  o una versión de la probabilidad como la establecida en la sección anterior.

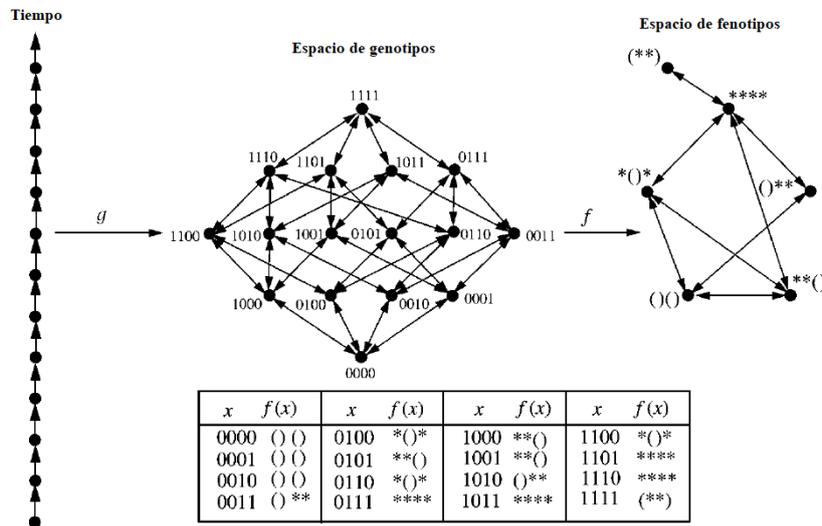


Figura 3.1: *Trayectoria evolutiva. Una trayectoria evolutiva será la composición  $f \circ g$  de una secuencia temporal de genotipos y el esquema genotipo-fenotipo  $f$ . Con una finalidad ilustrativa, el espacio de fenotipos se dota con la pretopología de la accesibilidad  $\mathcal{A}$ . El esquema  $f$  de genotipos-fenotipos viene resumido en la tabla inferior.*

Una trayectoria evolutiva es, por lo tanto, una función

$$\tau : (\mathbb{N}, \mathcal{T}) \rightarrow (Y, \mathcal{N}) \text{ tal que } t \mapsto \tau(t) = f(g(t))$$

cuya primera componente, las series temporales de genotipos  $g(t)$ , son continuas habitualmente, ya que los cambios del genotipo suceden a través de operadores genéticos básicos que definen la pretopología  $\mathcal{G}$  en  $X$ . De todas formas, no tiene por qué darse siempre este caso. Por ejemplo, si  $\mathcal{G}$  se obtiene a partir de mutaciones puntuales, entonces las mutaciones múltiples (es decir, aquellas en las que  $g(t)$  y  $g(t+1)$  difieren en más de una posición de la secuencia), las inserciones o las pérdidas constituyen discontinuidades en  $g : (\mathbb{N}, \mathcal{T}) \rightarrow (X, \mathcal{G})$ . Además, si no restringimos el caso de la continuidad en la pisada genética de la trayectoria evolutiva, todo vale y no podríamos establecer nada de manera clara. Si el esquema de genotipos-fenotipos  $f : (X, \mathcal{G}) \rightarrow (Y, \mathcal{N})$  es continuo en todo punto ( $\mathcal{N} = \mathcal{A}$ ), solo las discontinuidades genéticas pueden propiciar discontinuidades en el fenotipo.

