

# Filogenia molecular

---

## Resumen

La filogenia molecular es la rama de la filogenia que analiza las diferencias moleculares hereditarias en las secuencias de ADN, ARN y proteínas para obtener información sobre las relaciones evolutivas de un organismo. El resultado de un análisis filogenético molecular se expresa en un árbol filogenético.

# Índice general

<b>I</b>	<b>Introducción a la filogenia: principios y conceptos</b>	<b>2</b>
I.1	Conceptos básicos . . . . .	3
I.2	Politomías . . . . .	4
I.3	Tipos de árboles filogenéticos . . . . .	5
I.4	Inferencia filogenética . . . . .	6
I.5	Homología . . . . .	6

# Capítulo I

## Introducción a la filogenia: principios y conceptos

La filogenia es la determinación de la historia evolutiva de los organismos. La filogenética es el estudio de la filogenia utilizando árboles filogenéticos de los distintos organismos y estudiar las relaciones entre ellos. Ha habido varias iniciativas a lo largo de la historia (como [ToL Web](#)) que han intentado lograr crear árboles de todas las especies, cuyo número ronda los 3-5 millones. Los escarabajos son los que más especies tienen, y con ellos se infieren las estimaciones sobre la biodiversidad.

La filogenia es una disciplina muy consolidada que empezó hace aproximadamente 200 años. La filogenia trabaja con árboles evolutivos, que son las **representaciones gráficas (patrones) de las relaciones ancestro-descendientes (relaciones históricas de parentescos) entre elementos**, que pueden ser especies, secuencias de genes, etc. Entender este patrón es esencial para realizar estudios comparativos de cualquier tipo, porque existen **dependencias estadísticas entre los elementos que comparten ancestros comunes**. Conforme va pasando el tiempo, se van aplicando diferentes modelos evolutivos y se van depurando. Cuantos más datos se añadan (más especies o más secuencias), se obtiene una mejor aproximación.

La filogenia sirve, entre otros, para:

- Evolución de los seres vivos
- Genómica
- Ingeniería genética
- Farmacia
- Epidemiología
- Biología de la conservación
- Control de plagas
- Lingüística

La filogenética puede ser estudiada de diversas maneras. A menudo ha sido estudiada utilizando registros fósiles, que contienen información sobre la morfología de los antepasados de especies actuales y la cronología de divergencias. Esto permite datar las filogenias. Sin embargo, el uso de registros fósiles tiene muchas limitaciones: pueden estar disponibles sólo para determinadas especies, los datos existentes de fósiles pueden estar fragmentados, la recolección de datos está limitada por la abundancia, hábitat, rango geográfico y otros factores, y las descripciones de los rasgos morfológicos son a menudo ambiguas (múltiples factores genéticos). Por todo esto, utilizar registros fósiles para determinar relaciones filogenéticas puede producir sesgos. Además, los fósiles de microorganismos son prácticamente inexistentes, imposibilitando el uso de este enfoque. Afortunadamente, los datos moleculares que están en la forma de secuencias de ADN o de proteínas pueden ser también muy útiles para proporcionar una perspectiva de la evolución de los organismos, como el ARN 16S. Debido a que los genes son el medio para registrar las mutaciones acumuladas, éstos pueden servir como "fósiles moleculares". A través del análisis comparativo de secuencias de ADN de una serie de organismos relacionados, la historia evolutiva de los genes e incluso de los organismos puede ser revelada. La ventaja de utilización de datos moleculares es que son más numerosos que los registros fósiles y más fáciles de obtener. Además, no hay ningún sesgo de muestreo, como el que hay en los registros fósiles reales. Por tanto, es posible construir árboles filogenéticos más precisos y robustos utilizando datos moleculares.

La filogenia representa un registro indirecto del proceso evolutivo al ser una reconstrucción de la evolución de caracteres. Se deben realizar test de homología para ver que los caracteres se pueden comparar entre sí al compartir un origen común. De esa forma se obtiene información para construir clasificaciones y hacer predicciones dentro de un marco temporal cuando es posible obtenerlo.

## I.1. Conceptos básicos

Los árboles filogenéticos suelen ser binarios, estando compuestos por **nodos externos o terminales** y **nodos internos** unidos por **ramas** que parten de una **raíz**. A través de las diferentes ramas se van reconstruyendo las relaciones entre las especies. Los **nodos internos son hipótesis evolutivas de posibles ancestros comunes** de los cuales normalmente faltan datos para confirmar o descartar la teoría. En las distintas ramas se pueden representar la transformación de caracteres que aparecen a nivel genético y que se transmiten por herencia.

