GAYANA Gayana (2026) vol. 90, No. 1, 1-16

https://doi.org/10.4067/s0717-6538202600010001

Review Article

Limitaciones, controversias y recomendaciones en los análisis filogenéticos: una guía para inferencias evolutivas

Limitations, controversies, and recommendations in phylogenetic analyses: a guide for evolutionary inferences

Christian M. Ibáñez[1,*] & Moisés A. Valladares[2,3,]**

[1]Instituto One Health, Facultad de Ciencias de la Vida, Universidad Andres Bello, Santiago, Chile.
[2]Laboratorio de Evolución, Sistemática y Biodiversidad (LESiB), Departamento de Ciencias Básicas, Universidad del Bío-Bío, Chillán, Chile.
[3]Grupo de Investigación en Biodiversidad & Cambio Global (GIBCG), Departamento de Ciencias Básicas, Universidad del Bío-Bío, Chillán, Chile.
Corresponding authors: *ibanez.christian@gmail.com; **mvalladares@ubiobio.cl

RESUMEN

Las reconstrucciones filogenéticas son esenciales para comprender la evolución de los organismos y clasificarlos según sus relaciones de ancestría-descendencia. Los estudios de biología comparada incorporan cada vez más información filogenética para explicar cómo han evolucionado los seres vivos. El avance de técnicas moleculares y la bioinformática han facilitado los análisis filogenéticos, lo que ha generado controversia entre los biólogos evolutivos, al considerarse que algunos nuevos métodos olvidan principios básicos. En este ensayo se muestra que los enfoques de máxima verosimilitud e inferencia Bayesiana permiten evidenciar la monofilia de los grupos de estudio. El artículo aborda problemas comunes en la interpretación de árboles filogenéticos, como la confusión entre nodos "basales" y "derivados" y el uso inadecuado de los valores de bootstrap. Se comparan métodos basados en caracteres —como la parsimonia, la máxima verosimilitud y la inferencia Bayesiana— con enfoques basados en distancias, como Neighbor-Joining. Este estudio presenta recomendaciones prácticas para mejorar la robustez de las inferencias filogenéticas, enfatizando la importancia de integrar múltiples métodos y evaluar críticamente los indicadores de soporte para obtener interpretaciones evolutivas más confiables.

Palabras clave: bootstrap, métodos de distancia, monofilia, sinapomorfía, sistemática.

ABSTRACT

Phylogenetic reconstructions are essential for understanding the evolution of organisms and classifying them according to their ancestry-descent relationships. Comparative biology studies are increasingly incorporating phylogenetic information to explain how living organisms have evolved. Advances in molecular techniques and bioinformatics have facilitated phylogenetic analysis, which has generated controversy among evolutionary biologists, as some new methods are seen as neglecting basic principles. This essay shows that maximum likelihood and Bayesian inference approaches provides evidence for the monophyly of the study groups. The article addresses common issues in the interpretation of phylogenetic trees, such as the confusion between "basal" and "derived" nodes and the inappropriate use of bootstrap values. Character-based methods –such as parsimony, maximum likelihood and Bayesian inference– are compared with distance-based approaches, such

Editor: Grisel Cavieres
Received: 03.03.2025; Accepted: 06.04.2026

Open Access

©2026 The Author(s). This open access article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits any noncommercial use, distribution, and reproduction, as long as appropriate credit is given to the original author(s) and the source.

as Neighbor-Joining. This study offers practical recommendations to improve the robustness of phylogenetic inferences, emphasizing the importance of integrating multiple methods and critically evaluating supporting indicators to achieve more reliable evolutionary interpretations.

Keywords: bootstrap, distance methods, monophyly, synapomorphy, systematics.

INTRODUCCIÓN

Las filogenias son herramientas fundamentales para comprender la historia de la vida, pues constituyen una de las bases de la clasificación de los organismos y permiten evaluar diversas hipótesis acerca de cambios evolutivos (Futuyma & Kirkpatrick 2023). Conocer y comprender las relaciones evolutivas ancestro-descendiente de un grupo de especies es una tarea clave en el estudio de la biología comparada (Ibáñez & Méndez 2014; Kapli *et al.* 2020; Wadas & Domingues 2025). Por consiguiente, la interpretación de los árboles filogenéticos se ha convertido en un componente esencial de la enseñanza de la biología, incluso en estudiantes de los primeros años en la universidad (Dees *et al.* 2014). Sin embargo, a pesar de su amplia aceptación y de la aparente simplicidad gráfica de los árboles filogenéticos, persisten problemas en su correcta interpretación (Gregory 2008; Omland *et al.* 2008; Dees *et al.* 2014).

Las relaciones filogenéticas entre distintas especies pueden ser reconstruidas con datos morfológicos y/o moleculares utilizando distintos métodos, como la Parsimonia (comúnmente llamada Máxima Parsimonia, MP), la Máxima Verosimilitud (Maximum Likelihood en inglés, ML) y la Inferencia Bayesiana (IB) (Peña 2011; Yang & Rannala 2012; Irisarri & Zardoya 2013; Ibáñez & Méndez 2014; Kapli *et al.* 2020). Los diferentes métodos para reconstruir una filogenia tienen sus ventajas y desventajas y se crearon para resolver distintos problemas. El método de la parsimonia busca el árbol que requiere el menor número de cambios evolutivos (Ibáñez & Méndez 2014), mientras que la máxima verosimilitud busca las mejores estimaciones puntuales de los valores de los parámetros utilizando los largos de ramas y un modelo de sustitución nucleotídica. En contraste, en la inferencia Bayesiana se genera una distribución de probabilidad completa sobre todos los posibles valores de parámetros, reflejando así la incertidumbre inherente en los árboles filogenéticos y las cuantificaciones del modelo como los largos de ramas y modelo de sustitución (Cheon & Liang 2014).

Aunque Parsimonia, Máxima Verosimilitud e Inferencia Bayesiana comparten el objetivo de reconstruir relaciones filogenéticas, difieren sustancialmente en sus fundamentos teóricos, manejo de la incertidumbre y sensibilidad a los datos. La Parsimonia, al no requerir un modelo evolutivo explícito, es conceptualmente simple, pero puede ser vulnerable a fenómenos como la homoplasia o la atracción por ramas largas (long branch attraction; LBA) (Yang & Rannala 2012). En cambio, la Máxima Verosimilitud incorpora modelos de sustitución y permite estimar parámetros evolutivos con mayor precisión, aunque su desempeño depende críticamente de la adecuación del modelo de sustitución utilizado (Irisarri & Zardoya 2013). La Inferencia Bayesiana, por su parte, ofrece una aproximación probabilística que integra la incertidumbre en la estimación de árboles y parámetros, pero requiere mayor poder computacional y una cuidadosa especificación de priors (Chen *et al.* 2014).

Independientemente del método empleado, la forma correcta de estudiar las relaciones filogenéticas es mediante la búsqueda de homologías entre las especies (indicios de descendencia con modificación; véase Pardo-Gandarillas 2014), ya sea utilizando caracteres morfológicos, moleculares o ambos. Las homologías reflejan el grado de parentesco entre organismos y constituyen evidencia de descendencia común. Es decir, caracteres que presentan los organismos que pertenecen a un mismo linaje (Pardo-Gandarillas 2014). Por ejemplo, la estructura ósea compartida en el brazo humano, el ala de un murciélago y la aleta de una ballena es una homología que evidencia un ancestro común. En cambio, las similitudes debidas a la evolución convergente o paralela, denominadas analogías u homoplasias, no aportan información útil para la reconstrucción filogenética y deben evitarse; un ejemplo de ello es la aparición independiente de alas en aves, mamíferos e insectos.

En este trabajo se presentan recomendaciones para

la correcta interpretación de los árboles filogenéticos, analizando tanto el uso, las ventajas y desventajas o limitaciones de cada método. Asimismo, se aborda y clarifica la controversia metodológica en la identificación de grupos monofiléticos. Además, se incluye un glosario (ver Apéndice) con los conceptos utilizados en este artículo.

Problemas de interpretación: nodos basales y especies derivadas

Los árboles filogenéticos poseen una terminología específica, en la cual cada componente tiene su nombre (Fig. 1). En un árbol filogenético resuelto (i.e., sin politomías), cada nodo de ramificación da lugar a dos linajes, considerados grupos hermanos, ya que se originan simultáneamente a partir de un ancestro común; por tanto, cada nodo representa un evento de especiación (Wiley *et al.* 1991). Una vez que dos linajes han divergido, cada uno desarrolla nuevos caracteres de forma independiente, aunque conserva características del ancestro. Esto hace que las especies de ambos linajes presenten similitudes con dicho ancestro común (Crisp & Cook 2005). Estas similitudes heredadas, expresadas como caracteres compartidos entre taxones, conforman la señal filogenética que los distintos métodos de reconstrucción buscan identificar y modelar.

