Proyecto de tesis Valdivia, diciembre de 2024

**Historia evolutiva de las señales de comunicación acústica de las aves modernas: Integración de métodos filogenéticos y bioacústicos para la reconstrucción de estados ancestrales.**

Gabriel Morales Tello

Doctorado en Ciencias Mención Ecología y Evolución

Universidad Austral de Chile

**Menú de navegación:** [**Introducción**](#Introduccion) **|** [**Problema**](#Problema) **|** [**Investigación**](#Investigacion) **|** [**Metodología**](#Metodologia) **|** [**Referencias**](#Referencias)

**Resumen**

Las señales acústicas que utilizan las aves para comunicarse muestran una intrigante diversidad de formas que ponen en evidencia la multidimensionalidad y complejidad del rasgo. La siringe, el órgano de fonación exclusivo de las aves, cuya única función es la producción de sonido, es una prueba contundente de que la señalización acústica es un rasgo ancestral y el análisis comparativo filogenético proporciona una alternativa para la inferencia de su historia evolutiva. Se propone una reconstrucción a gran escala de las estructuras acústicas ancestrales en los nodos internos de la filogenia molecular más actual y completa de todas las aves modernas, así como una evaluación de la influencia de restricciones neuro-mecánicas del sistema de producción vocal, como el tipo de siringe y la capacidad de aprendizaje vocal. La codificación acústica reúne métodos de parametrización de bajo nivel y una extracción de características de alto nivel mediante una red neuronal previamente entrenada en más de 6000 especies de aves. El proyecto busca reconstruir las estructuras acústicas que se han conservado a través del tiempo y determinar cómo han evolucionado en los diversos linajes de aves. Un aspecto original del proyecto propone sintetizar estas estructuras ancestrales hipotéticas en espectrogramas reconstruidos con técnicas de morfometría geométrica. Adicionalmente se plantea la exploración de un método de síntesis temporal de la señal, la que puede resultar del procesamiento inverso a partir del espectrograma reconstruido en complemento con algoritmos de recuperación de fase. Los resultados de este proyecto permitirán una mejor comprensión de cómo evoluciona la morfología de un fenotipo intangible como las señales acústicas de comunicación y podrían contribuir a la emergencia de un nuevo campo de investigación en bioacústica evolutiva.

1. **Introducción**

Teorías evolutivas de la señalización acústica

Se puede definir una señal acústica producida por un organismo, como la articulación de características particulares de energía (Wiley, 2015) que se propagan en forma de ondas longitudinales a través del aire (Meyer & Neumann, 1972) y alteran el comportamiento actual o futuro de otro organismo (Wiley, 2015). Debido a este efecto, las señales acústicas han evolucionado en eficacia comunicativa gracias a que la respuesta del receptor también ha evolucionado. Esto implica que, para ser eficaz enprovocar la respuesta esperada, la señal debe ser honesta la mayor parte del tiempo, de lo contrario el receptor la ignoraría (Maynard-Smith & Harper, 2003c). Por lo tanto, este rasgo se comprende dentro de un **sistema de comunicación** integrado por emisores, receptores, un medio de transmisión y señales especialmente evolucionadas para la transferencia eficaz de información (Rubenstein & Alcock, 2019a).

Hasta hace pocos años se creía que la comunicación acústica había surgido de forma independiente en los principales clados de vertebrados terrestres, con orígenes entre hace 100 y 200 millones de años (Chen & Wiens, 2020; Kingsley et al., 2018; Russell & Bauer, 2021). Posteriormente, se ha contra argumentado con nuevos datos, que la comunicación acústica sería un **rasgo homólogo** con un origen en el ancestro común más reciente de todos los vertebrados coanatos(es decir, todos aquellos linajes de vertebrados con pulmones: tetrápodos y peces pulmonados) que data del Paleozoico hace 407 millones de años (Jorgewich-Cohen et al., 2022). Centrándonos en las aves modernas (Neornithes) hay evidencia reciente que confirma la presencia de comunicación acústica ampliamente conservada en este clado, incluyendo su ancestro común en el período Cretácico (Chen & Wiens, 2020). **Pero, cómo fueron las formas tempranas de señalización que dieron origen a los sofisticados sistemas de comunicación acústica que observamos en las aves actuales, cómo evolucionan y cómo se han diversificado, son preguntas que siguen abiertas.**

Una teoría bien aceptada para explicar el **surgimiento** de cualquier sistema de señalización es la «**ritualización**» de acciones preexistentes que individuos receptores utilizaban como guía para una acción futura (Maynard-Smith & Harper, 2003b; Tinbergen, 1952). En el caso de la señalización acústica es probable que los orígenes se encuentren durante la transición evolutiva de la vida acuática a la terrestre, cuando la vejiga natatoria adquirió una función respiratoria produciendo sonidos involuntarios que eran utilizados por receptores capaces de percibirlos como fuente de información (Morton, 2017b). Estos sonidos sin propósito que perturbaban incidentalmente el entorno habrían sido cooptados para fijar o moldear una forma reconocible a esta señal incipiente (es decir, una ritualización), confiriendo ventajas en supervivencia y éxito reproductivo (Dunlop et al., 2022). Posteriormente, las modificaciones evolutivas habrían originado nuevas especializaciones anatómicas para la fonación, así como nuevas capacidades auditivas (Dunlop et al., 2022). En consecuencia, las señales resultantes de este proceso adoptarían una forma más **conspicua, redundante y estereotipada** que el sonido original y a menudo irían precedidas de elementos o notas de **alerta** para llamar la atención sobre el mensaje que sigue (Maynard-Smith & Harper, 2003b). Estos aspectos de diseño que favorecen la eficacia de la señal están ampliamente documentados en diversos taxones, como en aves (Aubin et al., 2004; Price, 2013; Zollinger & Brumm, 2015), en ballenas (Cusano et al., 2021), en peces (Amorim et al., 2013) y en anuros (Toledo et al., 2015). **Por lo tanto, es razonable suponer que el linaje que dio origen a las aves modernas ya poseía estos aspectos de forma adquiridos durante procesos precursores como la ritualización, estableciéndose así la asociación entre emisor, mensaje y receptor** (Chen & Wiens, 2020; Jorgewich-Cohen et al., 2022)**.**

A partir de los trabajos de Nicholas Collias (1963), Eugene Morton (1977) dio una explicación de cómo las consecuencias de las vocalizaciones afectan la estructura acústica como resultado de la **evaluación de los receptores**. Es decir, la manera en cómo se evalúan las señales y se retroalimenta al remitente sería la principal fuente de selección. Este mecanismo daría como resultado nuevas o modificadas señales que optimizan la obtención de recursos, como por ejemplo al reemplazar el uso de la fuerza debido a supuestos en el ahorro de tiempo y energía, así como evitar posibles lesiones. Es así como una nueva codificación habría surgido tempranamente en arcosaurios y sinápsidos, precursores de las aves y mamíferos, respectivamente. El surgimiento de la endotermia durante el Triásico Temprano (Benton, 2021), seguida del comportamiento social complejo habrían actuado como impulsores clave de esta nueva codificación que modificaba la señalización acústica basada hasta ese entonces en simples índices de tamaño y fuerza asociados a la frecuencia del sonido. Las presiones que imponía la vida en grupo exigieron mayores niveles de complejidad en la comunicación (Morton, 2017a). En concreto, Morton propone unas **reglas generales que codifican la motivación** de un individuo en patrones de modulación de frecuencia y grado de «rugosidad» en la voz (es decir, pérdida de tonalidad). Estas reglas predicen el grado de agresividad con respecto a un cierto grado de apaciguamiento. Por ejemplo, un individuo que emite una señal rugosa y de baja frecuencia se asocia constantemente con una motivación agresiva, mientras que una señal tonal que aumenta su frecuencia codificaría una motivación pacífica (Collias, 1963; Morton, 1977, 2017a).

Así como las reglas de motivación de Morton dan una explicación evolutiva basada en los efectos en el receptor, otra explicación plausible que ha sido ampliamente estudiada en aves oscinas, pone el foco en el **emisor** y su «**performance vocal**», entendida como el grado en que se desafían las limitaciones neuromecánicas del sistema de fonación (Cardoso, 2017). Esta aproximación plantea la existencia de restricciones a la morfología acústica de la señal asociadas al control, la producción y la percepción (Podos & Nowicki, 2004); componentes del sistema de fonación que estarían bien conservados en aves oscinas y ayudarían a explicar una forma de evolución de fenotipos vocales en este suborden en particular (Düring & Elemans, 2016; Podos & Sung, 2020). Se propone un mecanismo evolutivo en el que los fenotipos vocales exigentes son favorecidos por selección intra e intersexual debido a que reflejan la alta calidad del individuo o porque estimulan de manera más efectiva los sesgos perceptivos del receptor (Podos & Sung, 2020). Las métricas para evaluar tentativamente esta cualidad se basan en observaciones espectrográficas en escala lineal de frecuencia e incluyen supuestos de restricciones biofísicas tales como: (1) compromiso en la ejecución de intervalos grandes de frecuencia entre duraciones breves de silencio (Dudouit et al., 2022; Geberzahn & Aubin, 2014), (2) dificultad para mantener elementos o sílabas estereotipadas (Sakata & Vehrencamp, 2012), (3) producción precisa del canto aprendido (Wan et al., 2024) y (4) compromisos y limitaciones asociadas a trinos rápidos (Podos, 1997; Podos et al., 2016; Sockman et al., 2024; Zollinger & Suthers, 2004. Véase tambien la Fig. 11.2 en Podos & Nowicki, 2004 y la Fig. 9.2 en Podos & Sung, 2020). La variabilidad resultante de los cantos que desafían estas restricciones daría lugar a cambios a nivel fonológico (es decir, cambios en la estructura acústica de cada elemento que compone una frase) y a nivel sintáctico (es decir, orden de los elementos que componen una frase). Estos cambios graduales, si cumplen las condiciones de heredabilidad entre generaciones, podrían explicar grandes diferencias entre especies estrechamente relacionadas –como en el caso de los gorriones del nuevo mundo (Podos & Nowicki, 2004)– en contraste con los hallazgos de alta señal filogenética (es decir, rasgos vocales supuestamente bien conservados) en representantes de la misma familia (Cicero et al., 2020).

Además de codificar la condición del individuo, las señales acústicas también pueden transmitir simultáneamente información sobre la identidad del señalizador (Wilkins et al., 2013). Si bien, la estructura que subyace a estas señales puede ser similar entre especies que comparten un ancestro común reciente, los fenotipos acústicos resultantes nunca son idénticos, de tal manera que sean suficientes para el **reconocimiento entre conespecíficos** (Gupta et al., 2021)**.** En particular, se sostiene que las señales reproductivas o «señales sexuales» cumplen un rol clave en la divergencia acústica entre poblaciones, facilitando la especiación, ya sea al principio o al final del proceso (Wilkins et al., 2013). Las señales sexuales de cada especie incipiente –especialmente aquellas que viven en simpatría– actúan como un fuerte mecanismo de aislamiento reproductivo pre-apareamiento, evitando así la hibridación comúnmente asociada a bajos valores de fitness en comparación con las especies parentales (Gupta et al., 2021). Habitualmente se sugiere que las diferencias en las señales de reconocimiento de pareja evolucionan por selección sexual dentro de las especies y luego funcionan en la capacidad de reconocimiento de especies, conectando de esta manera los procesos micro y macro evolutivos (Mendelson & Shaw, 2012; Ryan & Rand, 1993). En el contexto acústico, se ha sugerido que las presiones selectivas favorecerían la diversificación de señales sexuales en el dominio del tiempo (Gupta et al., 2021), en lugar de modificaciones en el dominio de la frecuencia como las presiones asociadas a la eficiencia energética (Riede et al., 2019). Por lo tanto, se deduce de lo anterior que, mientras evolucionan las señales sexuales en una dimensión específica, ciertas estructuras en otros dominios acústicos se conserven al finalizar el proceso de especiación.

Dos ideas íntimamente relacionadas que merecen la atención si el interés está en la evolución de señales acústicas, son la «**generalización**» y el «**peak shift**» que plantean una evolución continua fuera del equilibrio estable. Lo atractivo de estas teorías, en particular la generalización, es que sigue unas reglas básicas que son independientes de la especie, el contexto y la dimensión sensorial (Enquist et al., 2002). La generalización se puede definir como la capacidad de dar una respuesta habitual ante señales conocidas, aunque estén ligeramente modificadas (Ghirlanda & Enquist, 2003). Los experimentos sobre la respuesta ante señales nuevas sugieren que la generalización sería el comportamiento más común (Maynard-Smith & Harper, 2003b). La estocasticidad propia de la generalización crearía las condiciones para un cambio evolutivo continuo de la señal mediante peak shift. El fenómeno de peak shift ocurre cuando la señal más preferida ­va cambiando gradual y ligeramente –ya sea por herencia genética o cultural– eludiendo el equilibrio estable (Ghirlanda et al., 2002). Aunque aún no se comprende bien las formas del gradiente de cambio continuo, este mecanismo evolutivo resulta menos idealizado que los modelos de optimización y menos determinista que los modelos de estrategias evolutivamente estables basados en teoría de juegos (Enquist et al., 2002; Ghirlanda & Enquist, 2003; Maynard-Smith & Harper, 2003b).

Otra propuesta para explicar la evolución de las señales acústicas, se centra en las propiedades acústicas del hábitat y en el supuesto de que la selección natural opera para minimizar la degradación de señales utilizadas para comunicarse a distancias relativamente largas (Morton, 1975). Esta **hipótesis de adaptación acústica** predice que las vocalizaciones con frecuencias más bajas, rangos de frecuencia más estrechos e intervalos de mayor duración entre elementos, deberían ocurrir con mayor frecuencia en hábitats con vegetación densa en comparación con hábitats abiertos (Boncoraglio & Saino, 2007). La evidencia acumulada indica que hay una tendencia hacia frecuencias más bajas en los repertorios de especies de hábitats cerrados y no se hallaron diferencias significativas en la duración de los intervalos entre elementos de la señal en distintos hábitats, desestimando la hipótesis en este aspecto (Boncoraglio & Saino, 2007; Ey & Fischer, 2009; Freitas et al., 2024). Las escasas pruebas experimentales han dado resultados mayoritariamente positivos en aves y menos concluyentes en mamíferos, insectos y anuros (Hardt & Benedict, 2021). Sin embargo, el efecto sería débil y otros factores, como los costos energéticos de la fonación y la amenaza de depredadores que podrían «escuchar» la señal, deben considerarse en los modelos de evolución acústica en aves (Boncoraglio & Saino, 2007; Ey & Fischer, 2009; Hardt & Benedict, 2021).

