UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

DOCTORADO EN CIENCIAS BIOMÉDICAS

INSTITUTO DE ECOLOGÍA

DIVERSIDAD DE ARTRÓPODOS COMO INDICADOR ESTRUCTURAL DE LOS BOSQUES DE LA FAJA VOLCÁNICA TRANSMEXICANA

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTORA EN CIENCIAS

PRESENTA:

NANCY GÁLVEZ REYES

TUTOR PRINCIPAL DE TESIS: DR. DANIEL PIÑERO DALMAU
INSTITUTO DE ECOLOGÍA, UNAM
COTUTORA PRINCIPAL DE TESIS: DRA. ALICIA MASTRETTA YANES
COMISIÓN NACIONAL PARA EL CONOCIMIENTO Y USO DE LA BIODIVERSIDAD
COMITÉ TUTOR: DRA. ELLA VÁZQUEZ DOMÍNGUEZ
INSTITUTO DE ECOLOGÍA, UNAM
COMITÉ TUTOR DR. ANTONIO GONZÁLEZ RODRÍGUEZ
INSTITUTO DE INVESTIGACIONES EN ECOSISTEMAS Y SUSTENTABILIDAD, UNAM

MÉXICO, CD. MX. **FEBRERO, 2024.**

Agradecimientos

Agradezco profundamente al Programa de Doctorado en Ciencias Biomédicas, al Instituto de Ecología y a la Universidad Nacional Autónoma de México por la excepcional formación doctoral que me brindaron y por su invaluable contribución para hacer posible la realización de esta tesis.

Expreso mi sincero agradecimiento al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la generosa beca número 401781 (CVU: 412884) que me otorgó, permitiéndome llevar a cabo mis estudios con dedicación y compromiso.

Agradezco al proyecto que respaldó mi investigación mediante el financiamiento del Consejo Nacional de Ciencia y Tecnología, a través del proyecto número 178245, titulado "Efectos del paisaje y la historia en la estructuración de la biodiversidad: un análisis filogeográfico de alta resolución en plantas de las montañas de México". Este respaldo financiero fue crucial para llevar a cabo mi trabajo de investigación de manera exitosa.

Quiero expresar mi gratitud a los distinguidos miembros de mi comité tutor: la Dra. Alicia Mastretta Yanes, el Dr. Daniel Piñero Dalmau, la Dra. Ella Vázquez Domínguez y el Dr. Antonio González Rodríguez. Por tan acertadas opiniones y constante apoyo, que fueron fundamentales para el desarrollo y éxito de esta tesis. Agradezco a cada uno, su dedicación y experiencia en sus respectivos campos que enriquecieron mi investigación y fueron una guía fundamental en el proceso de desarrollo de mi trabajo doctoral.

Agradezco al laboratorio de Genética y Ecología del Instituto de Ecología de la UNAM por brindarme acceso a sus instalaciones, permitiéndome llevar a cabo de manera efectiva el procesamiento de las muestras y análisis de datos.

Mi reconocimiento especial a la Dra. Tania Garrido, la Dra. Mariana Rojas, la Dra. Azalea Guerra y la Odontóloga Anabel Domínguez por su valioso apoyo técnico en el Laboratorio de Genética y Ecología del Instituto de Ecología de la UNAM.

Expreso mi sincero agradecimiento a mis distinguidas sinodales, quienes desempeñaron un papel fundamental para que mi tesis encontrara la luz y han dejado una huella invaluable en mi trabajo, brindándome comentarios valiosos y aportaciones significativas que han contribuido de manera sustancial a la mejora de mi tesis. Aprecio enormemente su tiempo, esfuerzo y compromiso con la excelencia académica, lo cual ha sido fundamental para alcanzar los estándares de calidad para mi tesis. A cada una de ellas, mi agradecimiento infinito por su generosidad intelectual y por su contribución a la culminación exitosa de este proyecto académico.

Dra. Valeria Souza Saldivar	Presidenta
Dra. Alicia Mastretta Yanes	Secretaria
Dra. Ana E. Escalante Hernández	Vocal
Dra. Eria A. Rebollar Caudillo	Vocal
Dra. Ek Del Val De Gortari	Vocal

Agradecimientos personales

Dedico esta tesis a mi querida hermana Sabina Gabrex y mi dulce sobrina Farah Mila que son mi fortaleza y mi esperanza.

El doctorado es como una montaña tropical con laderas y valles, aunque los últimos 50 metros para llegar a la cima de la montaña cuesta el doble, pero cuando llegas puedes ver el cielo más azul. Me doy cuenta de que he sido muy afortunada al encontrar a muchas personas en mi trayectoria en el doctorado que me tendieron la mano. Gracias a la colaboración de todos ellos, mi vida académica fue más amena.

Cuando decidí seguir mi carrera científica le pedí a la vida que por favor me hiciera conocer a personas con humanidad en la ciencia y me encontré con dos senséis llenos de humanidad y muchas virtudes, **Alicia Mistretta** y **Daniel Piñero**. Si llego a ser una gran científica, quiero ser como ustedes, porque gracias a ustedes mi parámetro y referencia son muy altos, de cómo hacer ciencia con humanidad, ética profesional y la capacidad de enseñanza. Gracias a ambos por su infinita serenidad y paciencia.

Alicia Mastretta, la vida te puso en mi camino para que me enseñaras a como necesita hacer ciencia, y trabajar en equipo en el campo. Gracias a ti aprendí cómo se tienen que hacer las cosas en el laboratorio de biología molecular, a enseñarme a escribir un artículo científico. Gracias por enviarme a una estancia a Tenerife y un congreso a Finlandia y pagarme el boleto redondo de avión al otro lado del mundo en esa estancia. Gracias por darme una beca después que se me acabó la beca de Conacyt, gracias por poner confianza en mí y tu proyecto con el Jardín Botánico, UNAM y la Universidad de Vermont. Gracias por brindarme tu confianza en el proyecto teocintle con el Jardín Botánico y el instituto IBG-4 Bioinformatics Forschungszentrum Jülich. También, mil gracias cuando enviamos el primer artículo y me diste alas para volar en organizarme y yo misma subir a una plataforma el envío de nuestro primer artículo, a que yo pusiera mis scripts en un repositorio de github para que fuera público (porque siempre dices que la ciencia tiene que ser reproducible y publica) y sí sé bioinformática en gran parte es gracias a ti. Aunque, algunas veces me he sentido que soy un huevo de Kiwii en tu vida (y tú el ave Kiwii). Recuerdo cuando muchas veces platicamos y me sirvió de destrabar la situación emocional y mis miles dudas bioinformáticas y de laboratorio. ¡Gracias a ti conocí cómo hacer ciencia con humanidad!.

Daniel Piñero, agradezco infinitamente en abrirme las puertas de tu laboratorio desde la maestría, me adoptaste y pusiste tu confianza en mí sin conocernos. En mi doctorado me proveíste de tus recursos para poner en marcha el proyecto que gestamos Alicia, tu y yo, pagar mis salidas de campo, comprarme reactivos de biología molecular para poner en marcha las extracciones de ADN y la construcción de las bibliotecas, pagaste mis placas de secuenciación, y una beca cuando se me terminó el apoyo de mi beca conacyt, darme los mejores consejos asertivos en mi vida personal y profesional (siempre dices que la ciencia es divertida y se puede disfrutar). Cuando he sentido que se me acaba el mundo, tú siempre tienes una solución.