Un aspecto estrechamente vinculado a la interpretación de los árboles filogenéticos es la determinación de la polaridad de los caracteres, es decir, la dirección de los cambios evolutivos que se definen si los caracteres son plesiomórficos (ancestrales) o apomórficos (derivados) (Wiley *et al.* 1991). Este proceso es central en los enfoques cladísticos, particularmente en parsimonia, y para establecer dicha polaridad resulta indispensable definir la raíz del árbol, lo cual se logra habitualmente mediante la inclusión de un grupo externo (outgroup en inglés) (Wiley *et al.* 1991). La elección del grupo externo no es arbitraria: debe corresponder a un linaje suficientemente relacionado con los taxones de interés, pero claramente excluido del grupo interno, de modo que permita inferir qué estados de carácter son ancestrales y cuáles son derivados (Maddison *et al.* 1984; Yang & Rannala 2012). En ausencia de un grupo externo, el árbol permanece no enraizado y, por tanto, carece de información sobre la dirección de las transformaciones evolutivas, lo que puede inducir interpretaciones erróneas acerca de la

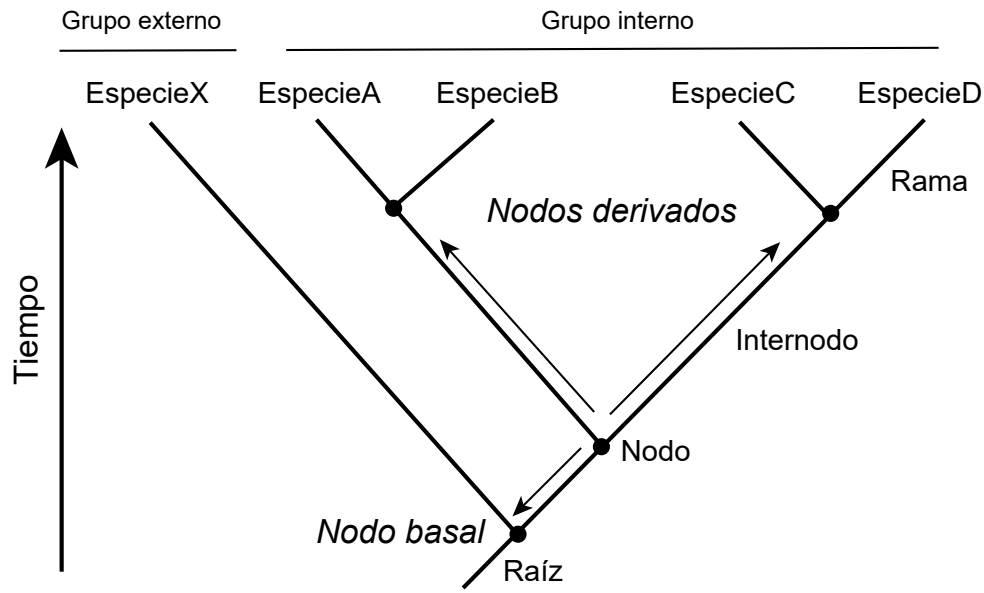

FIGURA 1. Descripción de un árbol filogenético y explicación gráfica de los nodos basales y derivados. En la figura, todas las especies existen en un mismo tiempo evolutivo y ninguna es basal. En cambio, los nodos pueden ser considerados basales o derivados solo en relación con un nodo específico de referencia o de interés (destacado en negrita). / Description of a phylogenetic tree and graphic explanation of basal and derived nodes. All species in the figure coexist at the same evolutionary time, and none is basal. Instead, nodes can be considered basal or derived only in relation to a specific reference or focal node (highlighted in bold).

posición “basal” de ciertos nodos o la supuesta condición “derivada” de determinadas especies. Sin embargo, dado que la mayoría de los modelos de evolución de secuencias asumen un proceso reversible, el puntaje de verosimilitud de un árbol no depende de la posición de la raíz (Irisarri & Zardoya 2013) permitiendo búsquedas más eficientes con árboles no enraizados (Abascal *et al.* 2013).

Diversos estudios han demostrado que la idoneidad del outgroup depende de su proximidad filogenética y del contenido de plesiomorfías que conserva, ya que un outgroup demasiado distante puede generar artefactos como la atracción de ramas largas (Lyons-Weiler *et al.* 1998; Schneider & Cannarozzi 2009). En este sentido, se recomienda seleccionar el grupo externo más cercano posible y, cuando es factible, incluir múltiples grupos externos para evaluar la estabilidad de las relaciones del grupo interno. Investigaciones recientes en filogenómica han reforzado estas buenas prácticas, mostrando que la inclusión de grupos externos cercanos y diversos mejora la robustez de los resultados y previene sesgos sistemáticos en la interpretación de las relaciones evolutivas (Lozano-Fernández 2022).

Erróneamente, algunos investigadores interpretan que un taxa que está más cercano a la raíz del árbol filogenético es “basal” en comparación con otro ubicado más lejos. Esta confusión, muy común en la literatura, surge porque no se han comprendido a cabalidad las deducciones que pueden hacerse a partir de un árbol filogenético. Todos los taxa de una filogenia coexisten en la actualidad y ninguno derivó antes que otro de su ancestro común, salvo en filogenias de fósiles. Ambas ramas que se originan en un nodo, es decir, los dos grupos hermanos, tienen la misma edad y han experimentado un cambio evolutivo equivalente (Krell & Cranston 2004). Se puede referir a un nodo interno como “basal” o “derivado” en relación con otro nodo (Fig. 1), puesto que estos nodos internos representan ancestros hipotéticos (Crisp & Cook 2005). En este sentido, se denomina “basal” a un nodo que se encuentra más próximo a la raíz que un nodo de referencia, y “derivado” a aquel que está más alejado, hacia las hojas del árbol. Sin embargo, estas designaciones no implican que existan taxa intrínsecamente basales o derivados (Crisp & Cook 2005).

Existen otros problemas de mala interpretación de los árboles filogenéticos y muy bien descritos en varios artículos (véase Gregory 2008; Omland *et al.* 2008; Dees *et al.* 2014). Entre estos destacan las incorrectas interpretaciones asociadas a los tiempos de divergencia molecular. Las especies divergen desde un ancestro común, técnicamente el Ancestro Común Más Reciente (ACMR), que en las filogenias es representado por el nodo de donde se separan las especies (Ho & Phillips 2009). Muchas veces este ACMR es interpretado como el origen evolutivo de cierto grupo, pero en realidad es la fecha estimada de la divergencia entre los linajes. Para encontrar el origen evolutivo temporal y espacial se utilizan análisis biogeográficos (dispersión, vicarianza), incluyendo la filogenia, que permiten conocer el área ancestral del taxa estudiado (Ibáñez *et al.* 2016; Ulloa *et al.* 2017; Pardo-Gandarillas *et al.* 2024).

Controversias sobre los grupos monofiléticos

Quizás la controversia más frecuente en filogenia, más que conceptual o interpretativa, se relaciona con la evidencia de monofilia utilizando los diferentes métodos mencionados. Se ha desarrollado una fuerte crítica de cómo se evidencia la monofilia con base en las sinapomorfías, caracteres derivados compartidos por dos o más taxones que se heredan de un ancestro común y permiten definir un clado (ver Apéndice) (Assis 2017). Sin embargo, ha resurgido el uso de la morfología en filogenia, asociado al amplio desarrollo de la taxonomía integrativa (Padial *et al.* 2010; Pulliandre *et al.* 2012), lo que ha renovado los métodos asociados a las sinapomorfías (véase Cornejo *et al.* 2018; Cona *et al.* 2022). En la Figura 2 se muestra cómo las sinapomorfías están asociadas a los caracteres derivados que sustentan los grupos monofiléticos (EspecieA, EspecieB, y EspecieC, EspecieD).

Libremente del método, podemos encontrar distintos tipos de relaciones entre las especies en un árbol. Un clado es un grupo “monofilético” que contiene al ancestro y todos sus descendientes, mientras que un grupo “polifilético” representa a un grupo artificial cuyos miembros son derivados de dos o más ancestros (Ibáñez & Méndez 2014). Asimismo, puede existir un grupo “parafilético” que corresponde a un grupo artificial que excluye a algunos miembros que comparten un ancestro común (Ibáñez & Méndez 2014). Se ha criticado este último término ya que reflejaría más los problemas taxonómicos que los filogenéticos de cierto grupo (Ebach *et al.* 2006; Ebach & Williams 2010). De hecho, Ebach & Williams (2010) sugieren usar el término “afilia” para referirse a grupos no-monofiléticos que requieren una revisión taxonómica exhaustiva. Un taxón es afilético

(artificial) cuando se considera que no es monofilético (parafilético o polifilético), por lo tanto, los taxones afiléticos son un problema taxonómico más que una anomalía evolutiva (Ebach & Williams 2010).

Varios autores han enfatizado que la cladística ha sido olvidada por la creciente adopción de análisis basados en modelos (inferencia Bayesiana, máxima verosimilitud) (e.g., Nelson 2004; Brower 2016; Assis 2017) e implementaciones informáticas (i.e., bootstrap) que no utilizan el criterio de las sinapomorfías para respaldar la monofilia. Además, se ha argumentado que los caracteres moleculares no siempre permiten identificar las sinapomorfías que sustentan los grupos naturales (Assis 2017). Una de las principales críticas radica en que el enfoque molecular reconoce clados basados en valores de bootstrap que no poseen sinapomorfías morfológicas conocidas (Assis 2017). En estos casos se pueden utilizar las sinapomorfías moleculares, regiones del genoma que definen grupos naturales (Herron & Freeman 2013). Los análisis moleculares realizados usualmente con el material hereditario (secuencias de ADN) comprenden cientos a miles de caracteres que sustentan los grupos monofiléticos y además estos caracteres son la base de la evolución. Aun así, algunos estudios reniegan de los resultados obtenidos cuando no son congruentes con la información morfológica o taxonómica (e.g., Ritschard *et al.* 2019).