Si bien la estructura del hábitat podría afectar la recepción de señales de largo alcance, las interferencias producidas por otras especies introducen una dimensión adicional, dando lugar a la **hipótesis de «nichos acústicos»** (Krause, 1987, 1993). Cuando las señales de diferentes individuos se superponen en frecuencia y tiempo, puede ocurrir interferencia acústica y enmascaramiento de señales, lo que reduce la capacidad del receptor para discriminar la información. Por lo tanto, se ha propuesto que las especies compiten por el espacio acústico y han desarrollado estrategias para minimizar la interferencia en sus comunicaciones, mediante selección natural sobre la estructura espectral de la señal y/o la capacidad del emisor para ajustar el momento de la señalización; cuestión especialmente relevante en comunidades con alta biodiversidad y densamente pobladas (Chek et al., 2003; Hart et al., 2021). Ciertos estudios han sugerido una segregación espectral en anuros (Allen-Ankins & Schwarzkopf, 2022; Villanueva-Rivera, 2014), una sincronización temporal como lo más frecuente en aves (Hart et al., 2021; Popp et al., 1985) o una fuerte organización entre taxones a nivel de clase (insectos, anfibios, aves y mamíferos) en bosques primarios prístinos (Monacchi, 2013). Pero, lo cierto es que hacen falta más estudios para identificar a qué escala temporal y en qué nivel taxonómico se da un soporte significativo a esta hipótesis. A la vez es necesario esclarecer la evidencia opuesta que sugiere que la comunicación entre especies en competencia puede impulsar la organización de conjuntos de señales similares en forma espectral y sincronía temporal (Tobias et al., 2014).

**Tabla 1.** Vista general de los cuerpos teóricos revisados hasta este punto, cada uno de los cuales propone una explicación parcial de la evolución de señales de comunicación acústica, desde diversos ángulos de análisis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Teoría** | **Definición** | **Mecanismo evolutivo** | **Rasgo(s) de cambio** | **Grupo** | **Referencias** |
| Ritualización | Proceso precursor de un sistema de comunicación | Selección natural | Conspicuidad, redundancia, estereotipia y elementos introductorios de alarma | Vertebrados coanatos | Amorim et al., 2013; Aubin et al., 2004; Chen & Wiens, 2020; Cusano et al., 2021; Dunlop et al., 2022; Jorgewich-Cohen et al., 2022; Maynard-Smith & Harper, 2003b; Morton, 2017b; Price, 2013; Tinbergen, 1952; Toledo et al., 2015; Zollinger & Brumm, 2015 |
| Reglas de Motivación | Regla de codificación de mensajes acústicos | Selección natural | Calidad tonal | Mamíferos y aves | Collias, 1963; Morton, 1977, 2017a |
| Performance vocal | Grado en que se desafían las limitaciones fisiológicas | Selección intra e intersexual | Fonología y sintaxis | Aves oscinas | Cardoso, 2017; Dudouit et al., 2022; Düring & Elemans, 2016; Geberzahn & Aubin, 2014; Podos, 1997; Podos et al., 2016; Podos & Nowicki, 2004; Podos & Sung, 2020; Sakata & Vehrencamp, 2012; Sockman et al., 2024; Wan et al., 2024; Zollinger & Suthers, 2004 |
| Reconocimiento entre conespecíficos | Impulsor de aislamiento reproductivo pre-apareamiento | Selección natural y sexual | Señales reproductivas en el dominio del tiempo | Anuros, aves, mamíferos e insectos | Gupta et al., 2021; Mendelson & Shaw, 2012; Ryan & Rand, 1993; Wilkins et al., 2013 |
| Peak shift | Cambio continuo del valor más favorecido de una señal en presencia de generalización | Transmisión cultural o genética de cambio gradual, ligero y no determinista | Valor máximo de un descriptor continuo de la señal | Insectos, peces, anfibios, reptiles, aves y mamíferos (humanos inclusive) | Enquist et al., 2002; Ghirlanda et al., 2002; Ghirlanda & Enquist, 2003; Maynard-Smith & Harper, 2003b, 2003a |
| Hipótesis de adaptación acústica | La estructura del hábitat influye en las señales acústicas para optimizar su transmisión | Selección natural | Frecuencia dominante, amplitud del rango de frecuencia y duración de intervalos entre elementos | Aves, mamíferos, insectos y anuros | Boncoraglio & Saino, 2007; Freitas et al., 2024; Hardt & Benedict, 2021; Morton, 1975 |
| Hipótesis de nichos acústicos | El espacio acústico es un recurso por el cual los organismos compiten y puede ser particionado tanto espectral como temporalmente | Selección natural | Rango de frecuencia y sincronía temporal | Aves, mamíferos, insectos y anuros | Allen-Ankins & Schwarzkopf, 2022; Chek et al., 2003; Hart et al., 2021; Krause, 1987, 1993; Monacchi, 2013; Popp et al., 1985; Tobias et al., 2014; Villanueva-Rivera, 2014 |

Habiendo revisado las principales propuestas que intentan explicar el origen y evolución de las señales acústicas, resulta evidente notar que el fenómeno es **difícil de predecir**, debido a que la estructura acústica evoluciona en un espacio multidimensional de rasgos tangibles e intangibles y enfrenta fuerzas evolutivas de naturaleza dispar ([Tabla 1](#Tabla1)). Mientras tanto, permanece indescifrable una necesaria integración o jerarquización de teorías, procesos y mecanismos que facilite una comprensión a gran escala de la evolución de señales acústicas, en una de las clases animales mejor estudiadas y monitoreadas por la ciencia: las aves (Fraixedas et al., 2020).

Retomando las preguntas planteadas anteriormente sobre las formas tempranas de señalización y cómo se han diversificado, es necesario introducir dos aspectos más en el análisis para una comprensión íntegra del problema. Estos aspectos representan las principales capacidades y restricciones neuro-mecánicas del sistema de producción vocal: (1) la capacidad de **aprendizaje vocal** como un rasgo determinante en la evolución acústica de ciertos clados (Jarvis, 2004) y (2) las restricciones impuestas por la **siringe como fuente de sonido** exclusiva de las aves modernas (Ten Cate, 2004). A continuación, se resumen estos rasgos desde un punto de vista evolutivo.

Capacidades y restricciones neuro-mecánicas del sistema de producción vocal

Ciertos clados aviares adquieren parte de su repertorio mediante **aprendizaje vocal**. Es decir, mediante imitación y no por instinto, a diferencia del aprendizaje auditivo, que es la capacidad de establecer asociaciones con los sonidos escuchados (Jarvis, 2004). El aprendizaje vocal es un rasgo que ha sido objeto de numerosos estudios que intentan resolver su distribución en la filogenia y comprender cómo pueden haber evolucionado los procesos y circuitos neuronales subyacentes (ten Cate, 2021). La hipótesis tradicional sostiene la presencia bien extendida de este rasgo en al menos tres grupos: loros, colibríes y aves oscinas (Rubenstein & Alcock, 2019b), y estudios comparativos han mostrado algunas homologías en circuitos neuronales y en la red de genes que subyacen al aprendizaje vocal en estos tres grupos (Jarvis, 2004). En lo que respecta a la evolución de las señales acústicas, se ha sugerido que el aprendizaje vocal **proporciona un medio rápido para la acumulación y transferencia de nuevas variantes dentro de las poblaciones** (Wilkins et al., 2013), conduciendo a la aceleración de procesos de divergencia acústica y especiación alopátrica (Lachlan & Servedio, 2004). Sin embargo, se ha demostrado que el fenotipo acústico resultante del aprendizaje no está determinado exclusivamente por herencia cultural, sino más bien habría una predisposición innata (es decir, heredada genéticamente) que guía el aprendizaje hacia el canto final desarrollado (Mets & Brainard, 2018; Rajan et al., 2024). Esto revela que, dependiendo de cómo interactúen la transmisión cultural y las predisposiciones innatas, podría ocurrir una evolución rápida del fenotipo acústico o, por el contrario, la divergencia acústica entre poblaciones podría frenarse. Esto ocurriría si el aprendizaje sucede después de un evento de dispersión, puesto que las predisposiciones genéticas podrían mantener ciertas estructuras acústicas comunes, facilitando la hibridación entre especies incipientes en contacto secundario (Wilkins et al., 2013). En resumen, el contexto, es decir, cómo y cuándo se desarrolla el aprendizaje, marca una diferencia significativa en el resultado evolutivo (Verzijden et al., 2012).

Por otro lado, la estructura de una señal acústica está inevitablemente limitada por el mecanismo generador de sonido. En las aves, esto corresponde a un sistema dinámico capaz de modular la presión de aire (fuente de energía), la tensión de labios y membranas en la siringe (fuente de sonido) y la morfología del tracto vocal superior (filtro) (Düring & Elemans, 2016). En particular uno de estos subsistemas, **la siringe**, ha demostrado ser útil en análisis filogenéticos basados en caracteres morfológicos para resolver relaciones de parentesco en distintos niveles taxonómicos (por ejemplo, en la familia Rhinocryptidae: (Mauricio et al., 2012), en el orden Falconiformes: (Griffiths, 1994) y en el clado de divergencia basal Palaeognathae: (McInerney et al., 2019). Sin embargo, nuestro conocimiento sobre sus orígenes es aún muy incompleto (Goller, 2022) y su relevancia en el estudio de las señales acústicas ancestrales es ineludible.

Todos los vertebrados terrestres, incluidos los parientes vivos más cercanos de las aves, los cocodrilos, utilizan la laringe como fuente de sonido, excepto las aves que usan la siringe (Goller, 2022). La posición de la siringe en la unión traqueobronquial, a diferencia de la laringe ubicada en el extremo superior, confiere ventajas en la eficiencia para la conversión de energía aerodinámica en energía acústica. Así lo sugiere un modelo físico y una simulación computacional que demostró mejoras en la eficiencia energética cuando la fuente de sonido se ubicaba en la base de la tráquea facilitando la sintonía entre la frecuencia fundamental y la primera resonancia del tracto vocal (Riede et al., 2019). Esta distribución de la energía en el dominio de la frecuencia refleja un aspecto estructural del órgano vocal y se define como «calidad tonal» (Nowicki & Marler, 1988; Peters & Nowicki, 1996). La calidad tonal se caracteriza desde estructuras de tono puro con gran parte de la energía concentrada en el armónico fundamental, también llamadas silbidos, hasta estructuras con mayor distribución espectral, es decir, con mayor riqueza de armónicos (Ames, 1971; Hoese et al., 2000). Las señales de alta pureza tonal son comunes y extendidas entre las aves. Se ha sugerido que esta estructura acústica sería el resultado de un mecanismo vibratorio que genera armónicos multifrecuencia en la fuente de sonido y luego filtrados en mayor o menor medida por un complejo sistema de cavidades con un comportamiento resonante difícil de predecir (Beckers et al., 2003; Fletcher & Tarnopolsky, 1999; Nowicki & Marler, 1988). Sin embargo, según un modelo físico en zebra finch (Taeniopygia guttata), habría un grado variable de tonalidad que emerge de la fuente del sonido (Sitt et al., 2008). Lo cierto es que la calidad tonal y, en términos generales, la articulación de formas acústicas depende de la interacción dinámica entre los componentes de la estructura siríngea, cuya morfología y ubicación varían significativamente entre los diversos linajes (Düring & Elemans, 2016).

Los clados aviares desarrollan estructuras cartilaginosas especializadas, varios músculos que pueden ser intrínsecos o extrínsecos, modificaciones de los anillos traqueales y bronquiales y variación en la cantidad de membranas como fuente de sonido (Chiappone et al., 2024; Larsen & Goller, 2002). Se reconocen **cuatro tipos de siringe** según la ubicación y cantidad de labios y membranas oscilatorias. La estructura siríngea más común es la traqueobronquial con dos fuentes de sonido, propia del suborden Oscine. La siringe traqueal con una sola fuente de sonido ubicada entre los anillos traqueales se observa en Columbiformes y algunos cicónidos (orden Pelecaniformes). Una tercera estructura siríngea con dos fuentes localizadas entre los anillos bronquiales, la podemos encontrar en Caprimulgiformes, Cuculiformes y en algunos Strigiformes (Mindlin & Laje, 2005). Finalmente, un cuarto tipo de siringe se ha descrito en paseriformes suboscinas de la familia Thamnophilidae y Furnariidae, pertenecientes al infraorden Furnariides. Un par de membranas traqueales que interactúa con dos pares de labios bronquiales conforman una estructura de tres fuentes de sonido (Garcia et al., 2017, revisado en Amador & Mindlin, 2023).

Pero, ¿qué sabemos sobre la estructura de la siringe ancestral?. Las aves modernas abarcan aproximadamente 11 mil especies en la actualidad y se dividen en dos clados hermanos: Paleognathae y Neognathae (Stiller et al., 2024; Widrig & Field, 2022). Paleognathae incluye 60 especies, como el avestruz, el emú y los tinamúes, mientras que Neognathae abarca más de 10 mil especies. A menudo se argumenta que, para comprender las características ancestrales de todas las aves actuales, es crucial estudiar la evolución de los rasgos en los linajes de divergencia temprana dentro de Paleognathae (Chiappone et al., 2024). El avestruz, taxón hermano de todos los demás paleognatos, posee una siringe prominente con una única fuente de sonido y un amplio repertorio con vocalizaciones que podrían implicar una superposición funcional entre laringe y siringe (Chiappone et al., 2024). En otros paleognatos como el ñandú grande (Rhea americana) se encontró dos fuentes de sonido de tipo traqueobronquial en contraste con una especie de tinamú (Eudromia elegans) con solo una (Chiappone et al., 2024).

Los únicos restos conocidos de una siringe fósil corresponden a un ejemplar recuperado de una parte de la radiación de aves vivas del Cretácico Superior (hace aprox. 66 a 69 millones de años) de la Antártica (Clarke et al., 2016). La transición de la fuente de sonido basada en la laringe a una siríngea ubicada en la unión traqueobronquial, habría ocurrido dentro de Dinosauria, antes o en el origen de Neornithes (Clarke et al., 2016; Kingsley et al., 2018) y, según el mismo informe paleontológico, una siringe compleja surgió mucho después del origen del vuelo y las innovaciones respiratorias. En particular, proponen una siringe ancestral con dos fuentes de sonido en línea con lo expuesto por Goller (2022). No obstante, los hallazgos de una única fuente de sonido en el avestruz con labios laterales prominentes y una membrana timpaniforme extensa, características poco comunes en las aves, sugieren la necesidad de reevaluar el número de fuentes de sonido ancestral (Chiappone et al., 2024), puesto que las capacidades acústicas estarían íntimamente ligadas a las radiaciones evolutivas tempranas de las aves (Goller, 2022). Una reevaluación que optimice los estados de caracter en toda la filogenia aviar esclarecería la presunción de que la siringe de algunos paleognatos como el avestruz y el casuario del sur (Casuarius casuarius) sería regresiva en lugar de basal (Düring & Elemans, 2016).