Quiero expresar mi sincero agradecimiento a mis estimad@s jef@s en mi trabajo actual: **Daniel Piñero, Ella Vázquez, Juan Pablo Jaramillo, Luis Osorio** y **Nadia Santini** (que fuiste parte del equipo). Les estoy profundamente agradecida por más que poner comida en mi mesa; su liderazgo y apoyo han sido fundamentales en mi vida profesional y académica. La gratitud que siento por tenerlos como jefes va más allá de las palabras. Su respaldo constante y palabras de aliento no solo me han motivado a concluir mi tesis, sino que también han sido un faro de inspiración en mi día a día. La combinación de su guía, experiencia y apoyo ha sido esencial para mi crecimiento tanto profesional como personal. Agradezco sinceramente el ambiente de trabajo que han cultivado, que va más allá de las responsabilidades laborales y se traduce en un espacio donde el estímulo mutuo y la colaboración florecen.

Nadia Santini, gracias a ti pude conocer cómo hacer una cover letter y escribir un artículo dentro de tu curso. Gracias por impulsar mi artículo de la maestría sobre la genética de la mosca de la fruta (que fue enviada a la revista que sugeriste). Gracias por darme esas palabras de ánimo en todo momento con esa frescura científica de siempre ver las cosas con optimismo. Gracias por apoyarme en obtener mi primer trabajo.

Quisiera expresar mi más profundo agradecimiento a **Luis Osorio** y a su querida esposa **Rusby Contreras** por su apoyo incondicional durante las últimas etapas de la elaboración de esta tesis. A pesar de los desafíos extraordinarios que han enfrentado debido a la batalla de Luis contra el cáncer, su presencia constante en el lab, aliento y apoyo han sido un faro de luz en mi camino, y deseo que la amistad de ambos continúe iluminando mi camino en los años venideros. La valentía y la determinación de Luis ante la adversidad son una inspiración para tod@s nosotr@s. Su optimismo y su fuerza de voluntad para enfrentar cada día con coraje y esperanza han dejado una huella indeleble en mi corazón y en mi vida. La presencia de Rusby ha sido igualmente reconfortante y alentadora. Su amor incondicional, cuidado y apoyo hacia Luis, ella siempre estar al pie del volcán, han sido una muestra del

verdadero significado del amor. Su generosidad y amistad han sido un regalo invaluable que siempre recordaré.

Agradezco inmensamente al equipo GUND, en especial a **Yolanda Chen, Alicia Mastretta, Ana Wegier, Jorge Ruiz** y **Ana Sofia Monroy** por darme la oportunidad de colaboración en el proyecto entre la Universidad de Vermont, CONABIO y Jardín Botánico, UNAM. Estoy muy agradecida por la oportunidad de aprender y crecer bajo su liderazgo, y valoro enormemente la influencia positiva que han tenido en mi trayectoria científica. **Ana Wegier** mil gracias por adoptarme en tu laboratorio y orientarme en la entrevista de mi actual trabajo, tus palabras en la llamada telefónica fueron asertivas en esa entrevista de trabajo.

Quiero expresar mi profundo agradecimiento a mis querid@s coleg@s y amig@s del laboratorio de Genética y Ecología. Compartir tiempo con ustedes ha sido una experiencia invaluable y su apoyo moral ha sido un pilar esencial en mi camino hacia la culminación de esta tesis. **Rocío González, Alfredo Villarruel, Ruth Percino, Raquel Hernández, Myriam Campos, Verónica González, Verónica Reyes, Alejandro Flores, Natalia Castillo, Elsa Piter, Fernanda Rosa, Gema Peña, Diana Cárdenas, Mariana Rojas,** y **Linda M. Martínez**, cada uno de ustedes ha dejado una huella imborrable en mi viaje académico y personal. Sus palabras de aliento y apoyo han sido un bálsamo en los momentos difíciles, y el sentido de comunidad que hemos construido juntos ha enriquecido mi experiencia de investigación de maneras que van más allá de las páginas de esta tesis. **Gustavo Ibrahim**, quiero agradecerte de manera especial. Tu presencia y tus palabras han sido una luz en los pasillos del laboratorio. En los momentos de dificultad, tus alentadoras palabras y tus hombros disponibles para consolarme han hecho una diferencia significativa. Tu habilidad para brindar ánimo y ofrecer perspectivas renovadoras ha sido un regalo invaluable que siempre recordaré con gratitud. A cada uno de ustedes, mi más sincero agradecimiento por formar parte de este viaje, por compartir risas, desafíos y triunfos. Este logro no habría sido lo mismo sin su presencia y contribución.

A mi querida hermana **Sabina Gálvez**, mi guía y cómplice, agradezco infinitamente tu apoyo constante. Siempre estás ahí para sostenerme cuando el mundo parece caer sobre Nancy, y con solo un abrazo tuyo, logras calmar mi ansiedad, brindarme amor y transmitir esa paz tan necesaria en tantas ocasiones. Tu dedicación va más allá de las palabras, como quedó demostrado al invertir tus vacaciones en pro de la biodiversidad y ayudarme a construir 3 mil trampas de caída para el muestreo. Este gesto no solo habla de tu generosidad, sino también de tu compromiso con mi trabajo y con nuestros ideales compartidos. Además, no puedo dejar de agradecerte por el mejor regalo del mundo: Farah. Ella ha llegado a ser una parte invaluable de nuestras vidas. Tu contribución no solo ha sido en apoyo logístico, sino

que has enriquecido mi existencia con esta adición tan especial. Gracias por ser una luz en mi camino y por compartir este viaje conmigo. Tu amor y apoyo son invaluables, y estoy agradecida por cada momento que hemos compartido y por aquellos que están por venir.

A mis amigos de senderismo **Julio Cesar García** y **Giezi Anthony Gálvez**, quien les tocó verme trabajar enfrente de una computadora y muchas veces me hicieron los desayunos y el café. Por enseñarme que los fines de semana puede ser un respiro e irnos de patas de perro. A ti **Cesar**, porque siempre estuviste desde que inicié mi carrera científica. A ti te contaba mis sueños de pisar la tierra de los pumas (UNAM) y siempre dijiste que yo tenía la capacidad de hacer lo que yo quisiera. A ti **Giezi**, gracias por tenerme mucha paciencia en enseñarme a manejar un carro, y gracias a tu espíritu de biólogo por guiarme al senderismo.

Al equipo de colecta, que sin ellos mi proyecto simplemente no hubiera ocurrido. A ti **Alicia Mastretta** por su enorme apoyo en salidas previas en buscar las comunidades para establecer el muestreo. A **Daniel Piñero** por su apoyo en la logística, manejo de vehículo en campo y colocar trampas. A **Julio Cesar García** por ir a todas mis salidas de campo, tanto poner las tramas, así como, levantar las trampas y coordinar el equipo correspondiente. A **Alfredo Villarruel, Azalea Guerra,** y **Jorge Jiménez** por ser los choferes de cabecera de todas las salidas y cargar los garrafones con alcohol y colocar trampas. A **Verónica Reyes**, por siempre decirme si como buena bióloga y me ayudó a ir a poner todas las trampas. A **Gustavo Ibrahim** por ayudarme a coordinar un equipo que le tocó medir los cuadrantes. A ustedes **Ana Silva, Myriam Campos**, **Idalia Rojas** y **Marco Garduño** por su apoyo en hacer hoyos y colocar y/o levantar las trampas con la mezcla.

Gracias al personal de las comunidades de **Santiago Tlacotepec, Estado de México** y al **Consejo Civil Mexicano para la Silvicultura Sustentable** en Amanalco Estado de México, por su apoyo logístico en el muestreo y permitirnos hacer la colecta en sus tierras. A la **Dra. Lucia Madrid, Biol. Carlos Jiménez,** y **C. Zeferino Espinosa** por apoyarnos en darnos asistencia en el muestreo en Amanalco.