Para comprender mejor las relaciones filogenéticas y las sinapomorfías que sustentan los grupos monofiléticos, algunos estudios filogenéticos en peces (Britz *et al.* 2014; Cornejo *et al.* 2018; Cona *et al.* 2022) han demostrado la importancia de comparar y combinar los resultados obtenidos a partir de datos morfológicos y moleculares –utilizando enfoques Bayesianos y de parsimonia, para evaluar la monofilia. Además, se utiliza frecuentemente más de un análisis en filogenias moleculares para sustentar las relaciones filogenéticas de las especies (e.g., Ibáñez *et al.* 2012; 2019; 2020; 2021; Morales *et al.* 2024; Pardo-Gandarillas *et al.* 2021; Valladares *et al.* 2022; Catalán *et al.* 2025), buscando alcanzar la congruencia entre análisis de parsimonia, máxima verosimilitud y Bayesianos. También es posible estimar, de manera independiente de la reconstrucción filogenética, la probabilidad de los caracteres sinapomórficos sin recurrir al método de parsimonia (Turjak & Trontelj 2012).

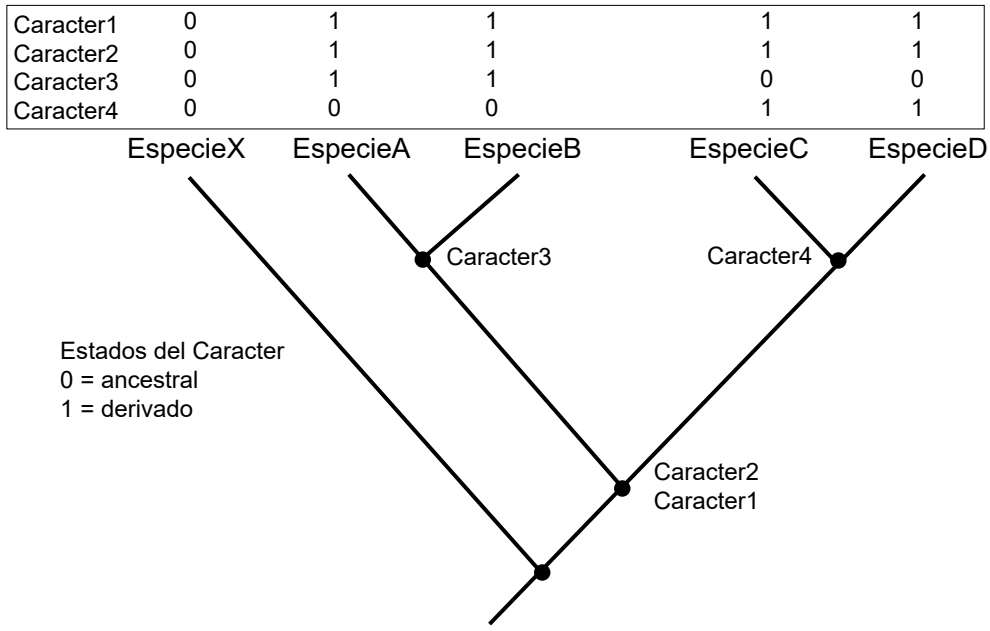

FIGURA 2. Representación gráfica del mapeo de sinapomorfías en un árbol filogenético. En este caso simplificado, utilizando cuatro caracteres binarios (0-1) los grupos monofiléticos (EspecieA, EspecieB) y (EspecieC, EspecieD) están sustentados por los rasgos derivados (sinapomorfías) Caracter3 y Caracter4, respectivamente. / Graphical representation of the mapping of synapomorphies on a phylogenetic tree. In this simplified case using four binary characters (0-1) the monophyletic groups (SpeciesA, SpeciesB) and (SpeciesC, SpeciesD) are supported by the derived traits (synapomorphies) Character3 and Character4, respectively.

Para esclarecer esta controversia, Wiley & Lieberman (2011) plantean algunas preguntas y tratan de responderlas: ¿Cumple solo la parsimonia, como algoritmo, el principio básico de Hennig (1966) de no asumir convergencia en ausencia de evidencia contraria? ¿Puede un enfoque probabilístico generar grupos monofiléticos confirmados por sinapomorfías? La agrupación por sinapomorfía (o principio auxiliar) subyace a toda la información filogenética. Pero ¿qué ocurre si la parsimonia y la argumentación Hennigiana clásica no son la única forma de cumplir el paradigma? En otras palabras, ¿existe más de una manera de abordar la filogenia dentro de la tradición Hennigiana?

Para contestar estas preguntas proponemos las siguientes argumentaciones lógicas. Si los análisis filogenéticos basados en máxima verosimilitud e inferencia Bayesiana intentan maximizar la homología y minimizar la homoplasia mediante modelos evolutivos, estos enfoques ya delimitan el grupo monofilético por medio de sinapomorfías, cumpliendo dos de los objetivos del paradigma Hennigiano. Además, para identificar las sinapomorfías de cada clado se pueden estimar independientemente de la reconstrucción del árbol y basadas en probabilidades (Turjak & Trontelj 2012), lo que sería realmente lo que Hennig quería conocer. Por lo tanto, la parsimonia no es la única forma de alcanzar los objetivos clásicos propuestos por Hennig (Wiley & Lieberman 2011). De hecho, Hennig (1966) enfatiza en múltiples ocasiones la necesidad de integrar toda la evidencia disponible (i.e., morfología, genética, ecología, fósiles) para obtener una mayor certeza sobre las relaciones filogenéticas entre las especies.

Nuestra recomendación para validar los grupos monofiléticos se basa en emplear diversos métodos de reconstrucción filogenética (MP, ML, IB) y en identificar las sinapomorfías utilizando tanto enfoques de parsimonia como de análisis probabilísticos. A pesar de que el uso de varios métodos se practica desde hace muchos años, "es frecuente que se utilice un único método" en sistemática filogenética. En casos particulares, como la estimación de tiempos de divergencia molecular o la reconstrucción de estados ancestrales, se utiliza un método filogenético, usualmente a través de inferencia Bayesiana, para estimar la incerteza tanto de la hipótesis filogenética como de los estados ancestrales (Pardo-Gandarillas *et al.*, 2018; Ibáñez *et al.*, 2018, 2023). Es fundamental evitar sesgos derivados de preferencias personales o tendencias prevalentes en la comunidad científica y comprender que cada método presenta ventajas y limitaciones. En este sentido, la correcta aplicación y el uso complementario de diferentes enfoques resultan más relevantes que seleccionar el método más usado o un método "superior", ya que calificar un método como el "mejor" es, en gran medida, una evaluación subjetiva.

De todas formas, no hay que olvidar que un árbol filogenético, independiente del método o de los caracteres utilizados, es considerado una hipótesis sobre la historia evolutiva de un grupo de organismos, ya que es aceptado hasta que tengamos nueva evidencia que lo modifique. Por lo tanto, todos aquellos métodos que permitan evaluar las relaciones ancestro-descendiente son útiles para evidenciar la evolución de los organismos vivos. No así los métodos de distancias (e.g., Neighbor-Joining, NJ) que solo permiten evaluar la similitud entre diferentes especies basándose en la cercanía genética más que en la homología (Peña 2011; Chen *et al.* 2014).

Utilidad y limitaciones de los métodos de distancia

Con el advenimiento de la genómica y la disponibilidad de grandes conjuntos de datos, se han desarrollado numerosos algoritmos y aplicaciones que agilizan el análisis de estos megadatos. En consecuencia, recientemente se ha recurrido a aproximaciones basadas en métodos de distancia, principalmente variantes de Neighbor-Joining (NJ) (Saitou & Nei 1987; véase también Desper & Gascuel 2006; Gascuel 2006), para reconstruir filogenias generadas a partir de datos genómicos. Para esto se utilizan tanto herramientas tradicionales para el análisis filogenético basado en métodos de distancia, como BIONJ (Gascuel 1997), FastME (Lefort *et al.* 2015) y MEGA (Kumar *et al.* 2024); así como nuevas implementaciones desarrolladas en entornos de análisis estadístico y de programación, tales como el paquete de R ape (Paradis *et al.* 2004) y RapidNJ (Simonsen *et al.* 2008). Más recientemente, se han propuesto optimizaciones de NJ enfocadas en datos genómicos, como las presentadas por Wang *et al.* (2023). Estos métodos ofrecen rapidez, escalabilidad y permiten obtener árboles en poco tiempo, lo cual es muy útil para explorar grandes conjuntos de datos. No obstante, esta renovada apuesta por los métodos de distancia no está exenta de limitaciones, especialmente al considerar los posibles problemas que pueden presentarse en ciertos escenarios evolutivos.