Otro paso evolutivo en la producción vocal, además del número de fuentes de sonido, corresponde a la adición de control neuromuscular fino. Como ya se mencionó, en algunos linajes de aves surgió varias veces de manera independiente una expansión neuronal, denominada aprendizaje vocal. Un elaborado circuito en el cerebro anterior y en el tálamo evolucionó para adquirir sonidos después del nacimiento y luego desarrollar las habilidades motoras para reproducirlos (Goller, 2022). La gran variabilidad acústica que se observa en uno de estos grupos aprendices, el suborden Oscine, contrasta con una siringe altamente conservada en todas las especies del grupo. Evidentemente, el factor más importante en la diversificación vocal en las aves oscinas es el control neuromuscular rico y diverso, en lugar de la estructura siríngea (Amador & Mindlin, 2023; Ames, 1971). En paralelo, las paseriformes suboscinas no aprendices, también muestran una diversificación vocal extendida en el grupo, pero, en contraste con su clado hermano, esta variabilidad es producto de una diversificación morfológica y funcional de la siringe (Goller et al., 2021).

De este modo, al integrar estos dos aspectos junto a los procesos evolutivos mencionados anteriormente ([Tabla 1](#Tabla1)), es posible obtener una perspectiva general de los factores que influyen en la evolución de las señales acústicas en aves. El marco conceptual se completa al entender las limitaciones morfológicas y las capacidades de aprendizaje como factores o rasgos que guían la dirección y velocidad evolutiva. Comprender cómo influyen estos dos aspectos a lo largo del tiempo y entre linajes nos permitiría instalar bases sólidas en la discusión que acompañe a una reconstrucción de formas acústicas tempranas y la complejidad de su evolución.

1. **Problema**

Reconstrucción de estados ancestrales

Las aves –al igual que todas las especies– llevan consigo una historia filogenética que se refleja en similitudes entre especies estrechamente emparentadas debido a características heredadas de un ancestro común. Por supuesto estas características pueden ser fenotipos acústicos susceptibles de ser parametrizados y cuantificados (Ten Cate, 2004). Aunque el estudio de la evolución acústica no puede beneficiarse directamente del registro fósil como en otras morfologías tangibles, se pueden obtener algunas inferencias sobre el pasado a partir del registro de audio de las aves actuales, mediante **métodos comparativos filogenéticos** (Morton, 2017b; Revell & Harmon, 2022). La omnipresencia de la siringe, cuya única función es la producción de sonido, es una prueba contundente de que se trata de un rasgo ancestral y el análisis comparativo proporciona una alternativa para la reconstrucción de la historia evolutiva de las señales acústicas aviares (Ten Cate, 2004).

Los análisis filogenéticos comparativos han facilitado el estudio de procesos evolutivos responsables de la diversidad biológica y se están volviendo populares en estudios sobre la evolución de señales, incluidas estructuras acústicas (Odom et al., 2021). En particular, un enfoque bien extendido y en constante actualización es el método conocido como **Reconstrucción de Estados Ancestrales**, el cual permite trazar los estados desconocidos de un rasgo fenotípico sobre árboles filogenéticos inferidos a partir de datos moleculares y/o morfológicos distintos al rasgo de interés (Revell, 2024a). De esta manera, por ejemplo, trabajando con filogenias moleculares, se controla la falta de independencia estadística causada por la historia evolutiva compartida (Felsenstein, 1985). El mapa trazado permite observar qué ramas del árbol dieron lugar a qué tipo de transiciones evolutivas y revelar, por ejemplo, componentes lábiles y estables de la estructura del rasgo (Price & Lanyon, 2002).

**Existe evidencia sólida de una dependencia estadística significativa entre las variables acústicas de los sistemas de comunicación vocal de las aves**, como consecuencia de sus relaciones filogenéticas y de los diversos procesos evolutivos implicados (Arato & Fitch, 2021; McCracken & Sheldon, 1997; Price, 2004; Price & Lanyon, 2002; Seddon, 2005). Ciertos componentes de la estructura acústica han probado ser filogenéticamente informativos en grupos taxonómicos diversos, como en paseriformes oscinas y suboscinas (Price & Lanyon, 2002; Seddon, 2005), pelecaniformes (garzas) (McCracken & Sheldon, 1997), columbiformes (tórtolas) (de Kort & ten Cate, 2004) y en psitaciformes (loros) (Medina‐García et al., 2015). El único trabajo que ha abordado el problema en especies de diferente orden taxonómico, bajo un mismo criterio de análisis, es el trabajo de Arato y Fitch (2021) quienes mostraron **evidencia de una fuerte señal filogenética** entre especies paseriformes y no paseriformes. Sin embargo, **una reconstrucción de estados ancestrales a gran escala, por ejemplo con datos representativos a nivel de familia de toda la clase Aves, continúa pendiente**. Quizás existe una falta de interés debido a la evidencia de cambios rápidos por evolución cultural y variación geográfica en poblaciones de especies oscinas ampliamente estudiadas (Ju et al., 2019; Nelson et al., 2004; Riebel et al., 2015; Williams et al., 2022), puesto que las altas tasas de cambio podrían aumentar las posibilidades de pérdida de señal filogenética (Odom et al., 2021). No obstante, los análisis interespecíficos en algunos clados aprendices han revelado aspectos del canto bien conservados (Price & Lanyon, 2002) y tasas evolutivas similares entre fenotipos acústicos y morfológicos (Medina‐García et al., 2015). **Si bien, estos hallazgos relajan los supuestos de un fenotipo evolutivamente lábil y proporcionan confianza para un análisis a gran escala, persisten otros desafíos respecto a la codificación acústica que se describen a continuación.**

Codificación acústica

En cualquier método comparativo filogenético se requiere una **hipótesis de homología primaria** que debe probarse en el análisis posterior. Los criterios habituales para homologar rasgos fenotípicos son basados en la anatomía comparada y el primer y más importante criterio es la similitud: las estructuras deben corresponder en posición y detalles estructurales (King & Rücklin, 2020). Este criterio de similitud estructural se ha discutido poco respecto a fenotipos intangibles como el sonido. Sin embargo, hay certeza de que, al menos el órgano generador de sonido es un rasgo homólogo en las aves (Ten Cate, 2004). McCracken y Sheldon (1997) discutieron este problema a la luz de comparar graznidos de una especie con una llamada de vuelo en otra. La solución propuesta fue evitar categorizaciones fonéticas y aprovechar las características cuantitativas del sonido (por ejemplo, frecuencia y energía), que permiten comparar señales de diversa naturaleza, como ruidos simples que son en gran medida función de la morfología siríngea o sonidos más complejos que presentan un mayor componente conductual (McCracken & Sheldon, 1997). Esta estrategia ha sido implementada en trabajos anteriores, pero los criterios fonéticos persisten parcialmente al extraer caracteres que implican fragmentar arbitrariamente la señal. Estas medidas incluyen caracteres discretos como, número de sílabas y estructura silábica binaria (McCracken & Sheldon, 1997), caracteres continuos de unidades fonéticas como, frecuencia máxima de cada nota y duración de cada nota (Seddon, 2005) o la transformación de caracteres continuos a caracteres binarios o de pocos estados (Price & Lanyon, 2002). **Estas caracterizaciones han sido fructíferas en comparaciones de especies estrechamente emparentadas que exhiben caracteres acústicos que pueden ser delimitados superficialmente mediante la inspección visual de espectrogramas. Sin embargo, ciertos caracteres extraídos bajo criterio fonético no son aplicables en comparaciones a gran escala y se deben descartar en la propuesta de homología primaria del rasgo acústico aviar en su conjunto.**

Una estrategia más universal la encontramos en codificaciones mediante parámetros tradicionales del sonido que han sido validadas para evaluar hipótesis filogenéticas en grupos taxonómicos diversos (Arato & Fitch, 2021; Gingras et al., 2013; Gingras & Fitch, 2013). La técnica se conoce como **parametrización acústica de bajo nivel**. Es decir, características básicas y directas de la señal de audio (Mojsilovic & Rogowitz, 2001) que suelen extraerse mediante algoritmos como la *Transformada de Fourier de Tiempo Corto* (*STFT*, por su sigla en inglés). Algunas bien comunes son el centroide espectral (media ponderada por amplitud de las frecuencias presentes en la señal de audio), frecuencia dominante (punto de máxima amplitud a lo largo del espectro de frecuencias), tasa de cruce por cero (velocidad a la que la amplitud de la señal de audio pasa de positiva a negativa o viceversa) y la tonalidad espectral (relación entre la media geométrica y la media aritmética, indica si un espectro de frecuencias es plano o contiene máximos definidos), entre otras (Gingras & Fitch, 2013). **Estos parámetros de bajo nivel demostraron ser eficaces descriptores de señal filogenética y, por lo tanto, probaron ser buenos candidatos a caracteres acústicos homólogos en un marco taxonómico relativamente amplio.**

Si queremos aumentar la escala del análisis a toda una clase de vertebrados, como las aves, el nivel de complejidad en el conjunto de datos exige una extracción de características de mayor nivel que conserve la universalidad de una parametrización de bajo nivel y capture el contenido perceptual de la señal. Una codificación así debería extraer información estructural y fina asociada a la morfología de la fuente de sonido (siringe) y a la neurofisiología del sistema de fonación en su conjunto, así como los patrones de modulación tiempo-frecuencia que han evolucionado para el reconocimiento entre conespecíficos u otro mecanismo neutral o adaptativo ([Tabla 1](#Tabla1)). Estos requerimientos de mayor nivel ya están siendo mencionados y se intuye un buen camino a seguir usando Machine Learning (Arato & Fitch, 2021; Pearse et al., 2018; Provost et al., 2022).

En la actualidad están disponibles técnicas de **extracción de características de alto nivel** que codifican estructuras complejas en un espacio multidimensional a partir de parámetros acústicos de bajo nivel. En particular, un flujo de procesamiento que ofrece un enorme potencial, es aquel que trabaja sobre espectrogramas tratados como una imagen y extrae características de alto nivel mediante una red neuronal convolucional (McGinn et al., 2023). La arquitectura de una red neuronal se basa en una configuración por capas en donde los parámetros internos de cada capa se deben ajustar mediante un proceso de entrenamiento que suele ser computacionalmente costoso y requiere una gran cantidad y diversidad de datos (LeCun et al., 2015). La buena noticia es que ya están disponibles modelos previamente entrenados en conjuntos de datos de enorme tamaño y evaluados con altos puntajes de precisión y sensibilidad. Por ejemplo, BirdNet, una red neuronal convolucional, optimizada para identificar más de seis mil especies de aves de todos los órdenes a partir de imagen espectrográfica de las señales acústicas (Kahl et al., 2021). Además de disponer en su salida un clasificador, es decir, una capa final que contiene las probabilidades de presencia de las especies, una red como ésta permite extraer la penúltima capa como un vector que contiene una codificación multidimensional de alto nivel (McGinn et al., 2023). El «alto nivel» quiere decir que es el resultado de una parametrización por capas que captura patrones desde los más simples y directos, hasta los más complejos y abstractos de la señal, tal como lo hacemos con nuestra corteza visual para aprender, reconocer y finalmente generalizar la identificación y clasificación de objetos nunca antes vistos (LeCun et al., 2015). La extracción de estas características de alto nivel está disponible en unas pocas líneas de código y su potencial está impulsando nuevas tareas bioacústicas como, por ejemplo, la detección de especies crípticas. Sin embargo, a la fecha no se han utilizado para responder preguntas de relevancia evolutiva, aunque los autores de BirdNet han sugerido en su repositorio, aplicabilidad en análisis de similitud (<https://github.com/kahst/BirdNET-Analyzer>).

Si bien, una codificación de este tipo suena tentadora, aún tenemos el **problema de la alta dimensionalidad**. Un vector de características obtenido con BirdNet, puede llegar a tener 1024 dimensiones que corresponden a características profundamente procesadas a partir de un espectrograma de 3 segundos en un rango de frecuencia de 0 a 15 kHz (McGinn et al., 2023). Si bien, esta dimensionalidad es dramáticamente inferior que la de una extracción de características directa del mismo espectrograma –que podría llegar a tener 67584 dimensiones (con tasa de muestreo de 48 Khz, 2048 puntos FFT, ventana temporal de 512 muestras, overlap al 25% y 64 bandas Mel)– persiste el problema de la alta dimensionalidad.

Habitualmente, para analizar espacios de alta dimensión y complejidad, se utilizan técnicas de aprendizaje no supervisado, como por ejemplo la técnica UMAP (Uniform Manifold Approximation and Projection), que reduce la alta dimensión preservando la estructura global y local de los datos de manera no lineal (McInnes et al., 2020). La figura 1 muestra una comparación visual de dos espacios de características reducidos con UMAP a solo 3 dimensiones. La optimización lograda con BirdNet seguida de UMAP, es notable. Las características obtenidas con BirdNet, extraen una síntesis optimizada de la estructura de la señal que se refleja en una alta separabilidad de los datos. El agrupamiento proyectado en un espacio de baja dimensión resulta parcialmente consistente con la organización taxonómica de las señales acústicas seleccionadas para el análisis, **razón principal para escoger este espacio vectorial como un escenario ideal para el análisis comparativo**.

**Figura 1.** Proyección UMAP de dos espacios de características:

1. [UMAP 3D a partir de características directas de un espectrograma en escala Mel-log.](STFT_embedding_3d.html)
2. [UMAP 3D a partir de características complejas extraídas de la penúltima capa de BirdNet.](BirdNet_embedding_3d.html)

Los datos corresponden a 3110 señales acústicas de 15 especies de la familia Strigidae (Buhos): 5 del género Bubo, 5 del género Glaucidium y 5 del género Otus. Se colorean los datos según género para observar agrupamiento en este nivel. Cada dato (es decir, cada punto en el espacio) es un segmento de 3 segundos extraído de un audio descargado de Xeno-Canto (https://xeno-canto.org). Cada segmento contiene la señal de una única especie en primer plano con una alta variabilidad de relación señal-ruido. Todos los segmentos fueron re-muestreados a una tasa de 48 Khz y 16 bits de cuantización en PCM-WAV.

Nota: Para visualizar las figuras, se adjuntan dos archivos html que deben estar en la misma carpeta de este documento, al pinchar el enlace se abre el navegador predeterminado con una figura interactiva.

Un desafío importante que surge al trabajar en un espacio de características de alto nivel es la interpretabilidad de las regiones o clusters que alojan los datos (puntos en el espacio), ya sea las coordenadas asociadas a datos acústicos de aves actuales o aquellas coordenadas ancestrales reconstruidas con métodos filogenéticos. La pregunta es, **¿cómo interpretamos las coordenadas reconstruidas en este espacio de características?**. Si bien, el espacio UMAP corresponde a un espacio de distancias euclidianas, las relaciones con el espacio original de los datos son no lineales (McInnes et al., 2020) y plantea una dificultad mayor si quisiéramos sintetizar una señal ancestral en formato de audio, o al menos en un espectrograma, a partir de puntos en un espacio de dimensiones reducidas. Sin embargo, existe un original enfoque de síntesis de señales bioacústicas que trabaja con análisis morfométrico de espectrogramas tridimensionales y análisis de componentes principales (PCA). El método está disponible en el paquete R *SoundShape* (Rocha & Romano, 2021) y facilita la reconstrucción de formas hipotéticas entre datos acústicos homólogos según las desviaciones de sus componentes principales. Esto significa que se podría implementar un flujo de procesamiento prometedor que dé como resultado una síntesis en audio de estructuras acústicas ancestrales. Tales inferencias evolutivas han sido motivo de investigaciones a gran escala en rasgos morfológicos, como se demostró en la reconstrucción de la flor ancestral de las angiospermas (Sauquet et al., 2017) y representan un enfoque inspirador en el presente proyecto.