A **Minerva Soto, Sabina Gálvez** y **Gala Soto,** por su apoyo en el corte y separación de las muestras junto conmigo.

En mi estancia por Tenerife, España, a **Brent Emerson** por permitirme llegar a su grupo de investigación Island Ecology and Evolution en el Instituto de Productos Naturales y Agrobiología (IPNA-CSIC). **Brent** gracias por pagarme el seguro de vida en mi estancia, empujar el primer artículo y siempre estar al pendiente de mi en la estancia. A **Paula Arribas** y **Carmelo Andújar**, por enseñarme a analizar sopas de la biodiversidad, gracias a ustedes

aprendí toda la parte bioinformática de este trabajo, gracias por su infinita paciencia y apoyo moral en todo momento. A **Martha López, Julien Piquet,** y **Víctor García** porque siempre estuvieron pendientes de mí, invitarme a sus fiestas y llevarme a sus colectas a Teno Bajo y Anaga, y conocer un Guachinche. A **Julien Piquet** gracias por rescatarme la noche en que perdí mis llaves. **Martha**, gracias a ti conocí el Teide. Les recuerdo con mucho cariño y aún sigo en espera de su llegada a México.

Agradezco a **Armando Sunny** y **Andrea González** por proporcionar el ráster de tipos de vegetación para analizar parte de los datos de las sopas de la biodiversidad.

A **Daniel Ortiz** y **Wolke Tobon** por su ayuda con las transformaciones de datos espaciales.

Los análisis se realizaron utilizando el clúster de cómputo de CONABIO con la asistencia de **Ernesto Campos** (Grillo), muchas gracias por el soporte de TI en todo momento. Aunque era un viernes por la tarde y en puente, tú estabas allí para levantar los nodos.

Contenido

RESUMEN

La deforestación de los bosques contribuye significativamente a la pérdida de biodiversidad y a la disminución de los servicios ecosistémicos. Como estrategia para combatir esta degradación, se ha propuesto la restauración de áreas deforestadas a través de plantaciones forestales, así como la reducción de la tasa de deforestación en los bosques aún existentes promoviendo el manejo forestal sustentable como alternativa a la tala ilegal. Sin embargo, para garantizar el cumplimiento de estos objetivos a nivel local, es necesario contar con una línea base de conocimiento sobre la diversidad local y realizar monitoreos periódicos para determinar si las plantaciones, el manejo forestal y los procesos de dispersión neutral al facilitar el flujo de genes entre poblaciones, puede favorecer en la configuración de la dinámica evolutiva de las especies y prevalencia de la biodiversidad de las comunidades faunísticas.

Los artrópodos son especialmente apropiados para este propósito, ya que la diversidad de estos organismos contribuye un sólido indicador de la integridad estructural de los ecosistemas debido a su riqueza, abundancia, diversidad de nichos y a la presencia de especies con historias biogeográficas diferentes en México. Por consiguiente, el análisis de comunidades completas de artrópodos puede proporcionar los datos necesarios para investigar si los cambios en la composición de las comunidades se deben a la distancia geográfica o a los tratamientos dentro del bosque, desde una perspectiva ecológica, y brindar recomendaciones de manejo. Por lo tanto, el objetivo principal de este estudio

consistió en detectar la variación espacial y temporal en la composición de la comunidad de artrópodos en los ecosistemas de la FVTM (Faja Volcánica Transmexicana), considerando las condiciones ambientales, las historias filogeográficas, las diferencias en el manejo y el estado de conservación a escala fina.

Para lograr este objetivo, realizamos una revisión de los conceptos de monitoreo y de las comunidades biológicas de artrópodos como indicador ecológico, además de considerar aspectos metodológicos que abarcan desde el diseño de muestreo hasta la bioinformática aplicada al metabarcoding de toda la comunidad (wcMBC, por sus siglas en inglés) y los flujos de análisis de datos basados en variantes de secuencias de ampliaciones (ASV, por sus siglas en inglés). Luego, evaluamos múltiples niveles jerárquicos en bosques bajo conservación, desde haplotipos individuales hasta linajes utilizando umbrales de similitud de 0.5, 1.5, 3.0, 5.0, 7.5%, esto nos permitió analizar patrones de riqueza, rotación y disminución de distancia de similitud utilizando enfoques de aislamiento por distancia y aislamiento por resistencia (este último, considerando los costos de dispersión determinados por las características del paisaje). Finalmente, exploramos múltiples niveles jerárquicos, desde haplotipos hasta linajes en umbrales de similitud de 3% y 5%. Evaluamos patrones de riqueza, rotación y disminución de la distancia de similitud con un enfoque de aislamiento por distancia para analizar el impacto de las limitaciones de dispersión, el manejo forestal, la reforestación, el disturbio y la conservación como impulsores de la diferenciación de las comunidades dentro de la isla del cielo tropical Nevado de Toluca.

Los resultados del primer estudio muestran que el éxito de wcMBC depende de buenas prácticas de laboratorio, que incluyen la correcta limpieza de las muestras a granel y la clasificación por tamaño de los artrópodos, la eficiencia de la extracción y cuantificación del ADN en la construcción de bibliotecas, la elección adecuada de marcadores genéticos para la identificación taxonómica, el filtrado de secuencias y curación de datos de referencias.

Los resultados del segundo estudio respaldan un modelo en el cual la diferenciación a nivel local está mediada por restricciones en la dispersión, combinada con la persistencia a largo plazo de los linajes, lo que constituye un importante factor impulsor de la diversidad dentro de las islas del cielo tropical. Finalmente, los resultados del tercer estudio muestran una baja riqueza de haplotipos en los bosques bajo conservación y manejo forestal, y revelan que la estructura y composición de las comunidades varían en función de los diferentes tratamientos dentro del bosque. En conclusión, la distancia, la elevación, el manejo forestal, la reforestación y el disturbio influyen en la disminución de la similitud en todos los niveles jerárquicos dado que cada grupo de artrópodos presenta diferentes habilidades de dispersión según la escala espacial y las actividades antropogénicas. Esto destaca su relevancia como impulsores de la diversidad dentro de una isla de cielo tropical. Todos estos hallazgos pueden informar sobre los índices de biodiversidad de las comunidades, promover políticas, conducir pautas efectivas de manejo y estrategias para la conservación de la biodiversidad en los ecosistemas forestales.

ABSTRACT

Deforestation of forests significantly contributes to biodiversity loss and the decline of ecosystem services. As a strategy to combat this degradation, the restoration of deforested areas through forest plantations has been proposed, along with reducing the deforestation rate in existing forests by promoting sustainable forest management as an alternative to illegal logging. However, to ensure the fulfillment of these objectives at the local level, it is necessary to have a knowledge baseline of local diversity and to conduct periodic monitoring to determine if plantations, forest management, and neutral dispersal processes facilitating gene flow among populations can favor the configuration of species' evolutionary dynamics and the prevalence of biodiversity in faunal communities.

Arthropods are particularly suitable for this purpose, as the diversity of these organisms provides a solid indicator of the structural integrity of ecosystems due to their richness, abundance, niche diversity, and the presence of species with different biogeographic histories in Mexico. Therefore, the analysis of complete arthropod communities can provide the necessary data to investigate whether changes in community composition are due to geographic distance or treatments within the forest, from an ecological perspective, and provide management recommendations. Thus, the main objective of this study was to detect spatial and temporal variation in the composition of the arthropod community in ecosystems of the FVTM (Transmexican Volcanic Belt), considering environmental

conditions, phylogeographic histories, differences in management, and conservation status on a fine scale.