Una de las principales críticas a los métodos de distancia radica en que reducen la información contenida en las secuencias de ADN o proteínas a una

matriz de distancias. Este proceso de compresión puede ocasionar la pérdida de información evolutiva clave, ya que se ignoran detalles relevantes a nivel de carácter (o sitio) que pueden ser esenciales para la discriminación fina de las relaciones filogenéticas (Felsenstein 2004). En contraste, métodos basados en caracteres, como parsimonia (MP), máxima verosimilitud (ML) o inferencia Bayesiana (IB) modelan explícitamente la evolución de cada sitio y pueden incorporar modelos de sustitución. Mientras que estos últimos se enfocan en identificar sinapomorfías y maximizar homologías (véase sección anterior), las aproximaciones de distancia se limitan a medir grados de similitud global entre secuencias y generalmente incorporan la homoplasia. Esta diferencia ontológica en la naturaleza de la comparación explica por qué los resultados de NJ pueden diferir sustancialmente de los obtenidos mediante métodos basados en caracteres. Esta deficiencia ha llevado a algunos biólogos evolutivos a cuestionar si estas aproximaciones pueden considerarse verdaderamente filogenéticas en sentido estricto (Swofford *et al.* 1996; Peña 2011).

En estudios con pocos marcadores, la conversión a una matriz de distancias puede conducir a la generación de relaciones espurias (Philippe *et al.* 2011), debido a la escasa señal filogenética y al impacto del ruido en los datos (e.g., saturación del marcador). En tales situaciones, la capacidad del método para captar la verdadera historia evolutiva se ve comprometida. Un problema que, cabe señalar, no es exclusivo de los métodos de distancia, sino que también puede afectar a MP, ML e IB cuando la cantidad de información es insuficiente. En cambio, cuando se dispone de un gran número de marcadores o incluso de genomas completos, la acumulación de datos tiende a reforzar la señal filogenética, lo que en teoría reduce el riesgo de obtener relaciones espurias. No obstante, este beneficio depende en gran medida de la estimación apropiada de las distancias, lo cual requiere seleccionar modelos evolutivos adecuados y considerar la heterogeneidad entre marcadores. Si el modelo no se ajusta bien a los datos, o si estas fuentes de variación no se manejan correctamente, incluso en grandes conjuntos de datos se pueden derivar inferencias erróneas.

Otra limitación importante de los métodos de distancia es la dificultad para incorporar la heterogeneidad entre marcadores moleculares. Los métodos basados en caracteres permiten asignar modelos evolutivos diferentes a cada partición del conjunto de datos. Esto es posible en análisis de ML o IB, donde se pueden modelar procesos evolutivos complejos y heterogéneos (Guindon *et al.* 2010; Stamatakis 2014). En contraste, los métodos de distancia no permiten esta asignación explícita. Si se trabaja con múltiples marcadores, se puede optar por la concatenación de secuencias o por calcular matrices de distancia individuales para cada marcador y luego combinarlas. Sin embargo, esta combinación no resulta tan directa ni robusta como el análisis particionado. Esta limitación reduce la capacidad de estos métodos para capturar la variación en los procesos evolutivos que pueden existir entre diferentes regiones del genoma o entre distintos marcadores. Asimismo, los métodos de distancia no ofrecen una manera directa de implementar un enfoque coalescente multiespecie (Degnan & Rosenberg 2009), ya que no modelan la historia genealógica individual de cada marcador ni la variación entre ellos (Maddison & Knowles 2006), limitando su aplicación en estudios de complejidad evolutiva elevada.

A pesar de estas limitaciones, es importante recalcar que la recomendación no es descartar el uso de métodos de distancia, sino utilizarlos con la debida cautela y conocimiento de sus potenciales deficiencias. De hecho, en el contexto de análisis exploratorios o para obtener un primer agrupamiento de las muestras, estos métodos pueden ser muy útiles debido a su rapidez y escalabilidad. Además, en los últimos años se han desarrollado implementaciones que combinan diferentes enfoques. Por ejemplo, existen herramientas híbridas que inician la construcción de un árbol mediante búsquedas heurísticas inspiradas en métodos de distancia y, posteriormente lo optimiza mediante técnicas basadas en máxima verosimilitud (Price *et al.* 2009; Price *et al.* 2010; Liu *et al.* 2011). Este enfoque permite aprovechar la velocidad de los métodos de distancia para obtener una topología inicial y, a la vez, mejorar la precisión del árbol a través de la optimización basada en modelos evolutivos complejos. No obstante, es fundamental no depender exclusivamente de estos resultados. En cambio, es recomendable contrastar los hallazgos con análisis adicionales para asegurar la robustez de las inferencias filogenéticas.

En síntesis, los métodos de distancia, y en particular las variantes de Neighbor-Joining, ofrecen una alternativa rápida y computacionalmente eficiente para la reconstrucción filogenética, especialmente en contextos donde se requiere rapidez en el análisis o cuando se busca una primera aproximación al patrón de similitud y divergencia entre muestras. Sin embargo, su dependencia

en la conversión de secuencias en una matriz de distancias, junto con la dificultad para incorporar la heterogeneidad entre marcadores y asignar modelos evolutivos diferenciados, constituyen limitaciones importantes. Por ello, mientras que los métodos de distancia son recomendables para exploraciones preliminares y para establecer patrones generales de divergencia, para la reconstrucción de filogenias moleculares robustas y para estudios sistemáticos se aconseja complementar o incluso optar por enfoques basados en ML o IB, que utilizan directamente la información genética y pueden modelar de forma más precisa los procesos evolutivos. En la Tabla 1 se presentan las principales características, ventajas, limitaciones y aplicaciones recomendadas de los métodos filogenéticos más utilizados. Este cuadro no pretende ser exhaustivo, sino una guía dentro del marco de este trabajo, que debe complementarse con la exploración de alternativas y la consideración de las preguntas biológicas específicas de cada investigación.

Controversias en la interpretación de los soportes de nodos

El bootstrap es una técnica estadística ampliamente utilizada para evaluar la confianza en los nodos de los árboles filogenéticos estimados, usualmente pero no exclusivamente, mediante parsimonia y máxima verosimilitud. Esta técnica consiste en generar múltiples réplicas del conjunto de datos original mediante remuestreo con reemplazo, de modo que algunos sitios pueden repetirse y otros omitirse en cada réplica. A partir de cada réplica se reconstruye un árbol filogenético, y el porcentaje de réplicas en que aparece un nodo en particular se utiliza como medida de soporte para ese nodo (Fig. 3). En los análisis estándar se suelen generar alrededor de 1000 réplicas, y los valores de bootstrap se incorporan al árbol de consenso como una estimación de

TABLA 1. Comparación de los principales métodos filogenéticos en términos de su principio de inferencia, ventajas, limitaciones y aplicaciones recomendadas. / Comparison of the main phylogenetic methods in terms of their inference principle, advantages, limitations, and recommended applications.

Método	Principio	Ventajas	Desventajas	Aplicaciones
Métodos de distancia	Basados en transformar secuencias en una matriz de distancias y construir el árbol que minimiza la suma de distancias entre taxones.	Muy rápido y escalable; útil para set de datos grandes; fácil de implementar.	Reduce datos a matrices de distancia; pérdida de información de sitio; no modela procesos evolutivos; resultados menos robustos.	Exploración preliminar; agrupamiento rápido; punto de partida para métodos híbridos u otros métodos de inferencia.
Parsimonia	Busca la topología del árbol que requiere el menor número de cambios en los caracteres para explicar las especies analizadas.	Conceptualmente simple; rápido en set de datos pequeños; no requiere modelos de sustitución; adecuada para set de datos morfológicos.	Sensible a homoplasia; sesgo hacia agrupaciones incorrectas (e.g., atracción de ramas largas); poco eficiente en set de datos grandes.	Reconstrucción rápida en set de datos pequeños; comparación preliminar de caracteres morfológicos o moleculares.
Máxima verosimilitud	Busca la topología que maximiza la probabilidad de observar los datos dados un modelo de sustitución molecular.	Basado en modelos evolutivos explícitos; permite particionar y usar diferentes modelos; alta precisión.	Computacionalmente costoso; sensible a modelos mal especificados.	Reconstrucción de filogenias robustas; análisis de set de datos medianos a grandes; estudios sistemáticos.
Inferencia Bayesiana	Estima la probabilidad posterior de árboles y parámetros usando modelos evolutivos y distribuciones a priori.	Integra incertidumbre de parámetros y topologías; produce distribuciones posteriores; permite modelos complejos.	Muy demandante en cómputo; requiere priors adecuados; convergencia puede ser difícil de evaluar.	Estimación de filogenias con intervalos de credibilidad; adecuada para set de datos medianos; análisis comparativos.

la consistencia de la señal en los datos. Sin embargo, es fundamental comprender que el bootstrap mide dicha consistencia bajo remuestreo y no debe interpretarse como la probabilidad directa de que un clado sea correcto (Felsenstein 1985) o que ese grupo sea monofilético (en ausencia de otra evidencia). Un error frecuente en la literatura es asumir que valores elevados de bootstrap reflejan certeza absoluta, cuando en realidad pueden deberse a sesgos en los datos o en el modelo; del mismo modo, valores bajos no implican necesariamente que el clado sea falso, sino que la señal disponible es inconsistente o limitada. Por ello, los valores de bootstrap deben interpretarse siempre en conjunto con otros indicadores y en el contexto del análisis realizado.