En resumen, están dadas las condiciones conceptuales, técnicas y metodológicas para reconstruir aquellos caracteres vocales conservados en las aves actuales, que estaban presente en aquellos nodos que dieron origen a los principales linajes. La integración de técnicas de aprendizaje supervisado y no supervisado, métodos comparativos filogenéticos y análisis morfométrico,proporciona una aproximación plausible al problema planteado. Más aún, la aplicación de este flujo bioinformático podría revelar aspectos ocultos de la historia evolutiva de la señalización vocal. Es decir, un mapa que indique novedades, pérdidas y cambios que han ocurrido en la estructura de las señales acústicas durante la evolución de las aves modernas (Ten Cate, 2004).

1. **Investigación**

**Estructuración científica: generalidades**

Pregunta general

¿Cómo eran y cómo han evolucionado las señales acústicas de comunicación en las aves?

Hipótesis general

Al menos una parte de las señales de comunicación vocal en los orígenes de las aves modernas (Neornithes) se ha conservado y está presente en las estructuras acústicas que observamos en las aves actuales. Esta estructura es compleja y multidimensional. Sin embargo, una codificación de alto nivel puede revelar caracteres estructurales que no son advertidos por una parametrización fragmentada de la señal. Por lo tanto, podemos reconstruirla en nodos ancestrales y mapear los cambios en filogenias moleculares bien soportadas. Una caracterización de alto nivel explicaría mejor el fenotipo acústico y potencialmente podría extraer información de la neuro-mecánica del sistema de fonación. Una reconstrucción que clarifique homologías, homoplasias y tasas de cambio, en relación con ciertas capacidades y restricciones conocidas, nos ayudaría a comprender mejor la evolución de fenotipos intangibles como las señales acústicas de comunicación, así como las fuerzas evolutivas que conducen el cambio.

Objetivo General

**Reconstruir** la historia evolutiva de las señales de comunicación acústica en las aves modernas, desde los orígenes del grupo corona hasta la actualidad, integrando el análisis comparativo filogenético, codificación acústica de alto nivel y técnicas de síntesis espectrográfica y auditiva.

**Estructuración científica: desglose**

**SECCIÓN 1: Codificación acústica**

¿Hay señal filogenética en los fenotipos acústicos presentes en la comunicación aviar y cómo podemos codificar convenientemente este rasgo para extraer el patrón filogenético en los datos?

**Hipótesis 1.1:** Dentro de las estructuras acústicas de las aves actuales, encontramos componentes que se han conservado a lo largo de la evolución y diversificación del clado. Estos componentes están codificados en un nivel perceptual que integra múltiples dimensiones acústicas.

**Predicción 1.1:** Deberíamos encontrar mayores valores de señal filogenética bajo una caracterización acústica de alto nivel que capture aquellos componentes perceptuales ancestrales que se han conservado y están presente en las señales acústicas de las aves actuales.

**Objetivo específico 1.1:** Evaluar señal filogenética comparando codificaciones de bajo y alto nivel para obtener el mejor criterio de caracterización acústica que refleje el patrón filogenético en los datos.

**SECCIÓN 2: Reconstrucción a gran escala**

¿Cómo fueron las señales acústicas en los nodos ancestrales que dieron origen a los principales linajes de las aves modernas? ¿Es posible reconstruir esas estructuras acústicas a partir de un conjunto de datos a gran escala de las aves actuales?

**Hipótesis 2.1:** Las estructuras acústicas en los nodos profundos de las aves modernas poseían componentes ritualizados por procesos precursores del sistema de comunicación que ya estaban presentes en el linaje que dio origen a las aves modernas.

**Predicción 2.1:** Las estructuras reconstruidas en los nodos más profundos no deberían ser radicalmente diferentes a las señales de las aves actuales y deberían contener componentes relativamente sofisticados, lejos de ser simples sonidos respiratorios sin función comunicativa.

**Objetivo específico 2.1:** Reconstruir los estados ancestrales en la filogenia más actual, bien soportada y de gran escala de las aves modernas, a partir de la caracterización acústica determinada en la sección anterior.

**Objetivo específico 2.2:** Sintetizar versiones espectrográficas y auditivas de los componentes ancestrales reconstruidos.

**Objetivo específico 2.3:** Determinar los componentes ritualizados o sus rastros revelados en la reconstrucción y síntesis ancestral, así como la presencia en los repertorios de las aves actuales.

**SECCIÓN 3: Restricciones evolutivas**

¿Cómo han evolucionado estas señales ancestrales en los diversos linajes y cuánto han influido ciertas restricciones neuro-mecánicas del sistema de fonación aviar, como el aprendizaje vocal o el tipo de siringe?

**Hipótesis 3.1:** Existen restricciones neuro-mecánicas en el sistema de producción vocal de las aves modernas, como el tipo de siringe y la capacidad de aprendizaje vocal, que determinan la dirección y tasa de cambio evolutivo, respectivamente.

**Predicción 3.1:** Es esperable observar consistencias entre linajes que comparten restricciones morfológicas de la siringe, respecto a tasas y correlación entre caracteres multidimensionales de alto nivel.

**Predicción 3.2:** Es esperable observar mayores tasas de cambio en los clados que poseen la capacidad de aprendizaje vocal respecto a grupos que producen sus señales acústicas de forma innata. Sin embargo, a la luz de evidencias contradictorias, habrían simultáneamente otras dimensiones acústicas desacopladas y resistentes al cambio.

**Objetivo específico 3.1:** Reconstruir los tipos de siringe en la filogenia de las aves, codificando el número de fuentes de sonido de las familias taxonómicas muestreadas.

**Objetivo específico 3.2:** Evaluar influencia de las restricciones morfológicas de la siringe en la evolución acústica de las aves mediante un análisis simultáneo de tasas y correlación de dimensiones sobre una filogenia previamente mapeada con los tipos de siringe reconstruidos.

**Objetivo específico 3.3:** Evaluar influencia del aprendizaje vocal mediante un análisis de tasas y correlación de caracteres de bajo nivel sobre filogenias de grupos aprendices y no aprendices.

1. **Metodología**

A continuación se describen las técnicas y métodos que se implementarán para alcanzar cada objetivo específico según el desglose anterior.

**SECCIÓN 1: Codificación acústica**

**Objetivo específico 1.1:** **Evaluar señal filogenética comparando codificaciones de bajo y alto nivel para obtener el mejor criterio de caracterización acústica que refleje el patrón filogenético en los datos.**

Como ya se mencionó, hay dos estrategias (bajo y alto nivel) que describen apropiadamente el patrón filogenético en un conjunto de datos acústicos a gran escala, siempre y cuando se trate a la señal de manera íntegra en su composición espectro-temporal, sin fragmentaciones fonéticas arbitrarias. Pero antes, debemos escoger una filogenia actualizada y construir el conjunto de datos.

Conjunto de datos

Las aves modernas (*Neornithes*) son el grupo corona de todas las aves actuales, para trabajar con esta clase completa, haremos uso de la filogenia molecular ultramétrica (calibrada en el tiempo) de Stiller et al. (2024), derivada con métodos coalescentes a partir de 63430 loci intergénicos seleccionados de alineaciones del genoma completo de 363 especies que representan a 218 familias taxonómicas (92% del total). Los scripts generadores de los árboles de especies calibrados en el tiempo, están disponibles en el repositorio público del autor (Stiller et al., 2024).

Los datos acústicos de las mismas 363 especies serán descargados de la colección Xeno-Canto ([https://xeno-canto.org](https://xeno-canto.org/)) mediante la librería *Python XenoPy* (<https://github.com/realzza/xenopy>). Se descargarán absolutamente todas las grabaciones disponibles de cada especie. Sin embargo, se seleccionará solo una grabación por especie para los análisis, según el siguiente **proceso que garantiza un muestreo estadístico representativo intra-especie**:

1. Escoger uno o varios *templates* (muestras de referencia) según tamaño de repertorio disponible, que garanticen buena relación señal-ruido.
2. Ejecutar algoritmo propio de correlación cruzada para buscar coincidencias con el *template* dentro de los audios descargados.
3. Ajustar las coincidencias a una duración fija de 3 segundos; requisito de entrada para el siguiente proceso (*BirdNet*) y criterio de estandarización temporal para todo el conjunto de datos.
4. Extraer características de alto nivel con *BirdNet* para cada segmento de 3 segundos.
5. Proyectar todas las coincidencias en un espacio de similitud bidimensional con técnica *UMAP*.
6. Analizar el agrupamiento con algoritmo *DBSCAN* (Schubert et al., 2017)
7. Seleccionar una muestra cercana al centroide (media) para clusters homogéneos y cercana a la mediana para clusters con valores atípicos significativos.

Preprocesamiento

Las 363 grabaciones seleccionadas serán re-muestreadas a una tasa de 48 kHz, cuantización de 16 bit y un filtrado pasa-banda entre 150 Hz y 15 kHz para reducir interferencias ajenas al rango de las aves (protocolo *BirdNet*). Con el ánimo de reducir la variabilidad en las técnicas y contextos de grabación (nivel de entrada, distancia del micrófono, potencia vocal, etc) se realizará una estandarización estadística normalizando los valores de amplitud en la señal discreta (serie de tiempo) a una media igual a 0 y una desviación estándar igual a 1,0 mediante la siguiente relación:

Donde,

: señal estandarizada

: señal original

: índice temporal de cada valor de amplitud

: media aritmética de todas las amplitudes de

: desviación estándar de todas las amplitudes de

Esto asegura que las diferencias en los niveles de amplitud no influyan en el análisis.

Luego se transformarán las señales discretas a un espacio tiempo-frecuencia (espectrograma) según *Transformada de Fourier de Tiempo Corto (STFT,* por su sigla en inglés*)* con el siguiente ajuste de parámetros: ventana temporal de 1024 muestras (~ 0,021 seg) con 50 % de superposición y 2048 puntos *FFT* con ventana *Hann*. Los espectros de magnitud cuadrática se mapearán en escala de frecuencia Mel utilizando un banco de filtros de 128 bandas. Luego se convertirán a unidades normalizadas de decibeles con referencia a la magnitud cuadrática máxima de cada espectrograma. Este ajuste sigue el trabajo previo de (Morales et al., 2022).

Extracción de características de bajo nivel

Se proponen cinco características acústicas de bajo nivel, sensibles a diversos atributos espectrales y temporales de la señal, siguiendo la parametrización de Arato & Fitch (2021):

1. Coeficiente de variación de la amplitud

Describe cuán variable o estable es la energía de una señal a lo largo del tiempo. Primero se obtienen los valores RMS para cada ventana temporal del espectrograma:

: magnitud de la amplitud en la frecuencia en la ventana temporal

: número total de bandas de frecuencia Mel

Luego se mide la relación entre la desviación estándar y el promedio de todos los RMS a lo largo del tiempo para obtener un coeficiente de variación adimensional:

Valores bajos indican que la energía de la señal es sostenida en el tiempo, mientras que valores altos indican mucha variabilidad, como en señales de pulsos o con intermitencias.

1. Flujo espectral

Mide cómo cambia el espectro de una señal a lo largo del tiempo, capturando la variabilidad entre ventanas temporales consecutivas del espectrograma, según la siguiente relación:

Luego se calcula la desviación estándar del flujo espectral a lo largo de todas las ventanas temporales proporcionando una medida de consistencia o irregularidad en los cambios espectrales y se escala logarítmicamente para reducir la sensibilidad a grandes variaciones:

Valores bajos indican cambios pequeños y/o consistentes en el espectro entre ventanas consecutivas, como pueden ser señales tonales con modulación lenta, mientras que valores altos indican cambios grandes y/o irregulares en la modulación espectro-temporal.

1. Entropía espectral

Mide la variabilidad espectral general de la señal (Arato & Fitch, 2021). Se utilizará la función *spectral\_entropy* de la librería *Antropy* (<https://github.com/raphaelvallat/antropy>). Ésta se basa en la entropía de Shannon (Shannon, 1948) aplicada sobre la densidad espectral de potencia tratada como una distribución de probabilidad, de modo que,

Donde,

: entropía espectral de la señal , con frecuencia de muestreo

: densidad espectral de potencia (*PSD*, por su sigla en inglés) calculada a partir del espectro *FFT* de una señal y normalizada para que la suma de los valores de potencia sobre todas las frecuencias sea igual a 1,0. Este proceso de normalización convierte la *PSD* en una «distribución de probabilidad» de la potencia en las diferentes frecuencias (Youngworth et al., 2005).

Valores bajos de indican que la energía está concentrada en pocas frecuencias específicas (baja complejidad), mientras que si la energía está dispersa entre muchas frecuencias (alta complejidad o incertidumbre), la entropía será alta (Shannon, 1948).

1. Coeficiente de tonalidad

El coeficiente de tonalidad o «planitud espectral» es una medida de cómo la energía está distribuida en un espectro de frecuencia, similar a la entropía espectral. Sin embargo, este coeficiente podría ser más sensible a las transiciones entre secciones tonales y no tonales (Dubnov, 2004). Se aplicará la función *librosa.feature.spectral\_flatness* para cada ventana temporal del espectrograma, luego se obtendrá la media escalada logarítmicamente. Valores bajos indican aleatoriedad (igual energía en todas las frecuencias), mientras que valores altos indican estructuras armónicas (energía concentrada en parciales)(Arato & Fitch, 2021). La formalización matemática puede ser revisada en detalle en (Dubnov, 2004).

1. Centroide espectral

Media ponderada de la distribución de energía espectral en una señal, que intuitivamente corresponde al centro de masa de dicha distribución y **se correlaciona con la frecuencia dominante (frecuencia con el** máximo valor de energía**) en señales acústicas de aves** (Arato & Fitch, 2021)**.** Para su cálculo, se evalúa el centroide espectral normalizado en cada ventana temporal del espectrograma, utilizando las magnitudes espectrales como pesos y sus frecuencias asociadas:

Luego, **se escala logarítmicamente y se obtiene la mediana. Este tratamiento permite** obtener un único valor representativo del **centroide espectral y reduce la influencia de valores atípicos y extremadamente grandes. Valores bajos indican concentración de energía en baja frecuencia, mientras que valores altos indican una concentración de la energía en alta frecuencia.**

Extracción de características de alto nivel

Las características de alto nivel se extraerán directamente de la penúltima capa de la red neuronal convolucional profunda *BirdNet* (Kahl et al., 2021) para cada señal (audio) del conjunto de datos. El resultado es un vector de la forma:

donde cada componente representa una característica numérica de punto flotante (números con decimales) extraída de una señal discreta con . Es decir, un vector de 1024 dimensiones que resume la representación interna de la red, capturando características complejas de la estructura acústica (McGinn et al., 2023) para cada una de las 363 señales del conjunto de datos.