To achieve this objective, we conducted a review of monitoring concepts and arthropod biological communities as ecological indicators, in addition to considering methodological aspects ranging from sampling design to bioinformatics applied to whole-community metabarcoding (wcMBC) and data analysis flows based on amplicon sequence variants (ASVs). We then evaluated multiple hierarchical levels in forests under conservation, from individual haplotypes to lineages using similarity thresholds of 0.5, 1.5, 3.0, 5.0, 7.5%, which allowed us to analyze patterns of richness, turnover, and similarity distance decrease using distance isolation and resistance isolation approaches (the latter considering dispersal costs determined by landscape characteristics). Finally, we explored multiple hierarchical levels, from haplotypes to lineages at similarity thresholds of 3% and 5%. We evaluated richness patterns, turnover, and similarity distance decrease with a distance isolation approach to analyze the impact of dispersal limitations, forest management, reforestation, disturbance, and conservation as drivers of community differentiation within the tropical sky island Nevado de Toluca.

The results of the first study show that the success of wcMBC depends on good laboratory practices, including proper cleaning of bulk samples and sorting of arthropods by size, efficiency of DNA extraction and quantification in library construction, appropriate choice of genetic markers for taxonomic identification, sequence filtering, and reference data

curation. The results of the second study support a model in which local-level differentiation is mediated by dispersal constraints, combined with long-term lineage persistence, which constitutes a significant driver of diversity within tropical sky islands. Finally, the results of the third study show low haplotype richness in forests under conservation and forest management and reveal that community structure and composition vary depending on different treatments within the forest. In conclusion, distance, elevation, forest management, reforestation, and disturbance influence the decrease in similarity at all hierarchical levels since each group of arthropods presents different dispersal abilities according to spatial scale and anthropogenic activities. This highlights their relevance as diversity drivers within a tropical sky island. All these findings can inform community biodiversity indices, promote policies, lead to effective management guidelines, and strategies for biodiversity conservation in forest ecosystems.

INTRODUCCIÓN

1.1. La deforestación y estrategias de conservación

La deforestación de los bosques contribuye a la pérdida de la biodiversidad y a la disminución de los servicios ecosistémicos. Hasta el 2002, México había perdido el 38% de su cobertura forestal original (Challenger y Dirzo, 2009). Como estrategia para abordar esta degradación, se ha propuesto la recuperación de áreas deforestadas a través de plantaciones forestales, y la reducción de la tasa de deforestación en los bosques que aún subsisten mediante la promoción del manejo forestal sustentable como una alternativa a la tala ilegal y el cambio climático (Back et al., 2020; Torres-Rojo et al., 2016). Estas estrategias, por ejemplo, forman parte de los argumentos en la reclasificación del Área Natural Protegida (ANP) Nevado de Toluca (Depraz et al., 2017; Mastretta-Yanes et al., 2014). A nivel internacional se ha postulado que las plantaciones pueden desempeñar un papel positivo en la protección y mejora de diversidad nativa, siempre y cuando no reemplacen por completo los ecosistemas naturales o se realicen con monocultivos de especies no nativas (Malcolm et al., 2001; Forestry Commission, 2011). No obstante, las plantaciones y otros bosques manejados intensivamente a menudo albergan menos biodiversidad que los bosques nativos (Wang et al., 2019; Zhang et al., 2016). En México, la Ley General de Desarrollo Forestal Sustentable establece que el manejo forestal puede tener por objeto tanto el aprovechamiento como la conservación de un ecosistema forestal y debe basarse en principios ecológicos, respetando la integralidad funcional e interdependencia de los recursos (Secretaría de Medio Ambiente y Recursos Naturales, 2018). Sin embargo, para

garantizar el cumplimiento de estos objetivos a nivel local, es esencial contar con una línea base de conocimiento de la diversidad local y realizar monitoreos periódicos.

1.2. Desafío en el monitoreo de la biodiversidad

El monitoreo de la biodiversidad consiste en documentar los cambios temporales en la distribución y presencia de las especies (Hajibabaei et al., 2016; Li et al., 2010). Sin embargo, la disponibilidad de información, la capacidad técnica y financiera para realizar el monitoreo y analizar las tendencias de cambio son limitados para permitir comparaciones a gran escala (Buss et al., 2015). Anteriormente, se han desarrollado indicadores alternativos, como el uso de especies "paraguas" (Lambeck, 1997; Launer y Murphy, 1994), y especies “banderas” (Noss, 1990), así como sustitutos que se refiere a medidas o indicadores alternativos basados en el hábitat (indicadores estructurales, e. g., volúmenes de madera muerta, cubierta de dosel, edad de los bosques, y complejidad estructural) que se utilizan como fuente de información sobre el estado de la comunidad biológica (Barsoum et al., 2019; Bruce, 2013; Coote et al., 2013). También se han utilizado los sustitutos basados en taxones (indicadores de composición, e.g., carábidos, sírfidos, arañas, y plantas vasculares) que son indicadores informativos de la riqueza de otros taxones en entornos forestales y se cree que pueden predecir la biodiversidad de manera beneficiosa (Barsoum et al., 2019; Bruce, 2013; Smith et al., 2008). Sin embargo, no está claro si estos indicadores predigan de manera adecuada, rápida y económica la verdadera riqueza de especies (Ji et al., 2013; Yu et al., 2012). Además, las evaluaciones tradicionales de biodiversidad requieren conocimientos taxonómicos especializados, lo que aumenta los costos y el tiempo necesario para procesar las muestras (Yu et al., 2012). En este contexto, se continúa desarrollando

esfuerzos para diseñar indicadores y estrategias eficiente de monitoreo de la biodiversidad que sean prácticos, efectivos, rápidos, rigurosos, confiables y aplicables a gran escala y de forma consistente (Ji et al., 2013; Yu et al., 2012). Uno de estos nuevos métodos de medición directa y rápida de la biodiversidad se denomina ADN metabarcoding de comunidades enteras (wcMBC, por sus siglas en inglés).

1.3. ADN metabarcoding en bosques mexicanos y monitoreo de artrópodos

El wcMBC es una técnica que permite la identificación del ADN de múltiples especies eucariotas en un ecosistema a partir de una sola muestra global compuesta por cientos o miles de individuos y especies (Ji et al., 2013; Taberlet et al., 2012; Yu et al., 2012). Esta técnica se basa en la capacidad de alto rendimiento de las plataformas de secuenciación de próxima generación (NGS, por sus siglas en inglés) para caracterizar muestras de ADN de comunidades mixtas, lo que permite identificar y verificar la presencia o ausencia de múltiples especies para evaluar la diversidad de forma rápida (Ji et al., 2013; Taberlet et al., 2012; Yu et al., 2012). Además, permite agrupar las lecturas de ADN en unidades taxonómicas operativas (OTU) para aproximar los perfiles de la comunidad a nivel de especie (Andújar et al., 2017; 2021; Arribas et al., 2020; Ji et al., 2013). Varios estudios han validado la utilización de este método en la evaluación de la biodiversidad de macroinvertebrados y artrópodos (Hajibabaei et al., 2012; Ji et al., 2013; Meyer et al., 2020; Yu et al., 2012; Zimmermann et al., 2015). Aunque la comparación entre los enfoques morfológicos y el wcMBC muestra que son complementarios, a menudo la evaluación morfológica es útil para la descripción de especies y para generar una biblioteca de referencia, mientras que los estudios de wcMBC son útiles para una evaluación rápida de la

biodiversidad (Ji et al., 2013; Zimmermann et al., 2015; Elbrecht et al., 2017; Meyer et al., 2020).