Es importante destacar que existen diversas aproximaciones al bootstrap. La implementación "original" (Standard Bootstrap o SBS) fue desarrollada por Felsenstein (1985), mientras que modificaciones posteriores han

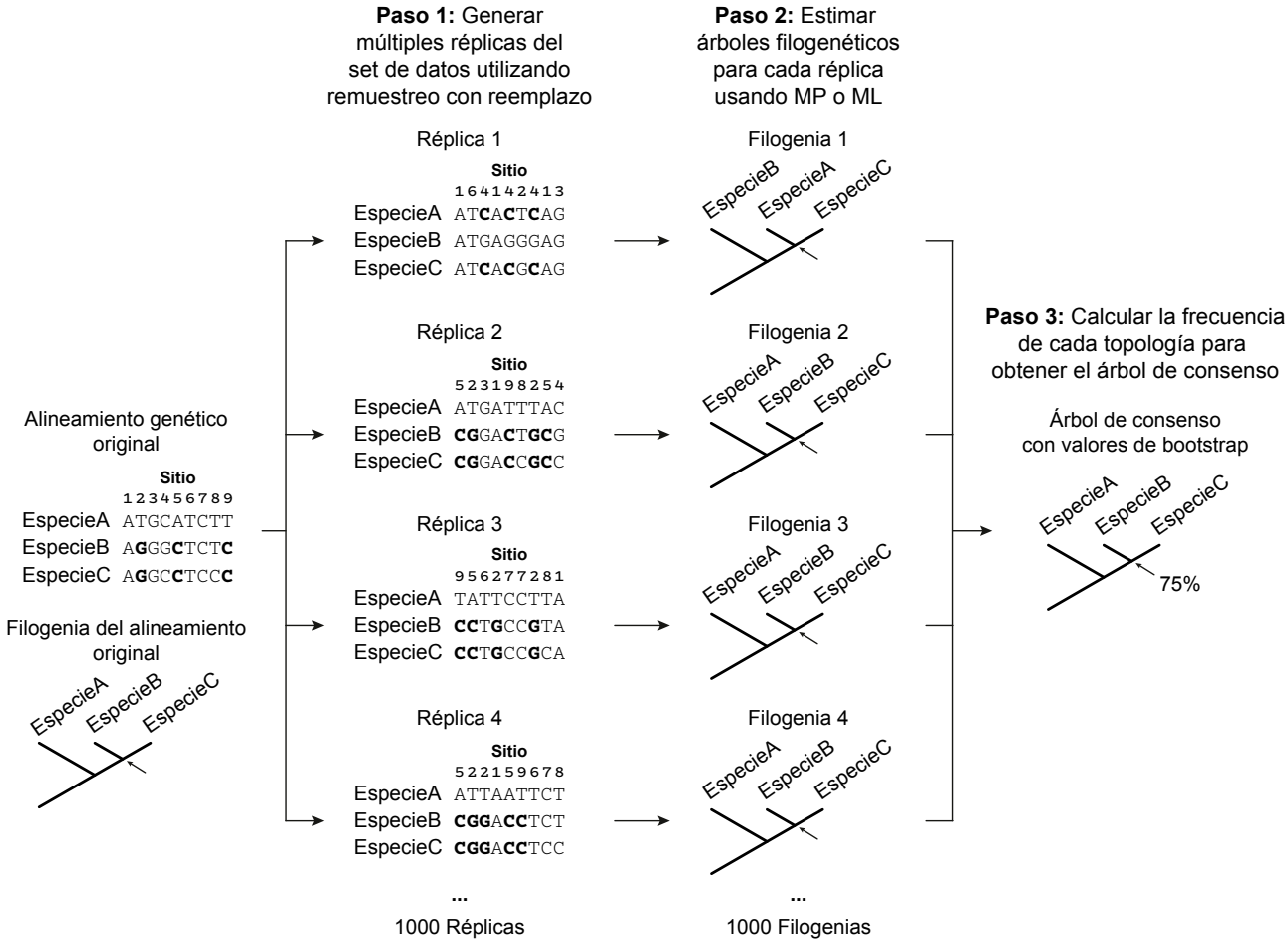

FIGURA 3. Proceso simplificado del análisis bootstrap. El método consta de tres pasos: (1) generar múltiples pseudo-réplicas del conjunto de datos original mediante muestreo con reemplazo, en las que algunos sitios pueden repetirse y otros omitirse. En el ejemplo, se muestra un alineamiento genético de nueve sitios nucleotídicos; (2) construir un árbol filogenético para cada réplica usando métodos como parsimonia (MP) o máxima verosimilitud (ML); y (3) calcular la frecuencia de aparición de los nodos para obtener un árbol de consenso con valores de soporte bootstrap. Por ejemplo, si el nodo que agrupa a EspecieB y EspecieC aparece en tres de cuatro árboles, se asigna un soporte del 75%. / Simplified process of bootstrap analysis. The method consists of three steps: (1) generate multiple pseudo-replicates of the original dataset by resampling with replacement, where some sites may be repeated and others omitted. In the example, a genetic alignment of nine nucleotide sites is shown; (2) build a phylogenetic tree for each replicate using methods such as parsimony (MP) or maximum likelihood (ML); and (3) calculate the frequency of nodes to obtain a consensus tree with bootstrap support values. For instance, if the node grouping SpeciesB and SpeciesC appears in three out of four trees, that grouping is assigned a support of 75%.

intentado mejorar su rendimiento y velocidad. Entre ellas se pueden mencionar el Rapid Bootstrap (RBS) (Stamatakis *et al.* 2008), la Ultrafast Bootstrap Approximation (UFBoot2) (Hoang *et al.* 2017) y el Transfer Bootstrap Expectation (TBE) (Lemoine *et al.* 2018). Aunque estos métodos pueden presentar en menor o mayor medida resultados similares, sus diferencias metodológicas pueden influir en la interpretación del soporte nodal.

Un error común en la literatura es considerar que valores altos de Bootstrap, por ejemplo, superiores al 70% (Hillis & Bull 1993; véase también Felsenstein & Kishino 1993), indican de manera absoluta la robustez de un nodo. Un alto soporte puede, en algunos casos, reflejar un sesgo en los datos. Por ejemplo, puede haber sitios con una señal filogenética homogénea pero inconsistente con la historia evolutiva real, lo que conduce a un soporte elevado para relaciones espurias. Un caso extremo de este tipo de error es la atracción de ramas largas (long branch attraction; LBA), donde linajes muy divergentes se agrupan artificialmente con altos soportes debido a limitaciones metodológicas (Felsenstein 1978; Bergsten 2005; Susko & Rogers 2021). Por otro lado, un valor bajo de bootstrap no necesariamente invalida la existencia de un nodo; puede ser consecuencia de datos insuficientes, del ruido estocástico presente en la muestra (Hillis & Bull 1993), o de procesos biológicos como la separación incompleta de linajes (incomplete lineage sorting; ILS), en la cual las genealogías de los genes difieren de la historia de especies (Maddison & Knowles 2006; Steenwyk *et al.* 2023; Liu *et al.* 2025). Así, el bootstrap ofrece una medida relativa de estabilidad de los nodos y debe ser interpretado en conjunto con otros indicadores, considerando además que puede verse afectado tanto por limitaciones metodológicas (LBA) como por procesos biológicos (ILS).

El método filogenético utilizado también influye en los valores de bootstrap. En el contexto de la parsimonia, el bootstrap se centra en evaluar la consistencia de la minimización de cambios evolutivos. Dado que este método busca la solución con el menor número de eventos, los valores de bootstrap en MP pueden estar fuertemente influenciados por la forma en que se distribuye la señal filogenética en los datos. Si la señal es homogénea, es posible obtener valores altos que, sin embargo, no garanticen la corrección del nodo. En sets de datos pequeños o con baja variabilidad, esta limitación puede ser especialmente problemática, ya que la sobreestimación de la señal conduce a soportes artificialmente inflados. Por otro lado, en la máxima verosimilitud, la interpretación del bootstrap depende en gran medida de la adecuación del modelo evolutivo empleado. Un modelo mal ajustado puede sesgar los valores de bootstrap, ya que el remuestreo evaluará la consistencia de la señal en función de parámetros que no reflejan adecuadamente el proceso evolutivo real (Yang 2006). Además, en ML es común observar que, en nodos con señal conflictiva, los valores de bootstrap pueden ser bajos, reflejando la dificultad del modelo para encontrar una solución única bajo esas condiciones. Sin embargo, es importante destacar que la señal conflictiva no siempre se origina en un mal ajuste del modelo, sino que también puede deberse a procesos biológicos genuinos, como la incongruencia entre datos nucleares y organelares (Kao *et al.* 2022; Quattrini *et al.* 2023), que generan diferentes historias evolutivas en el conjunto de datos.

Ante estas dificultades, se han propuesto diversas mejores prácticas para evaluar el soporte. Una recomendación es complementar los valores de bootstrap con otras métricas, como las pruebas de Approximate Likelihood-Ratio Test (aLRT) (Anisimova & Gascuel, 2006) o probabilidades posteriores en análisis Bayesianos (Huelsenbeck *et al.* 2001). Las probabilidades posteriores se calculan a partir del muestreo del espacio de parámetros mediante cadenas de Markov, proporcionando una estimación de la confianza en cada nodo basada en la distribución completa de probabilidades. No obstante, estas probabilidades tampoco están exentas de problemas, ya que pueden sobreestimar el soporte en ciertos casos (Suzuki *et al.* 2002). Sin embargo, esta combinación de indicadores o métricas de diferente naturaleza permite obtener una visión más equilibrada y crítica del soporte de cada nodo.