Como ya se adelantó más arriba ([Figura 1](#Figura1)), se utilizará la técnica de aprendizaje no supervisado *UMAP* para reducir la alta dimensionalidad de estos vectores a solo tres. La garantía de UMAP es que preserva la estructura global y local de los datos (McInnes et al., 2020) y está disponible su implementación en Python mediante la librería *umap-learn* (<https://github.com/lmcinnes/umap>)

Señal filogenética

La señal filogenética se ha definido de diferentes maneras, en este trabajo la vamos a considerar como una medida de la dependencia estadística entre los valores de los rasgos de las especies debido a sus relaciones filogenéticas (Revell et al., 2008). Hay varias métricas populares para calcular esta dependencia de rasgos continuos y se mide comúnmente en caracteres unidimensionales.

**En primer lugar, se aplicará la *K de Blomberg* tradicional para los caracteres unidimensionales de bajo nivel.** Tal como indica la derivación original de este indicador (Blomberg et al., 2003), se ajustará una expresión que relaciona los datos observados con los datos después de corregir sus dependencias filogenéticas esperadas bajo un modelo de evolución de movimiento browniano . Para esta corrección se utiliza una matriz de varianza-covarianza derivada de la filogenia de (Stiller et al., 2024). Luego, esta relación se estandariza por su expectativa bajo el modelo browniano, quedando de la siguiente forma:

Donde,

: suma ordinaria de cuadrados de los caracteres de bajo nivel de las especies. Es decir, la variación entre los datos observados sin ninguna corrección filogenética:

con,

: vector que contiene el rasgo caracterizado para especies, centrado en la media simple

: vector transpuesto

: suma de cuadrados corregida filogenéticamente:

con,

: matriz que contiene la varianza-covarianza filogenética derivada del árbol dado. Describe cuán cercanas están las especies. Según el modelo browniano, los elementos diagonales de contienen las distancias filogenéticas desde cada especie hasta la raíz de la filogenia, mientras que los elementos fuera de la diagonal contienen las distancias desde la raíz hasta el ancestro común más reciente para cada par de especies (Mitteroecker et al., 2024))

: residuales de después de haber sido ajustado por la media filogenética.

y, el denominador que estandariza la relación observada por la relación esperada bajo un modelo de movimiento browniano (Adams, 2014):

De esta forma, cuando implica que las especies relacionadas son menos similares entre sí respecto al carácter en cuestión que lo esperado por el movimiento browniano a lo largo de la filogenia. Cuando , la similitud del carácter entre especies relacionadas es consistente con el movimiento browniano y cuando , las especies relacionadas son más similares en el carácter medido que lo predicho por el modelo de movimiento browniano (Mitteroecker et al., 2024). Si bien, se suele atribuir estos valores de señal filogenética a ciertos procesos o tasas evolutivas, no se recomienda este tipo de interpretaciones debido a la alta complejidad de estas asociaciones (Revell et al., 2008)

**Hasta acá, tenemos los elementos para medir la *K* *de Blomberg* en los cinco caracteres de bajo nivel** (). Este indicador de señal filogenética se implementará con la función *phylosig* de la librería *R* *Phytools.* Los argumentos de esta función son: un árbol en formato *.tree* y un *.csv* con los datos de caracteres. Además, *phylosig* puede ejecutar una prueba estadística del valor medido de *K* comparándolo con una distribución nula de *K*, permutando aleatoriamente los valores del caracter en los extremos de la filogenia.

**Sin embargo, para medir la *K* *de Blomberg* de un rasgo multidimensional (características de alto nivel)**, es un desafío porque diferentes dimensiones del rasgo pueden mostrar diferentes grados de señal filogenética (Mitteroecker et al., 2024). Acá se propone aplicar un enfoque reciente que maximiza la señal filogenética en rasgos multidimensionales utilizando una extensión del estadístico *K* *de Blomberg* (Mitteroecker et al., 2024). El enfoque propone una descomposición multidimensional en *«K-componentes»,* que son combinaciones lineales de las dimensiones que maximizan (o minimizan) la señal filogenética, a la vez que mejora las limitaciones de una propuesta previa (Adams, 2014), que no distinguía si la señal estaba concentrada en pocas dimensiones o distribuida uniformemente. El método, llamado *Análisis de K-componentes* (*KCA*, por su sigla en inglés), no está disponible aún en librerías de programación. **Por lo tanto,** **se implementará un *script* propio siguiendo la formalización algebraica descrita por** Mitteroecker et al., (2024)**.**

Con estas medidas de señal filogenética, se espera alcanzar el objetivo de identificar una caracterización acústica que refleje de mejor manera la dependencia filogenética en el conjunto de datos.

**SECCIÓN 2: Reconstrucción a gran escala**

**Objetivo específico 2.1: Reconstruir los estados ancestrales en la filogenia más actual, bien soportada y de gran escala de las aves modernas, a partir de la caracterización acústica determinada en la sección anterior.**

Existen tres criterios principales para llevar a cabo una reconstrucción ancestral: máxima parsimonia, máxima verosimilitud y estimación bayesiana. Los estudios pioneros en poner en práctica la reconstrucción ancestral de rasgos acústicos usaron métodos de parsimonia. Sin embargo, en la actualidad se prefieren criterios bayesianos y de máxima verosimilitud por ser más sofisticados y tener supuestos más realistas basados en modelos evolutivos (Odom et al., 2021). En un espíritu dialogante con las preferencias actuales, se adoptan estos criterios en la metodología descrita a continuación.

Modelo evolutivo: movimiento browniano

Una vez construido nuestro conjunto de **caracteres continuos** de bajo y alto nivel y la filogenia asociada, se propone el modelo evolutivo estándar para reconstruir los valores ancestrales: el modelo de movimiento browniano. El movimiento browniano es un modelo de «paseo aleatorio» (Codling et al., 2008) en tiempo continuo y sin dirección. Bajo el movimiento browniano, los cambios evolutivos sucesivos son independientes y provienen de una distribución gaussiana con media igual a cero y varianza igual a , en la que es la tasa instantánea del proceso gaussiano que aumenta linealmente respecto al tiempo (Harmon, 2019b; Revell, 2024a).

Procedimiento de reconstrucción de caracteres continuos de bajo nivel

El procedimiento habitual es aplicar el modelo browniano simulando cambios a lo largo de las ramas del árbol, utilizando la distribución gaussiana mencionada, generando un conjunto de valores que reflejan cómo se espera que evolucione el caracter a lo largo del tiempo. **La estimación de los valores ancestrales se realiza comparando los valores observados en las especies actuales con los cambios simulados.**

**Bajo el criterio de máxima verosimilitud**, aplicaremos un algoritmo rápido para encontrar los valores en los nodos internos que maximicen la probabilidad de observar los valores en los nodos terminales (especies actuales) según el movimiento browniano (Revell, 2024a). Haremos uso de la función *fastAnc* de *phytools* para encontrar las estimaciones de máxima verosimilitud en los nodos internos junto con las varianzas e intervalos de confianza del 95% para cada nodo (Revell & Harmon, 2022).

**También aplicaremos un enfoque bayesiano bajo movimiento browniano** para estimar los estados ancestrales mediante la función *anc.Bayes* de *phytools.* Esta función utiliza un método MCMC (Cadena de Markov Monte Carlo) que permite inferir los parámetros del modelo y muestrear la distribución posterior de los estados ancestrales (Gilks et al., 1996). Como todo enfoque bayesiano, admite información previa sobre los estados en los nodos internos del árbol y entrega intervalos de confianza, así como la incertidumbre asociada a las estimaciones de los estados ancestrales (Revell, 2024b).

En definitiva, mientras que el enfoque de máxima verosimilitud se centra en la estimación puntual de los estados ancestrales basándose en la probabilidad de los datos, el enfoque bayesiano proporciona una evaluación más completa de la incertidumbre en las estimaciones. Ambos enfoques serán implementados en el análisis de la reconstrucción de caracteres continuos de bajo nivel suponiendo un modelo evolutivo de movimiento browniano (Revell & Harmon, 2022).

Por cierto, la implementación en *phytools* de los métodos mencionados admite la evaluación de diferentes modelos según criterios como el *AIC* (Criterio de Información de Akaike) o el *BIC* (Criterio de Información Bayesiano). Por lo tanto, **se evaluará qué modelo de movimiento browniano y sus extensiones se ajusta mejor a los datos**. En particular dos extensiones importantes de probar son el modelo de «*Estallido Temprano*» *(EB),* en el que la tasa de evolución varía como una función continua del tiempo desde la raíz del árbol y el modelo de *Ornstein-Uhlenbeck (OU),* en el que la evolución sigue un camino aleatorio, pero con la tendencia a revertir hacia un valor central concreto. Tales modelos se han asociado con el concepto de radiación adaptativa y con la selección natural direccional, respectivamente (Revell & Harmon, 2022). Esta asociación resulta particularmente interesante debido a la hipótesis bien aceptada de radiación adaptativa rápida o «estallido temprano» en el origen evolutivo de las aves de la corona (Stiller et al., 2024)

Procedimiento de reconstrucción de caracteres continuos de alto nivel

Una reconstrucción ancestral basada en caracteres continuos de dos o más dimensiones requiere un **modelo evolutivo multivariado**. Si consideramos el modelo de movimiento browniano en más de una dimensión al mismo tiempo, ya no se puede describir con un parámetro de una sola tasa, . Es necesario incluir una tasa para cada dimensión del caracter y una covarianza evolutiva entre ellas. Esto permite que la correlación entre las dimensiones sea diferente en diferentes partes de la filogenia (Revell & Harmon, 2022. Pag:122-124). Este modelo tiene sentido si queremos probar la correlación entre los ejes de similitud en un espacio vectorial como el proyectado con UMAP. De las librerías de programación que han implementado esta modalidad multivariada, uno bien actualizado lo encontramos en *phytools* con funciones que, incluso admiten un modelo diferente para cada dimensión (Revell, 2024b). Otra poderosa herramienta, la encontramos en la librería *ratematrix,* la que implementa en lenguaje *R* una estimación bayesiana de matrices de tasas y covarianzas evolutivas bajo el movimiento browniano utilizando MCMC. Con esta implementación es posible analizar simultáneamente la tasa y estructura de integración evolutiva entre dos o más dimensiones a lo largo de las ramas de un árbol filogenético (Caetano & Harmon, 2017). Se aplicarán ambos métodos en las estimaciones de los valores ancestrales del carácter multidimensional de alto nivel.

**Objetivo específico 2.2: Sintetizar versiones espectrográficas y auditivas de los componentes ancestrales reconstruidos.**

Un aspecto original que se plantea realizar, es la **síntesis de las estructuras acústicas reconstruidas** con los métodos anteriormente mencionados. El procesamiento de una reconstrucción de estados ancestrales arroja como resultado un contenedor de valores continuos que, en el caso de los caracteres de alto nivel, difícilmente pueden ser interpretados. Por lo tanto, es necesaria una síntesis en algún formato que permita una apreciación y comprensión biológica y acústica.

Morfometría geométrica y transformada inversa de Fourier

Se propone para estos efectos de síntesis, el método *eigensound* implementado en *R* con el nombre de *SoundShape* (Rocha & Romano, 2021). El método se basa en un análisis de morfometría geométrica y permite la síntesis espectrográfica de formas hipotéticas entre datos acústicos homólogos. En nuestro caso, los datos acústicos de alto nivel que nos interesa sintetizar estarán reducidos y encriptados en el espacio de características *UMAP* de tres dimensiones. En alguna parte de este espacio estarán los puntos reconstruidos de los nodos internos de la filogenia y para recuperar la información espectrográfica y auditiva se propone el siguiente procedimiento exploratorio:

**Si la predicción es cierta.** Es decir, las estructuras reconstruidas no son radicalmente diferentes a las señales de las aves actuales, los valores tridimensionales reconstruidos deberían estar dentro o relativamente cerca de algún agrupamiento de datos conocidos (puntos que representan las características de alto nivel de señales de 3 segundos). Si esto es así, mediante análisis morfométrico del «vecindario» más cercano (método *eigensound*) es posible recuperar las estructuras compartidas en el espacio tiempo-frecuencia original de los datos y estimar las distancias euclidianas entre ellas para sintetizar un espectrograma hipotético del punto reconstruido (Rocha & Romano, 2021).

**Si la predicción no es cierta,** los valores reconstruidos en el espacio *UMAP* estarán localizados muy lejos de algún dato conocido. Esto significa que la morfología acústica ancestral sería aparentemente irrecuperable.

Finalmente, si encontramos apoyo a la hipótesis planteada, se llevará a cabo un procesamiento para transformar estas reconstrucciones espectrográficas hipotéticas en una señal de audio, aplicando transformada inversa de Fourier (ISTFT). Además, se debe sintetizar la fase que, por definición, no existe en un espectrograma y es indispensable para reconstruir la señal temporal. La técnica tradicional para realizar esta recuperación de fase es el algoritmo de Griffin-Lim (Perraudin et al., 2013), implementado en la librería *Python Librosa* (McFee et al., 2015)

**Objetivo específico 2.3: Determinar los componentes ritualizados o sus rastros revelados en la reconstrucción y síntesis ancestral, así como la presencia en los repertorios de las aves actuales.**

Una vez reconstruidas las estructuras ancestrales hipotéticas en los nodos internos, mediante las siguientes métricas es posible determinar qué atributos acústicos son más o menos estables a lo largo de la evolución filogenética.

1. Autocorrelación: mide patrones globales de redundancia y periodicidad
2. Calcular la autocorrelación espectro-temporal de cada señal reconstruida por nodo para identificar patrones de periodicidad o redundancia intrínsecos.
3. Calcular la correlación cruzada de las autocorrelaciones entre nodos adyacentes y en diferentes niveles filogenéticos.
4. Calcular correlación de Pearson entre vectores de autocorrelación.
5. Caracteres de bajo nivel mencionados en la [sección anterior](#CaracteresBajoNivel)
6. Calcular para cada nodo interno
7. Calcular correlación de Pearson a lo largo de cada línea evolutiva (entre nodo raíz y cada nodo terminal)
8. Calcular las diferencias relativas entre nodos consecutivos en cada línea evolutiva para identificar qué tan rápido cambia cada carácter.
9. Representar los valores de cada caracter como un gradiente sobre la filogenia para observar patrones globales
10. Simple inspección
11. Observar e identificar cambios espectrográficos en las líneas evolutivas
12. Escuchar las reconstrucciones e identificar cambios perceptuales en las líneas evolutivas

**SECCIÓN 3: Restricciones evolutivas**

**Objetivo específico 3.1: Reconstruir los tipos de siringe en la filogenia de las aves, codificando el número de fuentes de sonido de las familias taxonómicas muestreadas.**

Como se mencionó en la [introducción](#Introduccion), se reconocen cuatro tipos de siringe según cantidad de labios y membranas oscilatorias (fuente de sonido). El tipo más común es la **siringe traqueo-bronquial** con dos fuentes de sonido. Luego encontramos la **siringe traqueal** con una sola fuente de sonido y la **siringe bronquial** con dos. Un cuarto tipo de siringe con un par de membranas traqueales que interactúa con dos pares de labios bronquiales conforman una estructura de tres fuentes de sonido que llamaremos **siringe bronquio-traqueal** (Amador & Mindlin, 2023; Chiappone et al., 2024; Garcia et al., 2017; Goller, 2022; Mindlin & Laje, 2005).