En el caso de los bosques mexicanos, en la presente tesis se aplicó el método de wcMBC a los artrópodos con propósito de monitorear el estado de conservación de los bosques y evaluar si las plantaciones y el manejo forestal realmente pueden favorecer a la biodiversidad nativa. Los artrópodos son un grupo taxonómico adecuado para este propósito debido a las siguientes razones: (i) su diversidad es un sólido indicador estructural de los ecosistemas (Obrist y Duelli, 2010). (ii) Los artrópodos son componentes significativos de los ecosistemas de montaña en México, dada su riqueza, abundancia y diversidad de nichos, además de presentar especies con historias biogeográficas neárticas y neotropicales (Morrone y Márquez, 2001). (iii) La composición de las comunidades de artrópodos se correlaciona con las comunidades de plantas (Basset et al., 2012; Lynggaard et al., 2020; Zhang et al., 2016). (iv) Los artrópodos presentan diferentes limitaciones en la dispersión, tanto alados como no alados, con efectos diferentes en la diferenciación entre localidades, como se espera según la teoría de la biodiversidad neutral (Baselga et al., 2013; 2015). (v) A pesar de la extinción de muchas de especies de megafauna, como los mamíferos (Ceballos et al., 2009), la diversidad de artrópodos podría seguir siendo alta y, por lo tanto, amerita ser conservada.

1.4. Rol de la limitación dispersiva y manejo forestal en la biodiversidad de artrópodos en bosques montañosos

La variación y similitud de las comunidades están influenciadas por procesos como la especiación, la dispersión, la extinción y las características del paisaje. En el caso de los

artrópodos, factores como la altitud, el aislamiento de la montaña, el estado de conservación y la edad del bosque, la complejidad del dosel, el sotobosque, el tipo de uso del suelo y la distancia espacial dentro del bosque contribuyen a moldear su composición. Por lo tanto, la composición de especies en montañas cercanas puede variar debido a la falta de continuidad climática y topográfica entre ellas, lo que resulta en aislamiento y potencial especiación (Graham y Fine, 2008; Kozak y Wiens, 2006; Uscanga et al., 2021; Wiens, 2004). El topoclima influye en la estructura de la comunidad dentro del hábitat (Noguerales et al., 2021), pero también lo hace el estado de conservación del bosque. Los bosques conservados presentan una mayor riqueza de especies en comparación con los bosques manejados (Lynggaard et al., 2020; Paillet et al., 2010; Wang et al., 2019; Zhang et al., 2016), ya que ofrecen un mayor número de nichos ecológicos (Graham y Fine, 2008). Los rodales bajo manejo forestal muestran composiciones de artrópodos diferentes a lo largo del tiempo (Barsoum et al., 2019).

Existen evidencias que sugiere que la biodiversidad de artrópodos es mayor en bosques nativos y plantaciones mixtas en comparación con los monocultivos (Wang et al., 2019). Se ha planteado la hipótesis de que la densa cobertura del dosel en los bosques permite una mayor diversidad de artrópodos (Lange et al., 2014; Zou et al., 2015). Asimismo, se ha observado que el sotobosque alberga una mayor diversidad de insectos en comparación con áreas donde la vegetación ha sido reducida (Rambo et al., 2014; Lange et al., 2014). Algunos grupos de artrópodos podrían estar correlacionados a bosques maduros, ya que se ha notado un aumento en la riqueza y diversidad de insectos en lugares con mayor complejidad estructural y heterogeneidad ambiental, como el incremento en la cobertura

del dosel, la presencia de hojarasca y madera muerta (Fuller et al., 2008). Diferentes tipos de uso del suelo también resultan en comunidades únicas de artrópodos, lo que se refleja en las marcadas diferencias en la estructura de las comunidades entre distintos tipos de bosques (Beng et al. 2016). Además, se ha observado que la composición de artrópodos experimenta cambios en ecosistemas sujetos a procesos de restauración (Edwards et al., 2014; Fernandes et al., 2019; Lynggaard et al., 2020). Adicionalmente, la limitación de la dispersión por procesos neutrales, que restringe la ubicación de un individuo basándose en la ubicación de su progenitor, ha sido destacada como un impulsor de la diversidad local (Hubbell, 2001; Rosindell et al., 2011). Esto significa que las limitaciones en la dispersión generan patrones espaciales a escalas más amplias, y pueden influir en las limitaciones de dispersión en los artrópodos del suelo (Arribas et al., 2020; McGill, 2010).

Para logra el éxito del biomonitoreo utilizando metabarcoding depende en gran medida de una sólida base de datos genéticos que facilite la comprensión de la biodiversidad de artrópodos a través de la FVTM y su historia filogeográfica. Esta cadena montañosa, conocida como el arco volcánico Neógeno más extenso de América del Norte, se extiende a lo largo de casi 1000 km de longitud, desde el Centro de México, el sur del Golfo de California hasta el Golfo de México, abarcando una superficie de aproximadamente 160,000 km2 de oeste a este a lo largo del territorio mexicano, comprendiendo una variedad de volcanes ubicadas entre las latitudes 18°30′ y 21°30′N (Ferrari et al., 2002; 2012). Con una formación que se remonta a unos 19 millones de años debido a la actividad tectónica asociada a la subducción de la Placa de Cocos bajo la Placa Norteamericana, lo cual ha generado una intensa actividad volcánica a lo largo de millones

de años (Ferrari et al., 2002; 2012; Gómez-Tuena et al., 2007). Entre los volcanes más destacados de esta región se encuentran el Popocatépetl, el Iztaccíhuatl, el Pico de Orizaba y el Nevado de Toluca, entre otros, con gradientes altitudinales que superan los 4,000 metros sobre el nivel del mar (Ferrari et al., 2012; Mastretta-Yanes et al., 2015). La FVTM no solo tiene gran importancia geológica, también en términos de biodiversidad, ya que ha sido moldeada por una intensa actividad volcánica, dando lugar a una variedad de paisajes y ecosistemas que contribuyen a la generación de endemismos y la complejidad en la distribución de especies. Por esta razón, se la considera una región biogeográfica relevante, vinculada a la zona de transición mexicana y reconocida por su papel en la diversificación, el endemismo y las transiciones biogeográficas (Morrone y Márquez, 2001; Rzedowski, 2006; Mastretta-Yanes et al., 2015; Uscanga et al., 2021). Este contexto geológico y ecológico permitirá comprender los procesos ecológicos e históricos que generan los hotspots de biodiversidad. Sin embargo, antes de determinar si el manejo forestal y las plantaciones tienen un impacto positivo en la biodiversidad local, es esencial considerar las limitaciones dispersivas de los artrópodos y los posibles cambios en el ensamblaje de las comunidades que podrían ser causados por procesos neutrales. En otras palabras, es necesario discernir si los cambios en la composición de las comunidades se deben a la distancia entre los sitios o a diferencias reales entre los tratamientos de manejo forestal, disturbio, reforestación y conservación del bosque. Aunque los artrópodos son una herramienta valiosa en el biomonitoreo, es necesario considerar las limitaciones dispersivas de cada grupo y su efecto en la diferenciación entre localidades, tal como lo predice la teoría de la biodiversidad neutral (Hubbel, 2001; Bell, 2001; Vellend, 2010; Rosindell et al., 2011).