Se denominan **grupos hermanos** a los nodos terminales que parten de un mismo nodo interno, es decir, dos taxones que compartan un ancestro común no compartido por ningún otro taxón. El **grupo externo (outgroup)** es aquel que se encuentra más alejado y parte de una rama distinta desde la raíz. Normalmente, este outgroup se elige de forma consciente para poder colocar la raíz donde se estima correcto. Todas las especies que se desarrollan desde una rama de la raíz se denomina **grupo interno o ingroup**.

Los árboles filogenéticos se pueden representar sin enraizar o enraizado. Un árbol filogenético sin raíz no asume conocimiento de un ancestro común, solo posiciones de los taxones para mostrar sus relaciones relativas (no hay dirección de un camino

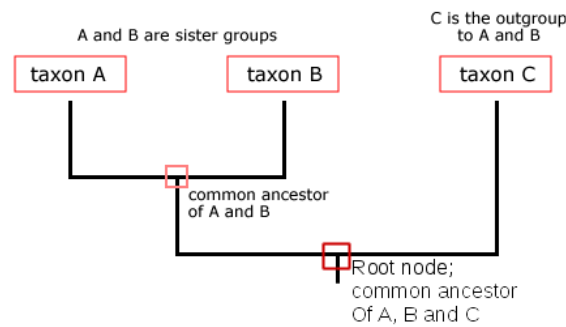


Figura I.1: Partes de un árbol filogenético.

evolutivo). Para definir la dirección de la evolución se necesita un árbol filogenético con raíz donde todas las secuencias bajo estudio tienen un ancestro o nodo raíz común (más informativo). Mientras que los árboles filogenéticos se centran en las relaciones evolutivas entre diferentes especies, las redes haplotípicas son representaciones gráficas sobre las relaciones evolutivas entre las diferentes poblaciones.

A la hora de visualización, hay varias formas de representar los árboles filogenéticos. Los distintos elementos no tienen un orden concreto; da igual si en un árbol los nodos terminales están en distinto orden mientras que las ramas sigan el mismo camino. En general, se suelen poner los nodos terminales de manera que sea más fácil de leer a simple vista.

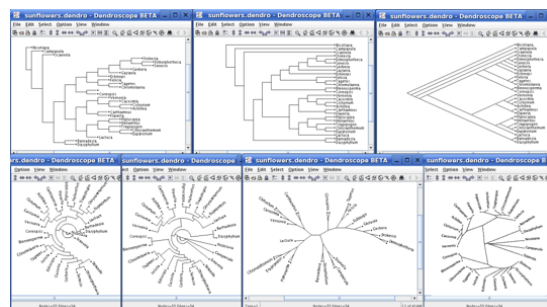


Figura I.2: Distintas representaciones de los árboles filogenéticos.

## I.2. Politomías

Topología es la forma en que se ramifica un árbol. Cuando todas las ramas se bifurcan en un árbol filogenético, éstas son denominadas como una **dicotomía**. Por el contrario, si una rama tiene más de dos descendientes, entonces se denomina **politomía**.

Los árboles filogenéticos se consideran resueltos cuando de un nodo interno salen las distintas terminales. En la mayoría de casos, los árboles son no resueltos y tienen politomías, es decir, que desde un nodo interno no se sabe cómo han avanzado las especies. A partir de ahí solo se pueden añadir más datos, pintar uniones con un bootstrap bajo (es decir, un bajo soporte de esa bifurcación), o justificarlo como que están en el momento de especiación. Dentro de las hipótesis filogenéticas siempre hay más de una solución (se producen varios árboles igualmente óptimos), así que el árbol

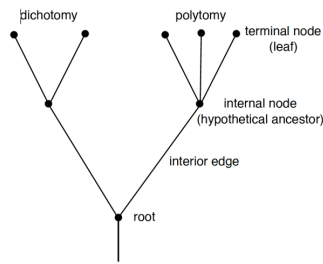


Figura I.3: Diferencia entre dicotomía y politomía.

final se debe elegir. Un árbol de consenso puede ser construido mostrando las porciones de bifurcación resueltas comúnmente y colapsando aquellas que no concuerdan entre los árboles. En un árbol de consenso estricto, todos los nodos en conflicto son colapsados.