Finalmente, es fundamental que los estudios presenten una discusión crítica de los valores de bootstrap obtenidos. En lugar de simplemente reportar los porcentajes, se recomienda discutir las limitaciones inherentes a la técnica. Esto incluye reconocer que un valor alto no garantiza la corrección del nodo y que un valor bajo puede deberse a factores como la insuficiencia de datos o el ruido estocástico. Una presentación crítica permite a la comunidad científica comprender mejor la solidez de las inferencias y evitar interpretaciones erróneas.

CONCLUSIONES

En conclusión, la identificación de grupos monofiléticos debe fundamentarse en homologías, las cuales pueden evaluarse mediante diferentes métodos de soporte. Esto resulta particularmente relevante ante la incertidumbre que generan factores como el tipo de caracteres analizados o el número de especies incluidas en el estudio. Independiente del método utilizado, los árboles filogenéticos permiten clasificar los organismos vivos en jerarquías taxonómicas y comprender su evolución bajo la tesis de Darwin del ancestro común.

El presente estudio subraya la importancia de una interpretación rigurosa de los árboles filogenéticos para avanzar en la comprensión de la evolución. Se ha demostrado que la confusión entre nodos "basales" y "derivados" conduce a interpretaciones equivocadas sobre las relaciones ancestrales. Es fundamental recordar que todos los taxones representados en un árbol existen actualmente y que la clasificación de un nodo como "basal" se refiere únicamente a su posición relativa, sin implicar que ese grupo sea más ancestral que otro.

Asimismo, para identificar o evaluar la monofilia de un grupo se recomienda emplear diferentes métodos de reconstrucción filogenética en combinación con el mapeo de sinapomorfías. Es fundamental contrastar enfoques como la máxima verosimilitud, la inferencia Bayesiana y la parsimonia para validar los resultados y minimizar los sesgos inherentes a cada método. Además, resulta esencial adoptar una perspectiva crítica y objetiva, evitando que preferencias metodológicas u opiniones controvertidas influyan en la interpretación de los datos.

Además, los métodos de distancia, como Neighbor-Joining, resultan muy útiles para análisis exploratorios debido a su rapidez y capacidad para manejar grandes volúmenes de datos. No obstante, al transformar los datos moleculares en matrices de distancias se pierde información evolutiva esencial, lo que puede llevar a inferencias erróneas, especialmente en estudios con pocos marcadores. Esta limitación resalta la necesidad de complementar estos métodos con enfoques basados en caracteres, que permiten incorporar modelos evolutivos complejos y capturar la heterogeneidad en los datos.

Por otro lado, se ha analizado la utilidad y las limitaciones del bootstrap como medida de soporte de nodos. Aunque esta técnica ofrece una forma de evaluar la estabilidad de la señal filogenética mediante remuestreo, sus valores no deben considerarse probabilidades directas de corrección. Valores altos pueden deberse a sesgos en los datos, mientras que valores bajos pueden reflejar limitaciones en la información o la presencia de ruido estocástico.

Finalmente, se recomienda adoptar un enfoque integrador, combinando distintos métodos y fuentes de evidencia, y evaluando críticamente los indicadores de soporte. De esta forma, se podrán construir hipótesis filogenéticas más robustas y refinadas, reconociendo que los árboles filogenéticos son, en esencia, hipótesis sujetas a revisión conforme se disponga de nueva evidencia. Esta estrategia multidisciplinaria fortalecerá la comprensión de la evolución y la confiabilidad de las inferencias filogenéticas.

AGRADECIMIENTOS

Christian M. Ibáñez agradece al proyecto Regular-UNAB 04-2020 y INACH RG 22-25. Moisés A. Valladares agradece a ANID Subvención a la Instalación en la Academia Convocatoria 2025 ANID-SIA 85250178; ANID-FONDECYT de Iniciación en Investigación 11261415.

REFERENCIAS

Abascal, F., Irisarri, I., Zardoya, R. 2013. Filogenia y evolución molecular. In: Sebastián, A., Pascual-García, A. (Eds.) Bioinformática: 175-208. Sociedad Española de Bioinformática, España.

Anisimova, M., Gascuel, O. 2006. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Systematic Biology 55(4): 539-552. https://doi.org/10.1080/10635150600755453

Assis, L.C. 2017. The jazz of cladistics. Systematics and Biodiversity 15(5): 385-390. https://doi.org/10.1080/14772000.2017.1279693

Bergsten, J. 2005. A review of long-branch attraction. Cladistics 21: 163-193. https://doi.org/10.1111/j.1096-0031.2005.00059.x

Catalán, J., Ibáñez, C.M., Carrasco, S.A., Sellanes, J., Díaz, A., Pardo-Gandarillas, M.C. 2025. Systematic Evaluation of Sea Stars of the Genus *Heliaster* from the Southeastern Pacific and Redescription of *Heliaster helianthus*. Taxonomy 5(4): 59. https://doi.org/10.3390/taxonomy5040059

Chen, M.H., Kuo, L., Lewis, P.O. 2014. Bayesian Phylogenetics: methods, algorithms, and applications. CRC Press.

Cona, S., Cornejo, C., Hernández, S., Ibáñez, C.M. 2022.

Combining morphological and molecular information to infer phylogenetic relationships of lamniform sharks. Revista de Biología Marina y Oceanografía 57: 1-13. https://doi.org/10.22370/rbmo.2022.57.Especial.3339

Cornejo, C., Ibáñez, C.M., Hernández, C.E. 2018. Evaluación sistemática del género Helcogrammoides (Blenniformes, Trypteriigidae): Pequeños peces con grandes problemas. Revista de Biología Marina y Oceanografía 53(S1): 15-24. https://doi.org/10.22370/rbmo.2018.53.0.1250

Crisp, M.D., Cook, L.G. 2005. Do early branching lineages signify ancestral traits? Trends in Ecology & Evolution 20(3): 122-128. https://doi.org/10.1016/j.tree.2004.11.010

Darwin, C. 1859. El origen de las especies por medio de la selección natural.

Dees, J., Momsen, J.L., Niemi, J., Montplaisir, L. 2014. Student interpretations of phylogenetic trees in an introductory biology course. CBE—Life Sciences Education 13(4): 666-676. https://doi.org/10.1187/cbe.14-01-0003

Degnan, J.H., Rosenberg, N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology and Evolution 24(6): 332-340. https://doi.org/10.1016/j.tree.2009.01.009

Desper, R., Gascuel, O. 2006. Getting a Tree Fast: Neighbor Joining, FastME, and Distance-Based Methods. Current Protocols in Bioinformatics 15: 6.3.1-6.3.28. https://doi.org/10.1002/0471250953.bi0603s15

Ebach, M.C., Williams, D.M. 2010. Aphyly: a systematic designation for a taxonomic problem. Evolutionary Biology 37(2-3): 123-127. https://doi.org/10.1007/s11692-010-9084-5

Ebach, M.C., Williams, D.M., Morrone, J.J. 2006. Paraphyly is bad taxonomy. Taxon 55(4): 831-832. https://doi.org/10.2307/25065678

Felsenstein, J. 1978. Cases in which Parsimony or Compatibility Methods will be Positively Misleading. Systematic Biology 27: 401-410. https://doi.org/10.1093/sysbio/27.4.401

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39(4): 783-791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Felsenstein, J. 2004. Inferring Phylogenies. Sinauer Associates, Oxford University Press, Oxford. 580 p.

Felsenstein, J., Kishino, H. 1993. Is There Something Wrong with the Bootstrap on Phylogenies? A Reply to Hillis and Bull. Systematic Biology 42(2): 193-200. https://doi.org/10.1093/sysbio/42.2.193

Futuyma, D.J., Kirkpatrick, M. 2023. Evolution. 5th ed. Sinauer, Sunderland, Massachusetts, USA.

Gascuel, O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Molecular Biology and Evolution 14: 685-695. https://doi.org/10.1093/oxfordjournals.molbev.a025808

Gascuel, O., Steel, M. 2006. Neighbor-Joining Revealed. Molecular Biology and Evolution 23(11): 1997-2000. https://doi.org/10.1093/molbev/msl072

Gregory, T.R. 2008. Understanding evolutionary trees. Evolution: Education and Outreach 1(2): 121-137. https://doi.org/10.1007/s12052-008-0035-x

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Systematic Biology 59(3): 307-321. https://doi.org/10.1093/sysbio/syq010

Hennig, W. 1966. Phylogenetic systematics. University of Illinois Press, Urbana.