El papel de la limitación dispersiva en la conformación de la biodiversidad ha sido objeto de estudio a nivel de comunidades (Baselga et al., 2013; 2015; Dong et al., 2016; Sarremejane et al., 2017; Arribas et al., 2020). Este enfoque se fundamenta en la teoría ecológica neutral, que propone que la estocasticidad en la dispersión es uno de los principales impulsores de la diversidad local y que debería actuar de manera análoga en todas las escalas taxonómicas (Hubbell, 2001; Bell, 2001; Vellend, 2010; Rosindel et al., 2011). Se ha utilizado este enfoque analítico para comparar las restricciones relativas de dispersión entre linajes de diferentes grupos taxonómicos (Gómez-Rodríguez et al., 2019, Múrria et al., 2017). Los estudios empíricos que han empleado este enfoque hasta la fecha (por ejemplo, Baselga et al., 2013; 2015; Gómez-Rodríguez & Baselga, 2018; Arribas et al., 2020) han encontrado respaldo para la generalidad de los procesos neutrales, identificando las restricciones en la dispersión como un factor predominante que impulsa la variación espacial en la estructura de la comunidad a múltiples niveles jerárquicos (Baselga et al., 2013; 2015). No obstante, para validar la generalidad de estos patrones observados, se requiere la recopilación de más datos empíricos, que incluyan comunidades enteras (en lugar de unos pocos linajes seleccionados), muestreadas a través de diferentes escalas geográficas, taxonómicas y regiones del mundo (Baselga et al., 2013; Arribas et al., 2020).

Las montañas son particularmente importantes para evaluar los patrones evolutivos, ya que las restricciones en la dispersión dentro de una única montaña pueden ser determinante en la generación de endemismos (Bray y Bocak, 2016). No obstante, se carece aún de datos empíricos que respalden esta hipótesis. Los pasos de montaña en zonas tropicales actúan como barreras de dispersión más efectivas en comparación con los pasos

ubicados en zonas templadas, debido a las diferencias de temperatura, lo que implica mayores costos para la dispersión de organismos tropicales (Janzen, 1967; Polato et al., 2018; Sheldon et al., 2018). Las denominadas islas en el cielo (sky-island) tropicales constituyen entornos propicios para investigar las limitaciones en la dispersión debido a las características particulares de su paisaje (Mastretta-Yanes et al., 2015; Rahbek et al., 2019). Estos hábitats, situados en sistemas montañosos a altitudes elevadas y geográficamente fragmentados, donde las tierras bajas actúan como barreras para la dispersión de organismos que residen en altitudes superiores. Esta dinámica favorece la divergencia de las poblaciones aisladas y confiere a las islas en el cielo un estatus reconocido por su singular diversidad biológica. Asimismo, estos ambientes pueden proporcionar información valiosa sobre la evolución y adaptación de las especies frente a los cambios climáticos a lo largo del tiempo (McCormack et al., 2009; Mastretta-Yanes et al., 2015; Wiens et al., 2019). Sin embargo, hasta el momento, no se han llevado a cabo estudios empíricos para determinar si la restricción en la dispersión dentro de una sola montaña tropical efectivamente constituye un factor determinante en la generación de endemismos.

La evaluación del estado de un bosque requiere contar con información confiable sobre su situación y condición (Chapela, 2012). Esto se logra a través de monitoreos que puedan informar sobre las tendencias de cambio de la flora y fauna nativa (Gollan et al., 2013). Los artrópodos son muy sensibles a los cambios en el uso del suelo y, por ende, pueden ser excelentes indicadores del efecto de las actividades humanas sobre la biodiversidad nativa. Sin embargo, debido a la complejidad biogeográfica de las islas en el cielo, la presencia de distintos tipos de vegetación y el conocimiento limitado de la

diversidad de artrópodos locales en los bosques mexicanos es necesario realizar un análisis específico que evalúe diferentes tipos de muestreo, variables topográficas y la historia filogeográfica de los taxa. En este trabajo, primero realizamos una revisión de las aplicaciones de wcMBC, con el propósito de proporcionar una guía para otros proyectos que busquen incorporar estas herramientas de nueva generación al biomonitoreo de bosques. Posteriormente, analizamos el papel de los procesos neutrales (dispersión) en la estructura de las comunidades de artrópodos a escalas geográficas finas, utilizando un muestreo de bosques de *Abies religiosa* bajo conservación dentro del Nevado de Toluca. Nuestro análisis de paisaje incluyó variables de diferentes tipos de vegetación y características topográficas, como elevación y pendiente. Finalmente, evaluamos los diferentes tratamientos de manejo forestal, disturbio y conservación dentro del bosque de *A. religiosa*, así como la reforestación y conservación en el bosque de *Pinus hartwegii*, con el objetivo de determinar si el manejo forestal y la reforestación realmente contribuyen a la conservación de la biodiversidad de artrópodos en la isla de cielo del Nevado de Toluca.

OBJETIVO

Detectar a una escala fina si existe limitación dispersiva debido a variaciones espaciales y topográficas, influyendo en la composición de la comunidad de artrópodos en los ecosistemas de la FVTM, mediante el método de *metabarcoding* y muestreos que consideraron condiciones ambientales, topográficas, diferencias en el manejo forestal y el estado de conservación. Se hizo énfasis en la comparación de la composición de la comunidad de artrópodos en bosques dedicados a la conservación, bosques bajo manejo y plantaciones forestales.

CAPÍTULO I

METABARCODING DE COMUNIDADES DE ARTRÓPODOS COMO HERRAMIENTA PARA BIOMONITOREO

Fotografía de Jorge Jiménez: Trampa de caída.
Fotografía de Nancy Gálvez-Reyes: Sopa de la biodiversidad en bosque bajo conservación.

PREFACIO

La biodiversidad, piedra angular de la salud y la resiliencia de la Tierra, enfrenta desafíos sin precedentes debido a la destrucción de hábitat, introducción de especies invasoras y cambio climático que conducen a cambios ambientales inducidos por la humanidad (Dobson et al., 2021). Preservar y monitorear la biodiversidad es crucial para abordar los problemas de conservación global y mantener el delicado equilibrio de nuestro planeta (Mace et al., 2014). El ADN metabarcoding, una técnica molecular de vanguardia se ha convertido en una poderosa herramienta para estudiar comunidades mixtas y evaluar la biodiversidad. Esta revisión profundiza en la aplicación del ADN metabarcoding a comunidades de artrópodos, arrojando luz sobre su potencial para monitorear la diversidad biológica.

En este capítulo, presentamos una revisión centrada en proporcionar fundamentos para el uso del ADN wcMBC como herramienta de biomonitoreo para evaluar la biodiversidad en comunidades de artrópodos. La atención se centra en la detección de secuencias de ADN, proporcionando una descripción general completa de las consideraciones metodológicas, desde el diseño de muestras hasta la bioinformática. Destacar las variantes de secuencia de amplicones (ASV) y su papel en la captura de la variación a nivel de haplotipo añade profundidad a la comprensión desde la diversidad intraespecífica (haplotipos) hasta rangos taxonómicos más altos. El capítulo también analiza los desafíos de los métodos tradicionales de identificación taxonómica, el potencial de los métodos moleculares para complementar estos métodos y la necesidad de adoptar

enfoques integradores de monitoreo de biodiversidad. Estos enfoques deben medir la biodiversidad repetidamente en grandes escalas espaciales y en diferentes niveles tróficos (multi-taxones), utilizando un diseño experimental de muestreo consistente, metodología estandarizada metodológica y estandarización de ADN wcMBC para obtener datos de calidad y robustos un biomonitoreo.