Con este caracter de cuatro estados (**1:** **siringe traqueal, 2: siringe bronquial, 3: siringe traqueo-bronquial y 4: siringe bronquio-traqueal**) se propone una reconstrucción ancestral a nivel de familia, mapeando los estados en la filogenia de las aves modernas de Stiller et al., (2024). Pero antes se debe hacer una **revisión bibliográfica exhaustiva para recuperar esta información morfológica en las 218 familias** representadas en el conjunto de datos y «podar» de la filogenia las familias sin información.

Una reconstrucción de **caracteres discretos**, es realizada habitualmente utilizando modelos basados en cadenas de Markov. Es decir, modelos que representan un sistema que transita entre una serie de estados, donde la probabilidad de ir a un próximo estado depende únicamente del estado actual y no del camino por el que se llegó a dicho estado (Harmon, 2019a). Un modelo bien popular es el modelo *Mk*, el cual describe una cadena de Markov en tiempo continuo con k posibles estados. Con un enfoque de máxima verosimilitud se ajustan los parámetros del modelo, es decir, las tasas instantáneas de transición entre los estados, de manera que se explique de la mejor manera posible la distribución de estados observada en los datos, dado el árbol. Una vez ajustado el modelo, la estimación de los estados ancestrales en los nodos de la filogenia se puede lograr utilizando diferentes procedimientos como la reconstrucción marginal, reconstrucción conjunta o el mapeo estocástico de caracteres(Revell, 2024a). Este último procedimiento, genera mapas aleatorios de los estados del caracter, incluyendo los cambios a lo largo de las ramas del árbol, a partir de la distribución de probabilidad de los datos. La acumulación de estos mapas permite calcular las probabilidades posteriores de que cada nodo esté en cada estado del carácter y estimar la distribución de probabilidad del número de cambios de cada estado en el árbol. El método supone que el tiempo evolutivo está relacionado con la cantidad de cambios de estado, lo que proporciona una representación más realista de cómo los caracteres pueden evolucionar a lo largo del tiempo. Otras ventajas incluyen la posibilidad de comparar modelos y tener en cuenta la incertidumbre en las tasas de transición.

En resumen, el **mapeo estocástico de caracteres** es una herramienta poderosa para estudiar la evolución de caracteres discretos al ofrecer una estimación flexible y robusta de los estados ancestrales (Huelsenbeck et al., 2003) y **es el método que se aplicará para estimar los estados de la siringe ancestral.**

Utilizando la librería *phytools* es posible realizar este método de varias maneras, acá se procederá con las siguientes acciones preliminares:

1. Ajustar, mediante máxima verosimilitud, con la función *fitMk*, los parámetros de tres modelos *Mk* de 4 estados (detalles de estos modelos en el documento anexo (material\_suplementario.docx)).
2. Generar, con la función *simmap*, 1000 mapas estocásticos («*entre 100 y 1000 pueden ser suficientes*» (Revell, 2024b)).
3. Calcular, utilizando el método de resumen genérico *summary,* las probabilidades posteriores de cada estado de carácter en cada nodo interno del árbol.
4. Calcular la distribución de frecuencias relativas de los tipos de transiciones entre estados en el conjunto total de mapas y a lo largo de todo el árbol, así como los intervalos bayesianos de densidad de alta probabilidad (HPD) del 95%. Esto se hace con la función *density.*
5. Graficar las probabilidades posteriores de cada uno de los estados a lo largo de las ramas de la filogenia, para visualizar la incertidumbre en la reconstrucción del estado ancestral en un carácter multiestado.

Se ejercita este procedimiento con una simulación simple en el documento anexo (material\_suplementario.docx).

**Objetivo específico 3.2: Evaluar influencia de las restricciones morfológicas de la siringe en la evolución acústica de las aves mediante un análisis simultáneo de tasas y correlación de dimensiones sobre una filogenia previamente mapeada con los tipos de siringe reconstruidos.**

La relación entre dimensiones, tasas y restricciones representa un grado de complejidad necesario de abordar en el problema acústico-evolutivo que se presenta aquí. Algunas funciones implementadas en *phytools* permiten que la tasa de evolución de cada dimensión de un carácter continuo y la correlación entre ellas pueda variar libremente entre diferentes restricciones discretas mapeadas previamente en la filogenia. Esto relaja la suposición de una coevolución constante en todas las ramas y clados (Revell et al., 2022; Revell & Collar, 2009), y nos permite evaluar la influencia de la morfología siríngea en la evolución de la señal acústica.

El método que se propone es el implementado en la función *evolvcv.lite* de *phytools,* el cual admite modelos multivariados de evolución browniana diferentes para cada dimensión del caracter (Revell et al., 2022). Se repetirá este mismo análisis con una función de la librería *R* *ratematrix*, que implementa una estimación bayesiana de matrices de tasas y covarianzas evolutivas bajo el movimiento browniano utilizando *MCMC* (Caetano & Harmon, 2017). Ambos análisis serán ejecutados a partir de los caracteres de alto nivel *(BirdNet + UMAP)* sobre la filogenia mapeada con los estados del carácter siríngeo, obtenida anteriormente ([Objetivo específico 3.1](#ObjEsp31))

En resumen, el análisis estará dirigido a medir la consistencia (o inconsistencia) entre la dirección y tasas del cambio evolutivo, respecto a las restricciones de la siringe mapeadas previamente en la filogenia. Siguiendo un análisis muy similar (aunque sobre rasgos y restricciones diferentes) realizado en Revell et al., (2022), proponemos probar la hipótesis de correlación entre señal y siringe ([hipótesis 3.1](#Hip31)), mediante una evaluación de modelos según el criterio de Akaike (*AIC*).

Modelos a evaluar:

* Modelo 1: Tasas y correlaciones comunes.

Asume que los caracteres continuos (dimensiones en el espacio *UMAP*) evolucionan con la misma tasa, sin importar el estado del carácter discreto (tipo de siringe) y las correlaciones entre los caracteres continuos también son constantes en todos los estados del carácter discreto. Si este modelo tiene respaldo, se puede interpretar que: las dimensiones del carácter acústico evolucionan al mismo ritmo en todas las especies, independientemente del tipo de siringe.

* Modelo 2: Tasas diferentes según los estados discretos.

Permite que la tasa de cambio de los caracteres continuos varíe según el estado del carácter discreto.

Sin embargo, las correlaciones entre los caracteres continuos son iguales en todos los estados. Si este modelo tiene respaldo, se puede interpretar que: las dimensiones del carácter acústico evolucionan más rápido en aves con un tipo de siringe que en aves con otro tipo, pero las correlaciones entre las dimensiones permanecen constantes.

* Modelo 3: Correlaciones diferentes según los estados discretos.

Asume que las tasas de evolución de los caracteres continuos son iguales en todos los estados discretos, mientras que las correlaciones entre los caracteres continuos pueden variar según el estado del carácter discreto. Si este modelo tiene respaldo, se puede interpretar que: las dimensiones del carácter acústico evolucionan al mismo ritmo, pero su relación cambia dependiendo del tipo de siringe (por ejemplo, en un estado, las dimensiones pueden estar más correlacionadas que en otro).

* Modelo 4: Tasas y correlaciones diferentes según los estados discretos.

Permite que tanto las tasas de evolución como las correlaciones entre los caracteres continuos varíen según el estado del carácter discreto. Si este modelo tiene respaldo, se puede interpretar que: en clados con un tipo de siringe, las dimensiones del carácter acústico evolucionan más rápido y están más correlacionadas, mientras que en otro tipo evolucionan más lentamente y con menor correlación.

Si la [predicción 3.1](#Pred31) es cierta, los modelos 2, 3 y 4 deberían obtener mejores puntajes y, dependiendo de cuál de estos tres modelos es el que mejor explica nuestros datos de caracteres de alto nivel (es decir, el con menor valor de *AIC*), podremos interpretar la influencia del tipo de siringe sobre la velocidad y/o correlación entre las dimensiones que evolucionan.

Se ejercita este procedimiento con una simulación simple en el documento anexo (material\_suplementario.docx).

**Objetivo específico 3.3: Evaluar influencia del aprendizaje vocal mediante un análisis de tasas y correlación de caracteres de bajo nivel sobre filogenias de grupos aprendices y no aprendices.**

En primer lugar, vamos a considerar dos conjuntos de datos nuevos en cuanto a filogenias y los datos acústicos asociados a las especies de ambas filogenias:

1. Tinamúes

Los tinamúes son una familia de aves neotropicales del clado *Palaeognathae,* no son aprendices vocales y representan un grupo de interés para comprender la evolución acústica de las aves modernas en relación a dos aspectos: (1) son bien conocidos por sus señales de alta pureza tonal (Laverde, 2016) con características aparentemente bien conservadas y (2) hay un fuerte resapaldo a una subdivisión entre tinamúes que habitan áreas abiertas (Nothurinae) y los de ambientes boscosos (Tinaminae) (Almeida et al., 2022). Este último aspecto nos proporciona una variable ecológica a considerar en las discusiones de nuestros resultados

La filogenia más actual del grupo que utilizaremos para un análisis intra-familiar a nivel de especie, es la informada por Almeida et al. (2022). Seleccionaremos el árbol bayesiano de consenso que fue derivado de tres loci mitocondriales y cinco loci nucleares (7316 caracteres alineados, 31 especies de la familia + 7 del grupo externo). El árbol ultramétrico, incluye calibración fosil en dos nodos y está disponible en el material suplementario (Almeida et al., 2022).

Los datos acústicos serán obtenidos de la colección Xeno-canto ([https://xeno-canto.org](https://xeno-canto.org/)), mediante el mismo procedimiento descrito en la sección [Conjunto de datos](#ConjDatos) del objetivo específico 1.1.

1. Ictéridos

Los ictéridos (tordos del Nuevo Mundo) se encuentran entre las familias de aves oscinas más estudiadas. Los aspectos que se han investigado aplicando métodos comparativos filogenéticos incluyen los sistemas de apareamiento, parasitismo de cría, dimorfismo sexual, patrón y color de plumaje, repertorio acústico femenino, divergencia acústica, migración, historia biogeográfica y divergencia de nichos ecológicos (referencias a todos estos trabajos en Powell et al., 2014). En nuestro caso, utilizaremos a este grupo de aves como un sistema modelo para investigar las implicancias del aprendizaje vocal en la evolución de sus señales acústica.

Utilizaremos como filogenia independiente para los análisis, el árbol molecular más completo informado a la fecha, correspondiente a 108 especies, secuenciadas con caracteres del genoma mitocondrial (completo o parcial, dependiendo de la especie, pudiendo alcanzar hasta 16775 pares de bases) y cuatro loci nucleares que abarcan un total de 5266 pares de bases. Seleccionamos el mejor árbol obtenido por máxima verosimilitud coincidente con análisis bayesiano según el mismo conjunto de datos concatenados (Figura 4 de Powell et al., (2014)). El árbol ultramétrico está disponible en el material suplementario del mismo artículo (Powell et al., 2014).

Los datos acústicos serán descargados de Xeno-canto con el mismo protocolo mencionado.

El siguiente procedimiento consiste en extraer características acústicas de bajo nivel () y reconstruir los estados ancestrales de los nodos internos según el procedimiento descrito para el [objetivo específico 2.1](#ObjEsp21) en ambos conjuntos de datos (tinamúes e ictéridos).

Luego se deben calcular las tasas de cambio (varianza) utilizando los valores reconstruidos en los nodos internos para cada carácter. El siguiente paso es comparar las tasas entre filogenias. Esto se puede hacer calculando la razón de tasas para identificar si un caracter evoluciona más rápido en aprendices o no aprendices.

Posteriormente, evaluamos la correlación de cambios evolutivos en ambos conjuntos de datos, de la siguiente manera: (1) calcular los cambios por rama para cada carácter (diferencias absolutas entre los valores reconstruidos en los nodos adyacentes); (2) calcular las correlaciones entre los vectores de cambios para cada par de caracteres; y (3) visualizar la matriz de correlaciones mediante un mapa de calor (*heatmap*) para identificar correlaciones fuertes entre caracteres.

Se ejercita este procedimiento con una simulación simple en el documento anexo (material\_suplementario.docx).

Otro método que se propone explorar para evaluar las tasas evolutivas es la función *ratebytrees*. El flujo de trabajo de esta función es ajustar un modelo bajo dos hipótesis: (1) tasa evolutiva común para todas las filogenias y (2) tasa evolutiva diferente para cada filogenia. Luego, aplica un test de razón de verosimilitudes para determinar si el modelo con diferentes tasas se ajusta significativamente mejor que el modelo con una tasa común.

Finalmente, se explorará un método llamado *Mínimos Cuadrados Generalizados Filogenéticos* (*PGLS*, por su sigla en inglés) que se puede utilizar para probar hipótesis sobre cómo los rasgos se relacionan entre sí (Revell & Harmon, 2022, cap:3).

1. **Referencias**

Adams, D. C. (2014). A Generalized K Statistic for Estimating Phylogenetic Signal from Shape and Other High-Dimensional Multivariate Data. *Systematic Biology*, *63*(5), 685–697. https://doi.org/10.1093/sysbio/syu030

Allen-Ankins, S., & Schwarzkopf, L. (2022). Using citizen science to test for acoustic niche partitioning in frogs. *Scientific Reports*, *12*(1), 2447. https://doi.org/10.1038/s41598-022-06396-0

Almeida, F. C., Porzecanski, A. L., Cracraft, J. L., & Bertelli, S. (2022). The evolution of tinamous (Palaeognathae: Tinamidae) in light of molecular and combined analyses. *Zoological Journal of the Linnean Society*, *195*(1), 106–124. https://doi.org/10.1093/zoolinnean/zlab080

Amador, A., & Mindlin, G. B. (2023). The dynamics behind diversity in suboscine songs. *Journal of Experimental Biology*, *226*(12). https://doi.org/10.1242/jeb.227975

Ames, P. L. (1971). The Morphology of the Syrinx in Passerine Birds . In *Encyclopaedia Britannica* (Vol. 37, pp. 133–152). Peabody Museum of Natural History Yale University.