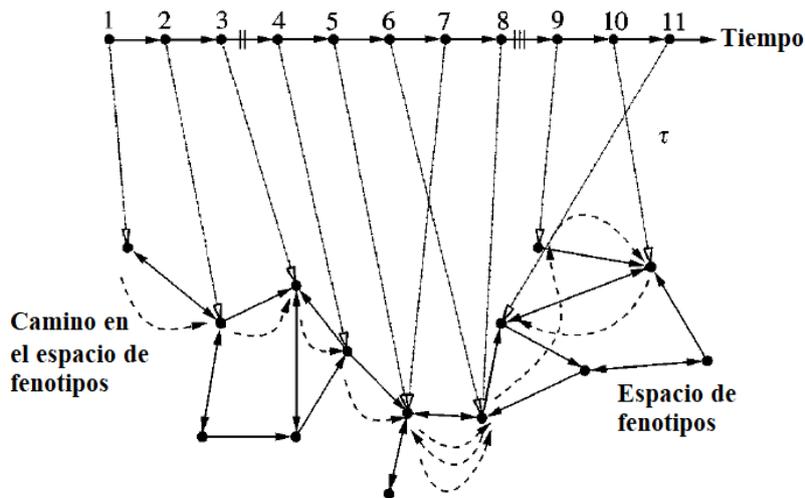


Figura 3.2: Visualización de la continuidad en una trayectoria evolutiva. Una trayectoria corta  $\tau : (\mathbb{N}, \mathcal{T}) \rightarrow (Y, \mathcal{N})$  es la que viene representada en este esquema. Las transiciones de  $t$  a  $t+1$  son continuas, o de forma más precisa,  $\tau$  es continua en  $t$ , si la transición  $[\tau(t) \rightsquigarrow \tau(t+1)]$  continua por la arista orientada en la pretopología del espacio de fenotipos. En este ejemplo hay dos transiciones no continuas, que denotaremos como  $\tau(3) \rightsquigarrow \tau(4)$  y  $\tau(8) \rightsquigarrow \tau(9)$ . Debemos destacar que la transición  $\tau(3) \rightsquigarrow \tau(4)$  pasa a ser continua cuando introducimos la topología  $\mathcal{N}$  ya que  $\tau(4)$  es accesible desde  $\tau(3)$  mediante una camino orientado. Por otra parte, la transición  $\tau(8) \rightsquigarrow \tau(9)$  se mantiene discontinua en todo caso.

En la práctica, la accesibilidad será más restrictiva que suponer  $\mathcal{N} = \mathcal{A}$  (y menos restrictiva que  $\mathcal{N} = \mathcal{C}$ ). Como hemos visto a lo largo del capítulo, la accesibilidad “efectiva” viene mejor descrita por una pretopología que sea (bastante) más fina que  $\mathcal{A}$ . Como consecuencia  $f$  no será continua en todo punto. Se puede dar el caso incluso de que para cualquier genotipo  $g$ , exista al menos una mutación de  $g$  que cambie su fenotipo de manera discontinua, haciendo que  $f$  sea continua en cualquier punto. Por lo tanto, ya que las mutaciones restantes sobre  $g$  cambian su fenotipo de forma continua, una trayectoria evolutiva  $\tau = f \circ g$ , compuesta por fenotipos limitados por la selección, debe resultar continua. Una transición en el tiempo  $t$  es continua si y sólo si  $\tau(t + 1) \in N(\tau(t))$ , es decir, si recorre alguna arista orientada del espacio de fenotipos. En el caso de un entorno continuo, la situación es cualitativamente similar, aunque mucho más difícil de visualizar. El margen temporal será el eje real  $\mathbb{R}$  con la topología estándar. De nuevo,  $\tau = f \circ g$  puede ser continua incluso si  $g$  o  $f$  no son continuas a lo largo de todo el espacio.

### 3.4. El caso del ARN

En esta última sección, sintetizaremos todo lo que hemos ido estableciendo en materia de evolución con respecto al ARN, para así resumir la idea de este trabajo. Nos restringiremos a un caso lo suficientemente simple y estable como para poder establecer una topología, algo que no es habitual como ya hemos comprobado a lo largo del trabajo. No nos adentraremos en plegamientos complejos de la molécula de ARN, por lo que solo trabajaremos con su estructura secundaria.

GUGAUGGAUU AGGAUGUCCU ACUCCUUUGC UCCGUAAGAU AGUGCGGAGU UCCGAACUUA CACGGCGCGC GGUUAC

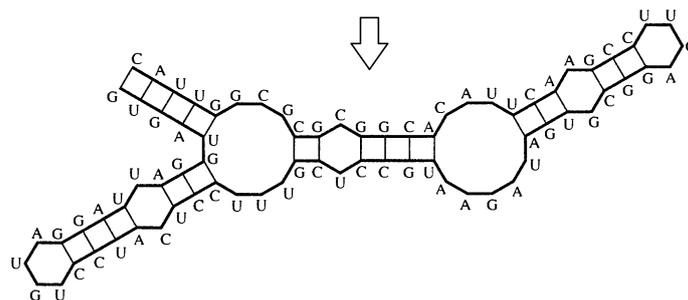


Figura 3.3: *Secuencia de ARN junto a su estructura.*

En este caso, el espacio de fenotipos viene dado por las moléculas de ácido ribonucleico o ARN y su plegamiento. Estas cadenas vienen dadas por nucleótidos que establecen un vínculo entre ellos para definir una cadena más fuerte y flexible. Los pares no adyacentes

de nucleótidos se pegan entre ellos aunque con menos fuerza, retorciendo la secuencia como podemos ver en la siguiente figura. Para simplificar el proceso, asumimos que se produce un único plegado a partir de una secuencia, aunque esto no sea del todo cierto.

A partir de aquí, usaremos solo cadenas relativamente cortas de nucleótidos para mostrar las estructuras y el modelo topológico, ya que las moléculas reales pueden tener desde diez hasta cientos de nucleótidos de longitud.

Dada una secuencia, su diagrama de plegado consiste en un diagrama sin ningún tipo de etiquetas representando la forma que adopta la secuencia de ARN. Este diagrama es lo que consideramos como fenotipo en nuestro modelo, también llamado anteriormente estructura secundaria o forma del ARN. Los espacios de fenotipos que introduciremos son conjuntos de formas de ARN donde definimos una topología.