Los resultados de la presente revisión evidencian que el éxito de la herramienta de ADN wcMBC están vinculados al método de muestreo y las buenas prácticas de laboratorio. Esto incluye la limpieza de muestras a granel y la clasificación de tamaño de los artrópodos, las extracciones eficientes de ADN y la cuantificación confiable durante la construcción de la biblioteca son pasos esenciales. La elección de marcadores genéticos y cebadores para la identificación taxonómica precisa, la utilización de datos de referencias, junto con un filtrado de secuencias eficaz, garantiza la fiabilidad de los resultados. El capítulo presenta el concepto innovador de incorporar datos de haplotipos en el monitoreo de la biodiversidad. Al revelar la variación intraespecífica, este enfoque complementa la diversidad a nivel de especie, mejorando los análisis convencionales. Estos aspectos deben ser considerados cuidadosamente para evitar subestimar la diversidad. La integración del ADN metabarcoding en la investigación de la biodiversidad a través de un marco conceptual, las hipótesis y pregunta de estudio presentadas aquí puede informar estimaciones para comunidades enteras. La información generada puede guiar las estrategias de conservación y gestión en el manejo forestal, proporcionando herramientas valiosas para abordar los desafíos ecológicos contemporáneos.

Artículo de revisión: Towards DNA metabarcoding-based haplotype for monitoring terrestrial arthropod communities. ***(Enviado)***

Revista: ***Insect Conservation and Diversity***

Manuscript running title: DNA metabarcoding for conservation monitoring.

Nancy Gálvez-Reyes*[1,2], Daniel Piñero[1] & Alicia Mastretta-Yanes*[3,4].

[1] Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad Universitaria, Coyoacán, 04510 CDMX, México.

[2] Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México, Ciudad Universitaria, CDMX, México.

[3] Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), Liga Periférico Insurgentes Sur 4903, Col. Parques del Pedregal, Tlalpan, CDMX, México.

[4] Consejo Nacional de Ciencia y Tecnología, Benito Juárez (CONACYT), CDMX, México, Avenida Insurgentes Sur 1582, Crédito Constructor, Benito Juárez, CDMX, México.

*nancygalvez@ecologia.unam.mx orcid: 0000-0003-2712-2377

pinero@unam.mx orcid: 0000-0002-2509-2445

*amastretta@conabio.gob.mx orcid: 0000-0003-2951-6353

***corresponding author**

Abstract

1. DNA metabarcoding has become a necessary and powerful tool to study mixed communities from bulk samples by detecting short DNA sequences.

2. This review integrates the use of biological communities of arthropods as ecological indicators using DNA metabarcoding. It also explores methodological considerations ranging from sampling design to bioinformatics and highlighting Amplicon Sequence Variants-based pipelines to allow for the incorporation of haplotype-level variation.

3. To ensure the success of DNA metabarcoding, it is crucial to follow specific sampling methods and good laboratory practices, involving the proper cleaning of bulk-sample, appropriate selection of arthropod sizes, efficient DNA extractions, reliable quantification during library construction, appropriate choice of genetic markers and primers for accurate taxonomic identification, and to perform and effective sequence filtering.

4. Incorporating haplotype data into monitoring unveils intraspecific variation that could complement species-level diversity, thus enhancing conventional analyses. Furthermore, the analyses should take into account either abundance- or incidence-based approaches, which condition the choice of statistical analyses to be employed.

5. Taking into account all these aspects, it is ensured that the hypotheses and research questions can inform biodiversity estimates for entire communities and guide ecosystem conservation and management strategies.

Keywords: Monitoring, Structural Indicator, Insect Communities, Amplicon Sequence Variants, DNA metabarcoding

Introduction

Biodiversity is crucial for Earth's health and well-being, benefiting both nature and humanity. Its preservation is a global responsibility integral to addressing environmental and conservation challenges (Wilson *et al.*, 2016; Hoban *et al.*, 2022). The commitment to halting the loss of biodiversity has emerged due to the transgression of the planetary boundaries, such as climate change, land-system change, and rate of biodiversity loss, among others (Rockström et al., 2009; Mace *et al.*, 2014). This situation has put human well-being and the resilience of the Earth's entire system at risk (Challenger *et al.*, 2009; Mace *et al.*, 2014; Dobson *et al.*, 2021). To secure our health and well-being and that of the planet, it is necessary to address the main drivers of nature loss considering the role of phylogenetic diversity, functional diversity, and biome integrity in determining a safe operating space (Mace *et al.*, 2014). To ensure that conservation actions are successful, we need to monitor biodiversity through the collection of extensive data and repeated analyses at various time points. In this context, comparing biological communities under diverse historical conditions, human management and conservation statuses is crucial, because it allows to contrast different sustainability strategies.

Biodiversity, encompassing diversity at various levels from genes to ecosystems, plays a pivotal role in ecological and evolutionary processes (Bailey *et al.*, 2009). Understanding the spatial and temporal distribution of ecological communities is crucial for assessing eco-evolutionary dynamics, identifying biodiversity hotspots, and guiding conservation efforts (Vellend, 2010; Baselga *et al.*, 2013; Gaston and Spicer, 2013; Baselga, Gómez-Rodríguez and Vogler, 2015; Weber *et al.*, 2017; Gómez-Rodríguez *et al.*, 2019; McGill *et al.*, 2019). Moreover, these distribution shifts offer insights into species'

evolutionary history and interactions, aiding in the study of speciation, adaptive radiation, and coevolution (Parmesan and Yohe, 2003; McGill *et al.*, 2007). Thus, while biodiversity monitoring is conducted for conservation or management purposes, it is essential to acknowledge that biodiversity distribution is influenced by ecological and evolutionary processes, independently of human intervention.

The integrated assessment of macroecological patterns at multiple levels of genotypes, genealogies, and species, known as multihierarchical macroecology (Baselga *et al.*, 2013; Baselga, Gómez-Rodríguez and Vogler, 2015), explores the factors influencing biodiversity patterns within an eco-evolutionary context. This approach investigates the spatial patterns of variation of entire assemblages at multiple hierarchical levels between haplotypes and species, exploring the patterns emerging across levels to discern between neutral and non-neutral controls of assemblage variation. It is firmly based on neutral ecological theory, which proposes that dispersal stochasticity is one of the main drivers of local diversity and that it should operate consistently across taxonomic scales (Bell, 2001; Hubbell, 2001; Vellend, 2010; Rosindell, Hubbell and Etienne, 2011). This theory extends to the neutral theory of molecular evolution, which suggests that genetic drift accumulates selectively neutral genetic changes over time (Kimura, 1968, 1983). By integrating these neutral theories and considering migration, speciation, and extinction, researchers have compared dispersal restrictions among different taxonomic lineages (Vellend and Geber, 2005; Papadopoulou *et al.*, 2011; Múrria *et al.*, 2017; Gómez-Rodríguez *et al.*, 2019). Empirical studies (Baselga *et al.*, 2013; Baselga, Gómez-Rodríguez and Vogler, 2015; Gómez-Rodríguez and Baselga, 2018; Arribas *et al.*, 2020; Gálvez-Reyes *et al.*, 2021) generally support the prevalence of neutral processes, with dispersal limitations being a significant driver of spatial variation in community structure across multiple hierarchical

levels. However, the generality of these observed patterns requires monitoring empirical data, encompassing entire communities, sampled contrasting geographic scales, and diverse taxa, and world regions (Baselga *et al.*, 2013; Arribas *et al.*, 2020).