Amorim, M. C. P., Pedroso, S. S., Bolgan, M., Jordão, J. M., Caiano, M., & Fonseca, P. J. (2013). Painted gobies sing their quality out loud: acoustic rather than visual signals advertise male quality and contribute to mating success. *Functional Ecology*, *27*(2), 289–298. https://doi.org/10.1111/1365-2435.12032

Arato, J., & Fitch, W. T. (2021). Phylogenetic signal in the vocalizations of vocal learning and vocal non-learning birds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1836). https://doi.org/10.1098/rstb.2020.0241

Aubin, T., Mathevon, N., Silva, M. L. da, Vielliard, J. M. E., & Sebe, F. (2004). How a simple and stereotyped acoustic signal transmits individual information: the song of the White-browed Warbler Basileuterus leucoblepharus. *Anais Da Academia Brasileira de Ciências*, *76*(2), 335–344. https://doi.org/10.1590/S0001-37652004000200022

Beckers, G. J. L., Suthers, R. A., & Cate, C. ten. (2003). Pure-tone birdsong by resonance filtering of harmonic overtones. *Proceedings of the National Academy of Sciences*, *100*(12), 7372–7376. https://doi.org/10.1073/pnas.1232227100

Benton, M. J. (2021). The origin of endothermy in synapsids and archosaurs and arms races in the Triassic. *Gondwana Research*, *100*, 261–289. https://doi.org/10.1016/j.gr.2020.08.003

Blomberg, S. P., Garland, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, *57*(4), 717–745. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x

Boncoraglio, G., & Saino, N. (2007). Habitat structure and the evolution of bird song: a meta‐analysis of the evidence for the acoustic adaptation hypothesis. *Functional Ecology*, *21*(1), 134–142. https://doi.org/10.1111/j.1365-2435.2006.01207.x

Caetano, D. S., & Harmon, L. J. (2017). ratematrix: An R package for studying evolutionary integration among several traits on phylogenetic trees. *Methods in Ecology and Evolution*, *8*(12), 1920–1927. https://doi.org/10.1111/2041-210X.12826

Cardoso, G. C. (2017). Advancing the inference of performance in birdsong. *Animal Behaviour*, *125*, e29–e32. https://doi.org/10.1016/j.anbehav.2016.11.034

Chek, A. A., Bogart, J. P., & Lougheed, S. C. (2003). Mating signal partitioning in multi‐species assemblages: a null model test using frogs. *Ecology Letters*, *6*(3), 235–247. https://doi.org/10.1046/j.1461-0248.2003.00420.x

Chen, Z., & Wiens, J. J. (2020). The origins of acoustic communication in vertebrates. *Nature Communications*, *11*(1), 369. https://doi.org/10.1038/s41467-020-14356-3

Chiappone, M., Rodriguez‐Saltos, C., Legendre, L. J., Li, Z., & Clarke, J. (2024). Ostrich ( *Struthio camelus* ) syrinx morphology and vocal repertoire across postnatal ontogeny and sex: Implications for understanding vocal evolution in birds. *Journal of Anatomy*, *244*(4), 541–556. https://doi.org/10.1111/joa.13992

Cicero, C., Mason, N. A., Benedict, L., & Rising, J. D. (2020). Behavioral, morphological, and ecological trait evolution in two clades of New World Sparrows ( *Aimophila* and *Peucaea* , Passerellidae). *PeerJ*, *8*, e9249. https://doi.org/10.7717/peerj.9249

Clarke, J. A., Chatterjee, S., Li, Z., Riede, T., Agnolin, F., Goller, F., Isasi, M. P., Martinioni, D. R., Mussel, F. J., & Novas, F. E. (2016). Fossil evidence of the avian vocal organ from the Mesozoic. *Nature*, *538*(7626), 502–505. https://doi.org/10.1038/nature19852

Codling, E. A., Plank, M. J., & Benhamou, S. (2008). Random walk models in biology. *Journal of The Royal Society Interface*, *5*(25), 813–834. https://doi.org/10.1098/rsif.2008.0014

Collias, N. E. (1963). A Spectrographic Analysis of the Vocal Repertoire of the African Village Weaverbird. *The Condor*, *65*(6), 517–527. https://doi.org/10.2307/1365510

Cusano, D. A., Paton, D., Noad, M. J., & Dunlop, R. A. (2021). Socially Complex Breeding Interactions in Humpback Whales Are Mediated Using a Complex Acoustic Repertoire. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.665186

de Kort, S. R., & ten Cate, C. (2004). Repeated decrease in vocal repertoire size in Streptopelia doves. *Animal Behaviour*, *67*(3), 549–557. https://doi.org/10.1016/j.anbehav.2003.07.004

Dubnov, S. (2004). Generalization of Spectral Flatness Measure for Non-Gaussian Linear Processes. *IEEE Signal Processing Letters*, *11*(8), 698–701. https://doi.org/10.1109/LSP.2004.831663

Dudouit, C., Maury, C., Bosca, J., Bakker, A., Gahr, M., Aubin, T., Rybak, F., & Geberzahn, N. (2022). Vocal performance during spontaneous song is equal in male and female European robins. *Animal Behaviour*, *193*, 193–203. https://doi.org/10.1016/j.anbehav.2022.08.012

Dunlop, R., Gannon, W. L., Kiley-Worthington, M., Hill, P. S. M., Wessel, A., & Thomas, J. A. (2022). Vibrational and acoustic communication in animals. In C. Erbe & J. A. Thomas (Eds.), *Exploring Animal Behavior Through Sound* (Vol. 1, pp. 389–417). Springer Nature.

Düring, D. N., & Elemans, C. P. H. (2016). Embodied Motor Control of Avian Vocal Production. In R. A. Suthers, W. T. Fitch, R. R. Fay, & A. N. Popper (Eds.), *Vertebrate Sound Production and Acoustic Communication* (pp. 119–158). Springer.

Enquist, M., Arak, A., Ghirlanda, S., & Wachtmeister, C. (2002). Spectacular phenomena and limits to rationality in genetic and cultural evolution. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1427), 1585–1594. https://doi.org/10.1098/rstb.2002.1067

Ey, E., & Fischer, J. (2009). The “Acoustic Adaptation Hypothesis”—A review of the evidence from birds, anurans and mammals. *Bioacoustics*, *19*(1–2), 21–48. https://doi.org/10.1080/09524622.2009.9753613

Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist*, *125*(1), 1–15. https://doi.org/10.1086/284325

Fletcher, N. H., & Tarnopolsky, A. (1999). Acoustics of the avian vocal tract. *The Journal of the Acoustical Society of America*, *105*(1), 35–49. https://doi.org/10.1121/1.424620

Fraixedas, S., Lindén, A., Piha, M., Cabeza, M., Gregory, R., & Lehikoinen, A. (2020). A state-of-the-art review on birds as indicators of biodiversity: Advances, challenges, and future directions. *Ecological Indicators*, *118*, 106728. https://doi.org/10.1016/j.ecolind.2020.106728

Freitas, B., D’Amelio, P. B., Milá, B., Thébaud, C., & Janicke, T. (2024). Meta-analysis of the acoustic adaptation hypothesis reveals no support for the effect of vegetation structure on acoustic signalling across terrestrial vertebrates. *BioRxiv*. https://doi.org/10.1101/2024.02.21.581368

Garcia, S. M., Kopuchian, C., Mindlin, G. B., Fuxjager, M. J., Tubaro, P. L., & Goller, F. (2017). Evolution of Vocal Diversity through Morphological Adaptation without Vocal Learning or Complex Neural Control. *Current Biology*, *27*(17), 2677-2683.e3. https://doi.org/10.1016/j.cub.2017.07.059

Geberzahn, N., & Aubin, T. (2014). Assessing vocal performance in complex birdsong: a novel approach. *BMC Biology*, *12*(1), 58. https://doi.org/10.1186/s12915-014-0058-4

Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour*, *66*(1), 15–36. https://doi.org/10.1006/anbe.2003.2174

Ghirlanda, S., Jansson, L., & Enquist, M. (2002). Chickens prefer beautiful humans. *Human Nature*, *13*(3), 383–389. https://doi.org/10.1007/s12110-002-1021-6

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1996). Introducing Markov chain Monte Carlo. In W. R. Gilks, S. Richardson, & D. J. Spiegelhalter (Eds.), *Markov Chain Monte Carlo in Practice* (pp. 1–20). Chapman & Hall/CRC.

Gingras, B., & Fitch, W. T. (2013). A three-parameter model for classifying anurans into four genera based on advertisement calls. *The Journal of the Acoustical Society of America*, *133*(1), 547–559. https://doi.org/10.1121/1.4768878

Gingras, B., Mohandesan, E., Boko, D., & Fitch, W. T. (2013). Phylogenetic signal in the acoustic parameters of the advertisement calls of four clades of anurans. *BMC Evolutionary Biology*, *13*(1), 134. https://doi.org/10.1186/1471-2148-13-134

Goller, F. (2022). The syrinx. *Current Biology*, *32*(20), R1095–R1100. https://doi.org/10.1016/j.cub.2022.08.034

Goller, F., Love, J., & Mindlin, G. (2021). Different frequency control mechanisms and the exploitation of frequency space in passerines. *Ecology and Evolution*, *11*(11), 6569–6578. https://doi.org/10.1002/ece3.7510

Griffiths, C. S. (1994). Monophyly of the Falconiformes Based on Syringeal Morphology. *The Auk*, *111*(4), 787–805. https://doi.org/10.2307/4088811

Gupta, S., Alluri, R. K., Rose, G. J., & Bee, M. A. (2021). Neural basis of acoustic species recognition in a cryptic species complex. *Journal of Experimental Biology*, *224*(23). https://doi.org/10.1242/jeb.243405

Hardt, B., & Benedict, L. (2021). Can you hear me now? A review of signal transmission and experimental evidence for the acoustic adaptation hypothesis. *Bioacoustics*, *30*(6), 716–742. https://doi.org/10.1080/09524622.2020.1858448

Harmon, L. J. (2019a). Beyond the Mk Model. In *Phylogenetic Comparative Methods* (pp. 132–146). Independent.

Harmon, L. J. (2019b). *Phylogenetic Comparative Methods*. Independent.

Hart, P. J., Ibanez, T., Paxton, K., Tredinnick, G., Sebastián-González, E., & Tanimoto-Johnson, A. (2021). Timing Is Everything: Acoustic Niche Partitioning in Two Tropical Wet Forest Bird Communities. *Frontiers in Ecology and Evolution*, *9*. https://doi.org/10.3389/fevo.2021.753363

Hoese, W. J., Podos, J., Boetticher, N. C., & Nowicki, S. (2000). Vocal Tract Function in Birdsong Production: Experimental Manipulation of Beak Movements. *Journal of Experimental Biology*, *203*(12), 1845–1855. https://doi.org/10.1242/jeb.203.12.1845

Huelsenbeck, J. P., Nielsen, R., & Bollback, J. P. (2003). Stochastic Mapping of Morphological Characters. *Systematic Biology*, *52*(2), 131–158. https://doi.org/10.1080/10635150390192780

Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, *1016*(1), 749–777. https://doi.org/10.1196/annals.1298.038

Jorgewich-Cohen, G., Townsend, S. W., Padovese, L. R., Klein, N., Praschag, P., Ferrara, C. R., Ettmar, S., Menezes, S., Varani, A. P., Serano, J., & Sánchez-Villagra, M. R. (2022). Common evolutionary origin of acoustic communication in choanate vertebrates. *Nature Communications*, *13*(1), 6089. https://doi.org/10.1038/s41467-022-33741-8

Ju, C., Geller, F. C., Mundinger, P. C., & Lahti, D. C. (2019). Four decades of cultural evolution in House Finch songs. *The Auk*, *136*(1). https://doi.org/10.1093/auk/uky012

Kahl, S., Wood, C. M., Eibl, M., & Klinck, H. (2021). BirdNET: A deep learning solution for avian diversity monitoring. *Ecological Informatics*, *61*, 101236. https://doi.org/10.1016/j.ecoinf.2021.101236

King, B., & Rücklin, M. (2020). A Bayesian approach to dynamic homology of morphological characters and the ancestral phenotype of jawed vertebrates. *ELife*, *9*. https://doi.org/10.7554/eLife.62374

Kingsley, E. P., Eliason, C. M., Riede, T., Li, Z., Hiscock, T. W., Farnsworth, M., Thomson, S. L., Goller, F., Tabin, C. J., & Clarke, J. A. (2018). Identity and novelty in the avian syrinx. *Proceedings of the National Academy of Sciences*, *115*(41), 10209–10217. https://doi.org/10.1073/pnas.1804586115

Krause, B. I. (1987). Bioacoustics, habitat ambience in ecological balance. *Whole Earth Review*, *57*(472), 14–18.

Krause, B. I. (1993). The niche hypothesis: a virtual symphony of animal sounds, the origins of musical expression and the health of habitats. *The Soundscape Newsletter*, *6*(5).

Lachlan, R. F., & Servedio, M. R. (2004). Song learning accelerates allopatric speciation. *Evolution*, *58*(9), 2049–2063. https://doi.org/10.1111/j.0014-3820.2004.tb00489.x

Larsen, O. N., & Goller, F. (2002). Direct observation of syringeal muscle function in songbirds and a parrot. *Journal of Experimental Biology*, *205*(1), 25–35. https://doi.org/10.1242/jeb.205.1.25

Laverde, O. A. (2016). *Visual and acoustic communication signals in birds-taxonomy, evolution and transfer between signals mediated by sensory drive* [Trabajo de grado - Doctorado]. Universidad de los Andes.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Mauricio, G. N., Areta, J. I., Bornschein, M. R., & Reis, R. E. (2012). Morphology-based phylogenetic analysis and classification of the family Rhinocryptidae (Aves: Passeriformes). *Zoological Journal of the Linnean Society*, *166*(2), 377–432. https://doi.org/10.1111/j.1096-3642.2012.00847.x

Maynard-Smith, J., & Harper, D. (2003a). *Animal Signals*. Oxford University Press.

Maynard-Smith, J., & Harper, D. (2003b). The evolution of signal form. In *Animal Signals* (pp. 68–89). Oxford University Press.

Maynard-Smith, J., & Harper, D. (2003c). What are signals? In *Animal Signals* (pp. 1–15). Oxford University Press.

McCracken, K. G., & Sheldon, F. H. (1997). Avian vocalizations and phylogenetic signal. *Proceedings of the National Academy of Sciences*, *94*(8), 3833–3836. https://doi.org/10.1073/pnas.94.8.3833

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015). Librosa: audio and music signal analysis in python. *In Proceedings of the 14th Python in Science Conference*, 18–25.

McGinn, K., Kahl, S., Peery, M. Z., Klinck, H., & Wood, C. M. (2023). Feature embeddings from the BirdNET algorithm provide insights into avian ecology. *Ecological Informatics*, *74*, 101995. https://doi.org/10.1016/j.ecoinf.2023.101995

McInerney, P. L., Lee, M. S. Y., Clement, A. M., & Worthy, T. H. (2019). The phylogenetic significance of the morphology of the syrinx, hyoid and larynx, of the southern cassowary, Casuarius casuarius (Aves, Palaeognathae). *BMC Evolutionary Biology*, *19*(1), 233. https://doi.org/10.1186/s12862-019-1544-7

McInnes, L., Healy, J., & Melville, J. (2020). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv:1802.03426v3*.

Medina‐García, A., Araya‐Salas, M., & Wright, T. F. (2015). Does vocal learning accelerate acoustic diversification? Evolution of contact calls in Neotropical parrots. *Journal of Evolutionary Biology*, *28*(10), 1782–1792. https://doi.org/10.1111/jeb.12694

Mendelson, T. C., & Shaw, K. L. (2012). The (mis)concept of species recognition. *Trends in Ecology & Evolution*, *27*(8), 421–427. https://doi.org/10.1016/j.tree.2012.04.001

Mets, D. G., & Brainard, M. S. (2018). Genetic variation interacts with experience to determine interindividual differences in learned song. *Proceedings of the National Academy of Sciences*, *115*(2), 421–426. https://doi.org/10.1073/pnas.1713031115

Meyer, E., & Neumann, E.-G. (1972). Theory of Sound Fields. In E. Meyer & E.-G. Neumann (Eds.), *Physical and Applied Acoustics. An Introduction* (pp. 1–2). Academic Press, Inc.