En el modelo que introducimos solo asociamos una forma a cada secuencia y en él, un cambio puntual puede tanto cambiar totalmente la forma asociada como no modificar nada en absoluto. En el caso de mutaciones ocasionales, la selección natural determinará si el fenotipo persiste o no.

Dadas  $r$  y  $s$  dos formas de ARN, podemos calcular lo probable que es para  $r$  mutar en  $s$  como ya vimos en el primer capítulo. El cálculo de esta probabilidad nos permitirá definir topologías sobre el conjunto de formas del ARN.

Definimos  $m_{r,s}$  como el número de mutaciones puntuales que trasladan una secuencia de la cadena neutra de  $r$  hasta una de la cadena neutra de  $s$ . En este caso se tiene que  $m_{r,s} = m_{s,r}$  ya que toda mutación puntual que realice ese cambio tiene una mutación inversa que relaciona las cadenas en una dirección opuesta. Sea entonces  $m_{r,*}$  el número de mutaciones puntuales que nos llevan desde la cadena asociada a  $r$  hasta cualquier otra cadena neutra. Definimos entonces la probabilidad de mutación  $p_{r,s}$  como:

$$p_{r,s} = \frac{m_{r,s}}{m_{r,*}}.$$

Aunque  $m_{r,s} = m_{s,r}$ , no tiene por qué darse que  $p_{r,s} = p_{s,r}$  ya que los valores del denominador pueden diferir mucho. Esta asimetría se verá reflejada en las topologías que definamos en el espacio de fenotipos: mientras una forma  $s$  de ARN puede estar próxima a  $r$ ,  $r$  no tiene por qué estar próxima a  $s$ . Debido a estas asimetrías, como una distancia no nos servirá para organizar correctamente el espacio, un espacio topológico es la alternativa adecuada.

**Ejemplo 3.3.** Veamos esto de forma explícita. Consideremos entonces una topología sobre  $GC_{10}$ , el conjunto de las formas de ARN de genotipos de longitud 10 compuestos solo por guanina y citosina.

No discutiremos los detalles de los valores  $m_{r,s}$  y  $m_{r,*}$  dados para las formas del ARN en  $GC_{10}$ . Las  $2^{10} = 1024$  secuencias de este conjunto adoptarán un total de 8 formas, recogidas en la figura 3.4. A partir de aquí, podemos calcular las probabilidades de mutación, recogidas a continuación en la figura 3.5.

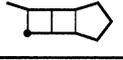
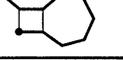
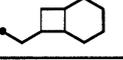
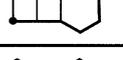
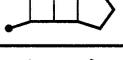
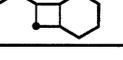
| $GC_{10}$ |   |     |       |   |    |       |  |    |
|-----------|---|-----|-------|---|----|-------|--|----|
| $s_1$     |  | 105 | $s_4$ |  | 26 | $s_7$ |  | 47 |
| $s_2$     |  | 128 | $s_5$ |  | 80 | $s_8$ |  |    |
| $s_3$     |  | 137 | $s_6$ |  | 70 |       |  |    |

Figura 3.4: El conjunto  $GC_{10}$ . Las 8 posibles formas que adoptan las secuencias, junto con su esquema y su incidencia.

|       | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $s_5$ | $s_6$ | $s_7$ | $s_8$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $s_1$ | —     | 0.13  | 0.15  | 0.08  | 0.07  | 0.09  | 0.04  | 0.44  |
| $s_2$ | 0.11  | —     | 0.15  | 0     | 0.11  | 0.18  | 0.05  | 0.40  |
| $s_3$ | 0.12  | 0.15  | —     | 0.03  | 0.09  | 0.06  | 0.05  | 0.50  |
| $s_4$ | 0.29  | 0     | 0.14  | —     | 0.07  | 0.09  | 0.06  | 0.35  |
| $s_5$ | 0.08  | 0.15  | 0.12  | 0.02  | —     | 0.08  | 0.08  | 0.47  |
| $s_6$ | 0.12  | 0.29  | 0.09  | 0.03  | 0.09  | —     | 0.03  | 0.35  |
| $s_7$ | 0.08  | 0.12  | 0.13  | 0.03  | 0.13  | 0.05  | —     | 0.46  |
| $s_8$ | 0.18  | 0.19  | 0.24  | 0.04  | 0.15  | 0.11  | 0.09  | —     |

Figura 3.5: Las probabilidades de mutación para  $GC_{10}$ .

En este caso, el valor de la celda  $(i, j)$  indica la probabilidad de llegar a la cadena neutra de  $s_j$  a partir de la de  $s_i$ .