Hillis, D.M., Bull, J.J. 1993. An Empirical Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis. Systematic Biology 42(2): 182-192. https://doi.org/10.1093/sysbio/42.2.182

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution 35(2): 518-522. https://doi.org/10.1093/molbev/msx281

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294(5550): 2310-2314. https://doi.org/10.1126/science.1065889

Ibáñez, C.M., Pardo-Gandarillas, M.C., Poulin, E., Sellanes, J. 2012. Morphological and molecular description of a new record of *Graneledone* (Cephalopoda, Octopodidea) in the southeastern Pacific Ocean. Revista de Biología Marina y Oceanografía 47: 439-450. https://doi.org/10.4067/S0718-19572012000300007

Ibáñez, C.M., Pardo-Gandarillas, M.C., Peña, F., Gleadall, I.G., Poulin, E., Sellanes, J. 2016. Phylogeny and biogeography of *Muusoctopus* (Cephalopoda: Enteroctopodidae). Zoologica Scripta 45: 494-503. https://doi.org/10.1111/zsc.12171

Ibáñez, C.M., Rezende, E., Sepúlveda, R.D., Avaria-Llautureo, J., Hernández, C.E., Sellanes, J., Poulin, E., Pardo-Gandarillas, M.C. 2018. Thorson's rule, life history evolution and diversification of benthic octopuses (Cephalopoda: Octopodoidea). Evolution 72(9): 1829-1839. https://doi.org/10.1111/evo.13559

Ibáñez, C.M., Eernisse, D.J., Méndez, M.A., Valladares, M., Sellanes, J., Sirenko, B., Pardo-Gandarillas, M.C. 2019. Phylogeny, divergence times, and species delimitation of *Tonicia* (Polyplacophora: Chitonidae) from the Eastern Pacific Ocean. Zoological Journal

of the Linnean Society 186: 915-933. https://doi.org/10.1093/zoolinnean/zlz006

Ibáñez, C.M., Fenwick, M., Ritchie, P.A., Carrasco, S.A., Pardo-Gandarillas, M.C. 2020. Systematics and Phylogenetic Relationships of New Zealand Benthic Octopuses (Cephalopoda: Octopodoidea). Frontiers in Marine Science 7: 182. https://doi.org/10.3389/fmars.2020.00182

Ibáñez, C.M., Pardo-Gandarillas, M.C., Méndez, M.A., Sellanes, J., Sigwart, J., Sirenko, B. 2021. Phylogenetic position and morphological descriptions of *Chiton* species from the Southeastern Pacific. Zoological Journal of the Linnean Society 191: 695-719. https://doi.org/10.1093/zoolinnean/zlaa067

Ibáñez, C.M., Márquez-Gajardo, C., Torres, F.I., Sales, J.B.L. 2023. Biological traits as determinants in the macroecological patterns of distribution in loliginid squids. Marine Biology 170: 133. https://doi.org/10.1007/s00227-023-04286-1

Irisarri, I., Zardoya, R. 2013. Phylogeny Reconstruction. In: eLS. John Wiley & Sons, Ltd, Chichester.

Kao, T., Wang, T., Ku, C. 2022. Rampant nuclear–mitochondrial–plastid phylogenomic discordance in globally distributed calcifying microalgae. New Phytologist 235: 1394-1408. https://doi.org/10.1111/nph.18219

Kapli, P., Yang, Z., Telford, M.J. 2020. Phylogenetic tree building in the genomic age. Nature Reviews Genetics 21(7): 428-444. https://doi.org/10.1038/s41576-020-0233-0

Kumar, S., Stecher, G., Suleski, M., Sanderford, M., Sharma, S., Tamura, K. 2024. MEGA12: Molecular Evolutionary Genetic Analysis Version 12 for Adaptive and Green Computing. Molecular Biology and Evolution 41(12): msae263. https://doi.org/10.1093/molbev/msae263

Lefort, V., Desper, R., Gascuel, O. 2015. FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program. Molecular Biology and Evolution 32(10): 2798-2800. https://doi.org/10.1093/molbev/msv150

Lemoine, F., Domelevo Entfellner, J.-B., Wilkinson, E., Correia, D., Dávila, F.M., De Oliveira, T., Gascuel, O. 2018. Renewing Felsenstein's phylogenetic bootstrap in the era of big data. Nature 556(7702): 452-456. https://doi.org/10.1038/s41586-018-0043-0

Liu, K., Linder, C.R., Warnow, T. 2011. RAxML and FastTree: Comparing Two Methods for Large-Scale Maximum Likelihood Phylogeny Estimation. PLoS ONE 6(11): e27731. https://doi.org/10.1371/journal.pone.0027731

Liu, Y.-X., Huang, X.-M., He, W.-C., Erst, A.S., Xiang, K.-L., Wang, W. 2025. Phylogenomic insights into incomplete lineage sorting and hybridization in early-diverging eudicots. Molecular Phylogenetics and Evolution 212: 108422. https://doi.org/10.1016/j.ympev.2025.108422

Lozano-Fernández, J. 2022. A practical guide to design and assess a phylogenomic study. Genome Biology and Evolution 14(9): evac129. https://doi.org/10.1093/gbe/evac129

Lyons-Weiler, J., Hoelzer, G.A., Tausch, R.J. 1998. Optimal outgroup analysis. Biological Journal of the Linnean Society 64(4): 493-508. https://doi.org/10.1111/j.1095-8312.1998.tb00346.x

Maddison, W.P., Knowles, L.L. 2006. Inferring Phylogeny Despite Incomplete Lineage Sorting. Systematic Biology 55(1): 21-30. https://doi.org/10.1080/10635150500354928

Maddison, W.P., Donoghue, M.J., Maddison, D.R. 1984. Outgroup analysis and parsimony. Systematic Biology 33(1): 83-103. https://doi.org/10.1093/sysbio/33.1.83

Morales, P., Gajardo, F., Valdivieso, C., Valladares, M.A., Di Genova, A., Orellana, A., Gutiérrez, R.A., González, M., Montecino, M., *et al.* 2024. Genomes of the *Orestias* pupfish from the Andean Altiplano shed light on their evolutionary history and phylogenetic relationships within Cyprinodontiformes. BMC Genomics 25: 614. https://doi.org/10.1186/s12864-024-10416-w

Omland, K.E., Cook, L.G., Crisp, M.D. 2008. Tree thinking for all biology: the problem with reading phylogenies as ladders of progress. BioEssays 30(9): 854-867. https://doi.org/10.1002/bies.20794

Paradis, E., Schliep, K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35(3): 526-528. https://doi.org/10.1093/bioinformatics/bty633

Pardo-Gandarillas, M.C. 2014. Evidencias de la Evolución. In: Méndez, M.A., Navarro, J. (Eds.) Introducción a la Biología Evolutiva: 42-67. ESEB-SOCEVOL, Chile. e-book. 229 pp.

Pardo-Gandarillas, M.C., Torres, F.I., Fuchs, D., Ibáñez, C.M. 2018. Updated molecular phylogeny of the squid family Ommastrephidae: insights into the evolution of spawning strategies. Molecular Phylogenetics and Evolution 120: 212-217. https://doi.org/10.1016/j.ympev.2017.12.014

Pardo-Gandarillas, M.C., Díaz-Santana-Iturrios, M., Fenwick, M., Villanueva, R., Ibáñez, C.M. 2021. Redescription of the flapjack octopod, *Opisthoteuthis bruuni* (Cephalopoda: Opisthoteuthidae), from the southeastern Pacific Ocean and evolutionary relationships of cirrate octopods. Malacologia 63(2): 155-169. https://doi.org/10.4002/040.063.0201

Pardo-Gandarillas, M.C., Carrasco, S.A., Varela, A.I., Ibáñez, C.M. 2024. Systematic and biogeography of two sympatric octopuses from the remote Juan Fernández

Archipelago, South Pacific Ocean. Reviews in Fish Biology and Fisheries 34: 1685-1706. https://doi.org/10.1007/s11160-024-09894-8

Peña, C. 2011. Métodos de Inferencia Filogenética. Revista Peruana de Biología 18(2): 265-267.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., Baurain, D. 2011. Resolving Difficult Phylogenetic Questions: Why More Sequences Are Not Enough. PLoS Biology 9(3): e1000602. https://doi.org/10.1371/journal.pbio.1000602

Price, M.N., Dehal, P.S., Arkin, A.P. 2009. FastTree: Computing Large Minimum Evolution Trees with Profiles Instead of a Distance Matrix. Molecular Biology and Evolution 26(7): 1641-1650. https://doi.org/10.1093/molbev/msp077

Price, M.N., Dehal, P.S., Arkin, A.P. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLoS ONE 5(3): e9490. https://doi.org/10.1371/journal.pone.0009490

Quattrini, A.M., Snyder, K.E., Purow-Ruderman, R., Seiblitz, I.G.L., Hoang, J., Floerke, N., Ramos, N.I., Wirshing, H.H., Rodríguez, E., *et al.* 2023. Mito-nuclear discordance within Anthozoa, with notes on unique properties of their mitochondrial genomes. Scientific Reports 13: 7443. https://doi.org/10.1038/s41598-023-34059-1

Ritschard, E.A., Guerrero-Kommritz, J., Sanchez, J.A. 2019. First molecular approach to the octopus fauna from the southern Caribbean. PeerJ 7: e7300. https://doi.org/10.7717/peerj.7300

Ronquist, F., Teslenko, M., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Systematic Biology 61(3): 539-542. https://doi.org/10.1093/sysbio/sys029

Saitou, N., Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4(4): 406-425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Schneider, A., Cannarozzi, G. 2009. Support measures for phylogenetic trees: The pOG score. Molecular Biology and Evolution 26(6): 1259-1263. https://doi.org/10.1093/molbev/msp034

Simonsen, M., Mailund, T., Pedersen, C.N.S. 2008. Rapid Neighbour-Joining. In: Crandall, K.A., Lagergren, J. (Eds.) Algorithms in Bioinformatics. WABI 2008. Springer, Berlin. https://doi.org/10.1007/978-3-540-87361-7_10

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9): 1312-1313. https://doi.org/10.1093/bioinformatics/btu033

Stamatakis, A., Hoover, P., Rougemont, J. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. Systematic Biology 57(5): 758-771. https://doi.org/10.1080/10635150802429642

Steenwyk, J.L., Li, Y., Zhou, X., Shen, X.-X., Rokas, A. 2023. Incongruence in the phylogenomics era. Nature Reviews Genetics 24: 834-850. https://doi.org/10.1038/s41576-023-00620-x

Susko, E., Roger, A.J. 2021. Long Branch Attraction Biases in Phylogenetics. Systematic Biology 70: 838-843. https://doi.org/10.1093/sysbio/syab001

Suzuki, Y., Glazko, G.V., Nei, M. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proceedings of the National Academy of Sciences 99(25): 16138-16143. https://doi.org/10.1073/pnas.212646199

Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M. 1996. Phylogenetic Inference. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.) Molecular Systematics: 407-514. Sinauer Associates, Oxford University Press, Oxford.