Mindlin, G. B., & Laje, R. (2005). Anatomy of the Vocal Organ. In *The Physics of Birdsong* (pp. 37–46). Springer.

Mitteroecker, P., Collyer, M. L., & Adams, D. C. (2024). Exploring Phylogenetic Signal in Multivariate Phenotypes by Maximizing Blomberg’s K. *Systematic Biology*. https://doi.org/10.1093/sysbio/syae035

Mojsilovic, A., & Rogowitz, B. (2001). Capturing image semantics with low-level descriptors. *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, 18–21. https://doi.org/10.1109/ICIP.2001.958942

Monacchi, D. (2013). Fragments of Extinction: Acoustic Biodiversity of Primary Rainforest Ecosystems. *Leonardo Music Journal*, *23*, 23–25. https://doi.org/10.1162/LMJ\_a\_00148

Morales, G., Vargas, V., Espejo, D., Poblete, V., Tomasevic, J. A., Otondo, F., & Navedo, J. G. (2022). Method for passive acoustic monitoring of bird communities using UMAP and a deep neural network. *Ecological Informatics*, *72*, 101909. https://doi.org/10.1016/j.ecoinf.2022.101909

Morton, E. S. (1975). Ecological Sources of Selection on Avian Sounds. *The American Naturalist*, *109*(965), 17–34. https://doi.org/10.1086/282971

Morton, E. S. (1977). On the Occurrence and Significance of Motivation-Structural Rules in Some Bird and Mammal Sounds. *The American Naturalist*, *111*(981), 855–869. https://doi.org/10.1086/283219

Morton, E. S. (2017a). Form and Function in Vocal Communication. In *Animal vocal communication: assessment and management roles* (2nd ed., pp. 84–137). Cambridge University Press.

Morton, E. S. (2017b). The Informationizing of Communication. In *Animal Vocal Communication: Assessment and Management Roles* (2nd ed., pp. 10–39). Cambridge University Press.

Nelson, D. A., Hallberg, K. I., & Soha, J. A. (2004). Cultural Evolution of Puget Sound White‐Crowned Sparrow Song Dialects. *Ethology*, *110*(11), 879–908. https://doi.org/10.1111/j.1439-0310.2004.01025.x

Nowicki, S., & Marler, P. (1988). How do birds sing? *Music Perception*, *5*(4), 391–426. https://doi.org/10.2307/40285408

Odom, K. J., Araya‐Salas, M., Morano, J. L., Ligon, R. A., Leighton, G. M., Taff, C. C., Dalziell, A. H., Billings, A. C., Germain, R. R., Pardo, M., de Andrade, L. G., Hedwig, D., Keen, S. C., Shiu, Y., Charif, R. A., Webster, M. S., & Rice, A. N. (2021). Comparative bioacoustics: a roadmap for quantifying and comparing animal sounds across diverse taxa. *Biological Reviews*, *96*(4), 1135–1159. https://doi.org/10.1111/brv.12695

Pearse, W. D., Morales-Castilla, I., James, L. S., Farrell, M., Boivin, F., & Davies, T. J. (2018). Global macroevolution and macroecology of passerine song. *Evolution*, *72*(4), 944–960. https://doi.org/10.1111/evo.13450

Perraudin, N., Balazs, P., & Sondergaard, P. L. (2013). A fast Griffin-Lim algorithm. *2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 1–4. https://doi.org/10.1109/WASPAA.2013.6701851

Peters, S., & Nowicki, S. (1996). Development of Tonal Quality in Birdsong: Further Evidence from Song Sparrows. *Ethology*, *102*(2), 323–335. https://doi.org/10.1111/j.1439-0310.1996.tb01128.x

Podos, J. (1997). A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). *Evolution*, *51*(2), 537–551. https://doi.org/10.1111/j.1558-5646.1997.tb02441.x

Podos, J., Moseley, D. L., Goodwin, S. E., McClure, J., Taft, B. N., Strauss, A. V. H., Rega-Brodsky, C., & Lahti, D. C. (2016). A fine-scale, broadly applicable index of vocal performance: frequency excursion. *Animal Behaviour*, *116*, 203–212. https://doi.org/10.1016/j.anbehav.2016.03.036

Podos, J., & Nowicki, S. (2004). Performance limits on birdsong. In P. Marler & H. Slabbekoorn (Eds.), *Nature’s Music: The Science of Birdsong* (pp. 318–342). Elsevier Academic Press.

Podos, J., & Sung, H.-C. (2020). Vocal Performance in Songbirds: From Mechanisms to Evolution. In J. T. Sakata, S. C. Woolley, R. R. Fay, & A. N. Popper (Eds.), *The Neuroethology of Birdsong* (pp. 245–263). Springer Nature.

Popp, J. W., Ficken, R. W., & Reinartz, J. A. (1985). Short-term temporal avoidance of interspecific acoustic interference among forest birds. *The Auk*, *102*(4), 744–748.

Powell, A. F. L. A., Barker, F. K., Lanyon, S. M., Burns, K. J., Klicka, J., & Lovette, I. J. (2014). A comprehensive species-level molecular phylogeny of the New World blackbirds (Icteridae). *Molecular Phylogenetics and Evolution*, *71*, 94–112. https://doi.org/10.1016/j.ympev.2013.11.009

Price, J. J. (2004). Patterns of song evolution and sexual selection in the oropendolas and caciques. *Behavioral Ecology*, *15*(3), 485–497. https://doi.org/10.1093/beheco/arh040

Price, J. J. (2013). Why is birdsong so repetitive? Signal detection and the evolution of avian singing modes. *Behaviour*, *150*(9–10), 995–1013. https://doi.org/10.1163/1568539X-00003051

Price, J. J., & Lanyon, S. M. (2002). Reconstructing the evolution of complex bird song in the oropendolas. *Evolution*, *56*(7), 1514–1529. https://doi.org/10.1111/j.0014-3820.2002.tb01462.x

Provost, K. L., Yang, J., & Carstens, B. C. (2022). The impacts of fine-tuning, phylogenetic distance, and sample size on big-data bioacoustics. *PLOS ONE*, *17*(12), e0278522. https://doi.org/10.1371/journal.pone.0278522

Rajan, S., Lamers, K. P., Both, C., & Wheatcroft, D. (2024). Translocated wild birds are predisposed to learn songs of their ancestral population. *Current Biology*, *34*(11), 2535-2540.e4. https://doi.org/10.1016/j.cub.2024.04.061

Revell, L. J. (2024a). Ancestral state reconstruction of phenotypic characters. *EcoEvoRxiv (Preprint, Version 2)*.

Revell, L. J. (2024b). phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, *12*, e16505. https://doi.org/10.7717/peerj.16505

Revell, L. J., & Collar, D. C. (2009). Phylogenetic analysis of the evolutionary correlation using likelihood. *Evolution*, *63*(4), 1090–1100. https://doi.org/10.1111/j.1558-5646.2009.00616.x

Revell, L. J., & Harmon, L. J. (2022). *Phylogenetic comparative methods in R*. Princeton University Press.

Revell, L. J., Harmon, L. J., & Collar, D. C. (2008). Phylogenetic Signal, Evolutionary Process, and Rate. *Systematic Biology*, *57*(4), 591–601. https://doi.org/10.1080/10635150802302427

Revell, L. J., Toyama, K. S., & Mahler, D. L. (2022). A simple hierarchical model for heterogeneity in the evolutionary correlation on a phylogenetic tree. *PeerJ*, *10*, e13910. https://doi.org/10.7717/peerj.13910

Riebel, K., Lachlan, R. F., & Slater, P. J. B. (2015). *Learning and Cultural Transmission in Chaffinch Song* (pp. 181–227). https://doi.org/10.1016/bs.asb.2015.01.001

Riede, T., Thomson, S. L., Titze, I. R., & Goller, F. (2019). The evolution of the syrinx: An acoustic theory. *PLOS Biology*, *17*(2), e2006507. https://doi.org/10.1371/journal.pbio.2006507

Rocha, P. C., & Romano, P. S. R. (2021). The shape of sound: A new R package that crosses the bridge between Bioacoustics and Geometric Morphometrics. *Methods in Ecology and Evolution*, *12*(6), 1115–1121. https://doi.org/10.1111/2041-210X.13580

Rubenstein, D. R., & Alcock, J. (2019a). Principles of communication. In *Animal Behavior* (11th ed., pp. 257–301). Oxford University Press.

Rubenstein, D. R., & Alcock, J. (2019b). The Integrative Study of Behavior. In *Animal Behavior* (11th ed., pp. 24–57). Oxford University Press.

Russell, A. P., & Bauer, A. M. (2021). Vocalization by extant nonavian reptiles: A synthetic overview of phonation and the vocal apparatus. *The Anatomical Record*, *304*(7), 1478–1528. https://doi.org/10.1002/ar.24553

Ryan, M. J., & Rand, A. S. (1993). Species recognition and sexual selection as a unitary problem in animal communication. *Evolution*, *47*(2), 647–657. https://doi.org/10.1111/j.1558-5646.1993.tb02118.x

Sakata, J. T., & Vehrencamp, S. L. (2012). Integrating perspectives on vocal performance and consistency. *Journal of Experimental Biology*, *215*(2), 201–209. https://doi.org/10.1242/jeb.056911

Sauquet, H., von Balthazar, M., Magallón, S., Doyle, J. A., Endress, P. K., Bailes, E. J., Barroso de Morais, E., Bull-Hereñu, K., Carrive, L., Chartier, M., Chomicki, G., Coiro, M., Cornette, R., El Ottra, J. H. L., Epicoco, C., Foster, C. S. P., Jabbour, F., Haevermans, A., Haevermans, T., … Schönenberger, J. (2017). The ancestral flower of angiosperms and its early diversification. *Nature Communications*, *8*(1), 16047. https://doi.org/10.1038/ncomms16047

Seddon, N. (2005). Ecological adaptation and species recognition drives vocal evolution in neotropical suboscine birds. *Evolution*, *59*(1), 200–215. https://doi.org/10.1111/j.0014-3820.2005.tb00906.x

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sitt, J. D., Amador, A., Goller, F., & Mindlin, G. B. (2008). Dynamical origin of spectrally rich vocalizations in birdsong. *Physical Review E*, *78*(1), 011905. https://doi.org/10.1103/PhysRevE.78.011905

Sockman, K. W., Lyons, S. M., & Caro, S. P. (2024). Vocal performance of one affects that of another. *Behaviour*, 1–28. https://doi.org/10.1163/1568539X-bja10256

Stiller, J., Feng, S., Chowdhury, A.-A., Rivas-González, I., Duchêne, D. A., Fang, Q., Deng, Y., Kozlov, A., Stamatakis, A., Claramunt, S., Nguyen, J. M. T., Ho, S. Y. W., Faircloth, B. C., Haag, J., Houde, P., Cracraft, J., Balaban, M., Mai, U., Chen, G., … Zhang, G. (2024). Complexity of avian evolution revealed by family-level genomes. *Nature*, *629*(8013), 851–860. https://doi.org/10.1038/s41586-024-07323-1

Ten Cate, C. (2004). Birdsong and evolution. In P. Marler & H. Slabbekoorn (Eds.), *Nature’s music: The Science of Birdsong* (pp. 296–317). Elsevier Academic Press.

ten Cate, C. (2021). Re-evaluating vocal production learning in non-oscine birds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1836). https://doi.org/10.1098/rstb.2020.0249

Tinbergen, N. (1952). “Derived” Activities; Their Causation, Biological Significance, Origin, and Emancipation During Evolution. *The Quarterly Review of Biology*, *27*(1), 1–32. https://doi.org/10.1086/398642

Tobias, J. A., Planqué, R., Cram, D. L., & Seddon, N. (2014). Species interactions and the structure of complex communication networks. *Proceedings of the National Academy of Sciences*, *111*(3), 1020–1025. https://doi.org/10.1073/pnas.1314337111

Toledo, L. F., Martins, I. A., Bruschi, D. P., Passos, M. A., Alexandre, C., & Haddad, C. F. B. (2015). The anuran calling repertoire in the light of social context. *Acta Ethologica*, *18*(2), 87–99. https://doi.org/10.1007/s10211-014-0194-4

Verzijden, M. N., ten Cate, C., Servedio, M. R., Kozak, G. M., Boughman, J. W., & Svensson, E. I. (2012). The impact of learning on sexual selection and speciation. *Trends in Ecology & Evolution*, *27*(9), 511–519. https://doi.org/10.1016/j.tree.2012.05.007

Villanueva-Rivera, L. J. (2014). Eleutherodactylus frogs show frequency but no temporal partitioning: implications for the acoustic niche hypothesis. *PeerJ*, *2*, e496. https://doi.org/10.7717/peerj.496

Wan, X., Wang, A. S., Storch, D.-S., Li, V. Y., & Sakata, J. T. (2024). Perineuronal nets in motor circuitry regulate the performance of learned vocalizations in songbirds. *BioRxiv: The Preprint Server for Biology. 2024.05.14.593930*.

Widrig, K., & Field, D. J. (2022). The Evolution and Fossil Record of Palaeognathous Birds (Neornithes: Palaeognathae). *Diversity*, *14*(2), 105. https://doi.org/10.3390/d14020105

Wiley, R. H. (2015). Evolution of signalers and receivers. In *Noise matters: the evolution of communication* (pp. 127–268). Harvard University Press.

Wilkins, M. R., Seddon, N., & Safran, R. J. (2013). Evolutionary divergence in acoustic signals: causes and consequences. *Trends in Ecology & Evolution*, *28*(3), 156–166. https://doi.org/10.1016/j.tree.2012.10.002

Williams, H., Scharf, A., Ryba, A. R., Ryan Norris, D., Mennill, D. J., Newman, A. E. M., Doucet, S. M., & Blackwood, J. C. (2022). Cumulative cultural evolution and mechanisms for cultural selection in wild bird songs. *Nature Communications*, *13*(1), 4001. https://doi.org/10.1038/s41467-022-31621-9

Youngworth, R. N., Gallagher, B. B., & Stamper, B. L. (2005). *An overview of power spectral density (PSD) calculations* (H. P. Stahl, Ed.; p. 58690U). https://doi.org/10.1117/12.618478

Zollinger, S. A., & Brumm, H. (2015). Why birds sing loud songs and why they sometimes don’t. *Animal Behaviour*, *105*, 289–295. https://doi.org/10.1016/j.anbehav.2015.03.030

Zollinger, S. A., & Suthers, R. A. (2004). Motor mechanisms of a vocal mimic: implications for birdsong production. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1538), 483–491. https://doi.org/10.1098/rspb.2003.2598