La noción de proximidad que definimos se basa en la probabilidad de una mutación de una forma a otra. Como solo tenemos 8 posibles formas, cada una potencialmente puede mutar en las otras siete. En este caso, fijamos una probabilidad para tomar como referencia en materia de proximidad. Así, tomando por ejemplo  $p_{i,j} > 1/7$  establecemos una probabilidad mayor que el promedio de mutaciones a  $s_j$ .

Para cada  $i = 1, \dots, 8$ , definimos  $R_i = \{s_i\} \cup \{s_j | p_{i,j} > 1/7\}$ . La colección  $\mathcal{R}_{1/7} =$

$\{R_i\}_{i=1}^8$  no es una topología en sí misma, pero la extendemos hasta definir  $\mathcal{T}_{1/7}$  como la topología mínima que contiene a  $R_{1/7}$ . Esta topología viene generada por una base formada tomando intersecciones finitas de los conjuntos de  $\mathcal{B}_{1/7}$ . El espacio resultante es lo que conocemos como espacio de fenotipos.

A partir de la tabla de probabilidades es sencillo determinar cual será la base minimal  $\mathcal{B}_{1/7}$ . Para empezar, en la siguiente tabla destacamos aquellos casos en los que  $p_{i,j} > 1/7$ .

|       | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $s_5$ | $s_6$ | $s_7$ | $s_8$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $s_1$ | ✓     |       | ✓     |       |       |       |       | ✓     |
| $s_2$ |       | ✓     | ✓     |       |       | ✓     |       | ✓     |
| $s_3$ |       | ✓     | ✓     |       |       |       |       | ✓     |
| $s_4$ | ✓     |       |       | ✓     |       |       |       | ✓     |
| $s_5$ |       | ✓     |       |       | ✓     |       |       | ✓     |
| $s_6$ |       | ✓     |       |       |       | ✓     |       | ✓     |
| $s_7$ |       |       |       |       |       |       | ✓     | ✓     |
| $s_8$ | ✓     | ✓     | ✓     |       | ✓     |       |       | ✓     |

Figura 3.6: Los elementos de la tabla de probabilidades de  $GC_{10}$  que verifican que  $p_{i,j} > 1/7$ .

Para cada  $i$ , los tics en la fila  $i$  se corresponden con los elementos de  $R_i$ . Para determinar el elemento de la base  $B_i$  tomamos la intersección de las filas que contienen un tic en la columna  $i$ . La base  $\mathcal{B}_{1/7}$  que obtenemos mediante este proceso queda resumida en la siguiente figura.

|       |  |   |
|-------|--|---|
| $B_1$ |   |  |
| $B_2$ |   |  |
| $B_3$ |   |  |
| $B_4$ |   |    |
| $B_5$ |   |    |
| $B_6$ |   |    |
| $B_7$ |  |  |
| $B_8$ |  |  |

Figura 3.7: Base  $\mathcal{B}_{1/7}$  para la topología  $\mathcal{T}_{1/7}$ .

En este ejemplo en particular,  $s_8$  está cerca de cada  $s_j$  en el sentido que el elemento básico  $B_j$  contiene a  $s_8$  y por lo tanto todo entorno de  $s_j$ . Esto refleja el hecho de que cada  $s_j$  muta hacia  $s_8$  con una probabilidad muy alta. Por otro lado,  $B_8$  solo contiene a  $s_8$ , y por lo tanto ningún otro  $s_j$  está tan cerca de  $s_8$  como  $s_8$  lo está de ellos.

En este ejemplo hemos tomado como referencia la probabilidad  $p = 1/7$  para construir todo, pero podríamos fijar cualquier otra y obtener así diferentes topologías sobre el mismo conjunto de formas de ARN.

Por todo esto, los investigadores consideran que establecer una noción de proximidad usando topologías puede ayudar a aclarar los procesos biológicos en los que se basa la evolución de los organismos. Este modelo tan simple, usando formas del ARN, puede trasladarse a un contexto más general y complejo en materia de genotipos y fenotipos. Como ya hemos visto, la proximidad es un componente muy importante para la continuidad y de la misma forma lo es la estructura que nos permite determinarla. Como consecuencia, estos espacios de fenotipos nos proporcionan un marco adecuado para modelar e investigar cambios evolutivos.

# Bibliografía

- [1] Bärbel M. R. Stadler, Peter F. Stadler, Günter P. Wagner and Walter Fontana, *The Topology of the Possible: Formal Spaces Underlying Patterns of Evolutionary Change*, J. theor. Biol. (2001) 213, 241-274.
- [2] Colin Adams and Robert Franzosa, *Introduction to Topology: pure and applied*, First Edition, 2009, Pearson Education, ISBN 978-81-317-2692-1.
- [3] Steven A. Gaal, *Point set Topology*, 1964, Academic Press Inc.