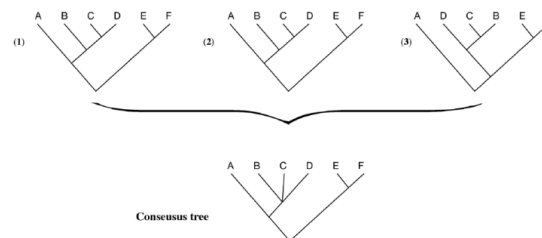


Figura I.4: Árbol de consenso.

### I.3. Tipos de árboles filogenéticos

Existen distintos tipos de árboles filogenéticos. Los **filogramas (phylogram)** miden en las ramas los cambios que ha habido por sitio, por lo que las longitudes de las ramas representan a escala la cantidad de divergencia evolutiva. Tienen la ventaja de mostrar tanto las relaciones evolutivas como la información sobre el tiempo relativo de divergencia de las ramas. Los **cladogramas (cladogram)** miden la similitud de los distintos elementos, pero las longitudes de sus ramas no son proporcionales al número de cambios evolutivos y, por tanto, no tienen ningún significado filogenético. Los cronogramas (chronogram) representan la relación de los elementos de forma temporal.

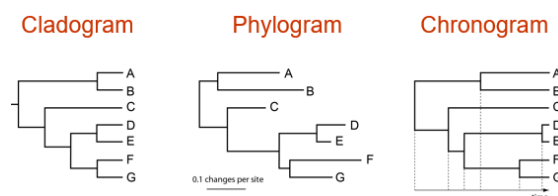


Figura I.5: Tipos de árboles filogenéticos.

## 1.4. Inferencia filogenética

Cualquier episodio histórico es, por definición, irrecuperable. La única forma que tenemos de reconstruirlo es a través del estudio de sus efectos. Por ello, la reconstrucción filogenética es un proceso de inferencia: se intenta obtener la mejor estima posible de una historia evolutiva basada en la información incompleta y con frecuencia ruidosa contenida en los datos. Las evidencias que se emplean pueden ser morfología (comparación entre caracteres de especies), ultraestructura, embriología, paleontología (registro fósil), etología (comportamiento animal), bioquímica y moléculas.

Un **carácter** es una característica de los taxones que supuestamente es heredada (si no es heredada, no se puede utilizar la filogenia). El **estado de carácter** es el valor específico que toma un carácter en un taxón concreto. Por ejemplo, un carácter sería tener ojos y el estado de ese carácter sería 2 para humanos y 8 para algunas arañas.

## 1.5. Homología

La **homología** es la relación que existe entre dos partes orgánicas diferentes de dos organismos distintos cuando sus determinantes genéticos tienen el mismo origen evolutivo, es decir, cuando un mismo órgano tiene diversas formas y funciones. Los caracteres que se estudian en filogenia deben ser homólogos. Se compara la semejanza de una estructura debido a la herencia común. Por el contrario, la analogía es una estructura semejante a otra o que tiene la misma función, pero cuyo desarrollo embrionario y origen son diferentes. No se presentan en un antepasado común (como en el caso de los caracteres homólogos), si no que es fruto de convergencia evolutiva.

Dentro de la homología se distinguen dos tipos: la ortología y la paralogía. Los **genes ortólogos** son semejantes por pertenecer a dos especies que tienen un antepasado común. Los **genes parálogos** son aquellos que se encuentran en el mismo organismo y cuya semejanza revela que uno procede de la duplicación del otro (y puede adquirir funciones diferentes del gen original). La ortología requiere que se haya producido especiación, mientras que esta no es necesaria en el caso de la paralogía, que puede producirse solo en los individuos de una misma especie. Por ello, idealmente se deben comparar caracteres ortólogos para hacer las reconstrucciones filogenéticas.

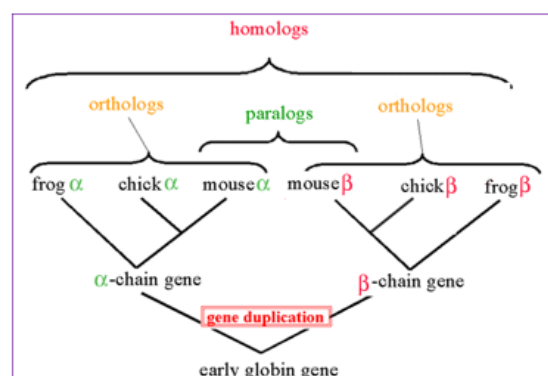


Figura 1.6: Homología en genes de la hemoglobina.