Turjak, M., Trontelj, P. 2012. A method for measuring support for synapomorphy using character state distributions on phylogenetic trees. Cladistics 28(6): 627-638. https://doi.org/10.1111/j.1096-0031.2012.00403.x

Valladares, M.A., Fabres, A.A., Collado, G.A., Sáez, P.A., Méndez, M.A. 2022. Coping with dynamism: phylogenetics and phylogeographic analyses reveal cryptic diversity in *Heleobia* snails of Atacama Saltpan, Chile. Frontiers in Ecology and Evolution 10: 869626. https://doi.org/10.3389/fevo.2022.869626

Wadas, I., Domingues, I. 2025. Systematic Review of Phylogenetic Analysis Techniques for RNA Viruses Using Bioinformatics. International Journal of Molecular Sciences 26(5): 2180. https://doi.org/10.3390/ijms26052180

Wang, W., Barbetti, J., Wong, T., Thornlow, B., Corbett-Detig, R., Turakhia, Y., Lanfear, R., Minh, B.Q. 2023. DecentTree: scalable Neighbour-Joining for the genomic era. Bioinformatics 39: btad536. https://doi.org/10.1093/bioinformatics/btad536

Wiley, E.O., Lieberman, B.S. 2011. Phylogenetics: theory and practice of phylogenetic systematics. 2nd Ed. John Wiley & Sons, Hoboken.

Wiley, E.O., Brooks, D.R., Seigel-Causey, D., Funk, V.A. 1991. The compleat cladist: a primer of phylogenetic procedures. Natural History Museum, University of Kansas, USA.

Yang, Z. 2006. Computational Molecular Evolution. Oxford University Press, Oxford. 376 p.

Yang, Z., Rannala, B. 2012. Molecular phylogenetics: principles and practice. Nature Reviews Genetics 13(5): 303. https://doi.org/10.1038/nrg3186

APÉNDICE

Árbol filogenético (o filogenia): Representación gráfica que ilustra las relaciones evolutivas entre organismos, mostrando cómo descienden de ancestros comunes y se diversifican a lo largo del tiempo.

Atracción de ramas largas (Long Branch Attraction, LBA): Fenómeno en el que linajes muy divergentes (con ramas largas en el árbol) aparecen agrupados de manera incorrecta debido a un sesgo sistemático en el método de inferencia, especialmente en parsimonia y en algunos escenarios de máxima verosimilitud. Este artefacto puede generar altos valores de soporte para clados espurios.

Bootstrap: Técnica estadística que utiliza el remuestreo de los datos originales para evaluar la estabilidad y confiabilidad de los nodos en un árbol filogenético.

Cadenas de Markov: En filogenética, se refieren a secuencias de estados en un sistema probabilístico en el que la transición al siguiente estado depende únicamente del estado actual. Este principio es fundamental en métodos Bayesianos, especialmente en algoritmos de Monte Carlo por Cadenas de Markov (MCMC), que permiten explorar eficientemente el espacio de parámetros y árboles filogenéticos.

Coalescencia: Proceso en el cual las copias de un gen presentes en una población convergen en un ancestro común, reflejando la historia genealógica subyacente y siendo fundamental para modelar la variación genética a lo largo del tiempo en estudios filogenéticos.

Concatenación de secuencias: Proceso mediante el cual se unen alineamientos de diferentes marcadores genéticos en una sola secuencia larga, con el fin de aprovechar toda la información evolutiva conjunta para la reconstrucción filogenética.

Enfoque coalescente multiespecie: Método que integra el proceso de coalescencia para múltiples especies, modelando la divergencia genealógica individual de cada marcador y permitiendo inferir la historia evolutiva a nivel de población y especie, a pesar de la variación entre genes.

Escalabilidad: Capacidad de un método o software para manejar eficientemente incrementos en el tamaño de los datos y la carga computacional sin perder rendimiento, lo cual es fundamental en análisis filogenéticos y en la bioinformática para procesar grandes volúmenes de información.

Grupo monofilético: Conjunto de organismos que comparten un ancestro común, incluyendo a ese ancestro y todos sus descendientes, lo que constituye una unidad evolutiva natural.

Heterogeneidad de marcadores: Se refiere a las diferencias en las propiedades evolutivas—como tasas de evolución, patrones de sustitución y niveles de conservación—entre distintos marcadores genéticos. Esta variabilidad puede producir señales filogenéticas divergentes y afectar la congruencia y robustez de las inferencias evolutivas.

Heurística: Conjunto de estrategias y reglas prácticas utilizadas para encontrar soluciones aproximadas a problemas complejos, minimizando el tiempo y los recursos computacionales. En filogenia, las heurísticas se aplican para explorar grandes espacios de posibles árboles evolutivos cuando una búsqueda exhaustiva resulta inviable.

Homología: Característica compartida entre organismos que se debe a su descendencia de un ancestro común. Ejemplo: Las extremidades de vertebrados (como el ala de un murciélago, la aleta de una ballena y el brazo humano) son homólogas.

Homoplasia: Característica similar en diferentes organismos que surge de forma independiente, generalmente por convergencia evolutiva, y no por herencia común. Ejemplo: Las alas de insectos y las alas de aves son homoplásicas, ya que ambos grupos desarrollaron estructuras para el vuelo de manera independiente.

Linaje: Conjunto de organismos que descienden de un ancestro común, representando una línea continua de descendencia a lo largo del tiempo.

Modelos de sustitución: Son representaciones matemáticas que describen el proceso mediante el cual nucleótidos o aminoácidos se sustituyen a lo largo del tiempo. Estos modelos estiman las probabilidades y tasas de cambio entre caracteres, permitiendo ajustar los datos a un marco evolutivo y facilitar la inferencia de relaciones filogenéticas.

Monofilia: Grupo que incluye a un ancestro común y a todos sus descendientes. Ejemplo: Los mamíferos forman un grupo monofilético, ya que comprenden al ancestro común de los mamíferos y todos sus descendientes.

Parafilia: Grupo que incluye a un ancestro común, pero excluye a uno o más de sus descendientes, resultando en una representación incompleta de la historia evolutiva. Ejemplo: El grupo tradicional de reptiles es parafilético, ya que excluye a las aves, a pesar de que ambos comparten un ancestro común.

Polifilia: Grupo que agrupa organismos con características similares, pero que no comparten un ancestro

común inmediato, generalmente debido a la evolución convergente. Ejemplo: El grupo de "animales voladores" es polifilético, ya que incluye aves, murciélagos e insectos, que desarrollaron la capacidad de volar de manera independiente.

Politomía: Se refiere a un nodo interno de un árbol filogenético que presenta tres o más ramas descendientes, lo que indica una falta de resolución en las relaciones evolutivas entre los grupos implicados.

Principio auxiliar: En filogenia, se refiere al criterio que establece que los taxa deben agruparse en clados basándose en sinapomorfías, es decir, en caracteres derivados compartidos que evidencian un ancestro común. Este principio supone la homologación de dichas características, a menos que exista evidencia que indique lo contrario, y es fundamental para la reconstrucción y validación de relaciones evolutivas en los árboles filogenéticos.

Ruido filogenético: Variabilidad o información en los datos que no corresponde a la verdadera historia evolutiva, generada por errores, convergencia o saturación, y que puede distorsionar las inferencias filogenéticas.

Señal filogenética: Patrón de similitud heredada entre secuencias o caracteres que refleja la historia evolutiva común de los linajes y permite inferir relaciones de parentesco. En métodos no probabilísticos, la señal filogenética se expresa principalmente a través de sinapomorfías, mientras que en enfoques probabilísticos su recuperación depende del modelo evolutivo empleado y de su adecuación a los datos.

Separación incompleta de linajes (Incomplete Lineage Sorting, ILS): Proceso por el cual diferentes genes pueden tener historias evolutivas discordantes con la historia de especies. Esto ocurre cuando la coalescencia de los alelos precede al evento de especiación, provocando que los árboles génicos no coincidan con el árbol de especies. En filogenias multilocus o genómicas, ILS puede dar lugar a valores bajos de bootstrap en nodos que, evolutivamente, reflejan una relación real de ancestría y descendencia.

Sinapomorfía: Característica derivada compartida por dos o más taxones, que se hereda de un ancestro común y se utiliza para definir un clado. Ejemplo: La presencia de un pabellón auditivo en los mamíferos es considerada una sinapomorfía, ya que es un rasgo derivado compartido que ayuda a delimitar este grupo.

Topología: Se refiere a la estructura o patrón de ramificación de los nodos y ramas en un árbol filogenético, que representa las relaciones de parentesco entre los taxones sin considerar la longitud de las ramas.