Analyzing the spatial and temporal distribution of whole communities is essential to understand the biodiversity patterns drawn by eco-evolutionary processes and the effects of human-induced changes on ecosystems (Vellend, 2010). Monitoring community changes can help to detect disturbances caused by activities such as urbanization, deforestation, pollution, and climate change (Pimm *et al.*, 1995; Sala *et al.*, 2000). However, for this to be possible it is necessary to have integrative approaches that enable robust and reproducible assessment of biodiversity at different levels (genes, species, communities), repeatedly, at large spatial scales, and within a short time frame. Incorporating empirical data to meet these requirements is challenging due to the significant sampling effort and the temporal bottleneck for taxonomic identification (Favreau *et al.*, 2006; Yu *et al.*, 2012). Arthropods are particularly valuable for the task, as their sensitivity to environmental change, small size, large abundance, and presence in a wide array of niches across ecosystems means that entire communities of them can be sampled in short time. This is also the case of insects (Chua *et al.*, 2023), which also represent one of the most vulnerable groups of species to explore the loss of biodiversity during the Anthropocene and in the face of climate change. As for the challenge of taxonomic identification, this can be done using morphological identification, but complementing it with molecular methods can accelerate it (McCravy, 2018).

DNA barcoding allows to identify specimens using one or several short DNA fragments ("barcodes") (Hebert *et al.*, 2003). In turn, DNA metabarcoding involves identifying multiple specimens from a mixed sample, based on high-throughput sequencing of DNA barcodes (Taberlet *et al.*, 2012; Andújar *et al.*, 2017). Morphological assessments

are useful to describe species and generate a reference library (sequences assigned to specimens with known taxonomic identifications), whereas DNA metabarcoding studies are preferable for rapid biodiversity assessments (Ji *et al.*, 2013; Zimmermann *et al.*, 2015; Elbrecht *et al.*, 2017; Meyer *et al.*, 2021). In particular, advances in high-throughput molecular methods (Bourlat *et al.*, 2013) and bioinformatics have led to an explosion of studies generating large-scale ecological data that can be used for monitoring.

Metabarcoding involves sequencing one or several variable barcode regions and simultaneously profiling multiple taxa within a single sample (Taberlet *et al.*, 2018; Ruppert, Kline and Rahman, 2019; Miyata *et al.*, 2022). Metabarcoding can be based on DNA, RNA or environmental DNA (eDNA). After sequencing, bioinformatic tools and reference databases are subsequently employed to assign taxonomic identities to the sequenced fragments (Taberlet *et al.*, 2012; Yu *et al.*, 2012; Ji *et al.*, 2013). These methods were initially applied in the context of microbial community analysis of bulk samples using 16S sequencing (Schmidt, DeLong and Pace, 1991; Paliy and Shankar, 2016), which allowed for the assessment of the composition and dynamic changes of complex communities (Paliy and Shankar, 2016). Eukaryotic communities, such as arthropods or mollusks have been also studied from bulk samples, containing hundreds of individuals, which were sequenced focusing on fragments of the mitochondrial DNA Cytochrome c Oxidase subunit I (COI) gene (Yu *et al.*, 2012). This is called whole-community DNA-RNA metabarcoding (wcMBC) (Andújar, Arribas, Yu, *et al.*, 2018).

Although the method for taxonomic identification of multiple species extracted from a mixed sample can be from community DNA or eDNA, there are differences between the two (Deiner *et al.*, 2017). Particularly, in community DNA metabarcoding, the targeted groups are most often collected in bulk (e.g., arthropods from soil, malaise, pitfall or net

traps), and in eDNA, DNA it is isolated directly from an environmental material (e.g., soil or water) without prior segregation of individuals (Deiner *et al.*, 2017). In this study, we focus on community DNA metabarcoding. A recent refinement includes the use of community DNA metabarcoding which we will discuss further in the following sections.

DNA-RNA metabarcoding has revolutionized the understanding and monitoring of animal communities, particularly arthropods (Yu *et al.*, 2012; Ji *et al.*, 2013; Andújar *et al.*, 2015; Arribas *et al.*, 2016). Data can be obtained as abundance data for each species or as a presence-absence data for those species. Haplotype level data, which is often ignored, can also be extracted from DNA metabarcoding data in the form of hypothetically error-free sequences, called Exact Sequence Variants (ESVs) or Amplicon Sequence Variants (ASVs, sensu Callahan, McMurdie and Holmes, 2017), as surrogates for haplotypes (Callahan *et al.*, 2016; Edgar, 2016; Elbrecht *et al.*, 2019; Andújar *et al.*, 2021). ASV is a single inferred DNA sequence obtained from marker gene analysis that is generated by filtering out PCR and sequencing errors. They allow for precise identification of sequence variations due to individual nucleotide changes (Callahan *et al.*, 2017). Analyzing of species intraspecific genetic variability remains challenging from DNA metabarcoding data, but recent improvements to denoise barcoding datasets (e.g., Edgar, 2016; Callahan *et al.*, 2016) and evaluating the prevalence of sequencing errors and coamplified pseudogenes (Andújar *et al.*, 2021) raise the possibility of performing read-based haplotype-level analyzes with metabarcoding data from the mitochondrial COI community, which represents a radical change for the study of animal diversity patterns through genetic analysis of the entire community at different scales and biomes (Andújar *et al.*, 2017; Arribas *et al.*, 2020). The method is still in its infancy and few studies have evaluated the genetic component using haplotypes (Elbrecht *et al.*, 2018; Tsuji *et al.*, 2020; Turon *et al.*, 2020; Zizka, Weiss and

Leese, 2020; Arribas *et al.*, 2020; Gálvez-Reyes *et al.*, 2021; Noguerales *et al.*, 2021). Nonetheless, there are already a plethora of methodological options and applications that need to be considered to avoid generating faulty data (Liu *et al.*, 2020; Bohmann *et al.*, 2021; Creedy *et al.*, 2022).

Therefore, the focus of this review is to provide guidelines for using DNA metabarcoding to assess biodiversity status from haplotypes to higher taxonomic ranks. We focus on arthropods because they constitute the most diverse and abundant biological group, making them ideal for monitoring at both fine and large spatial scales (Gaspar, Gaston and Borges, 2010; Basset *et al.*, 2012). Here, we first summarize key concepts and gaps in biodiversity monitoring. Second, we justify why arthropod communities are meaningful in answering the relevant questions for monitoring using molecular based methods. Third, we describe the current *status quo* and main methodological considerations for bulk DNA metabarcoding of arthropod communities. Fourth, we conclude by reviewing current advances, applications, and future possibilities of DNA metabarcoding particularly using ASV-based approaches and other individuals-free sampling methods. Finally, we included supplemental material about Ecological statistics based on ASV-based data and DNA metabarcoding.

1. Key concepts of molecular-based biodiversity monitoring

Monitoring is a systematic approach using living organisms to assess ecosystem conditions and environmental changes (Li, Zheng and Liu, 2010). This method involves observing the effects of external factors on ecosystems over a long time period to identify patterns in biotic attributes and their causes (Markert *et al.*, 1999; Buss *et al.*, 2014). Successful monitoring requires measuring taxonomic diversity among sensitive to environmental disturbances, and

requires specialized taxonomic knowledge (Hajibabaei *et al.*, 2016). Biodiversity surveys track changes in indicator species' distribution and presence, studying population distributions over time and space (Niemelä, 2000; Baird and Hajibabaei, 2012; Calvignac-Spencer *et al.*, 2013).

On the other hand, monitoring global ecosystems for biodiversity is challenging with traditional methods for recording and identifying organisms (Baird and Hajibabaei, 2012). To address this, various ecological indicators have been developed as alternatives to quantify biodiversity (Barsoum *et al.*, 2019). These indicators include flag species, umbrella species, structural indicators, and composition indicators, which offer insights into habitat changes and progress (Noss, 1999). Additionally, subsets of the total community based on ecological guild (Roslin, 2001), and biological communities (Vellend, 2010; Stroud *et al.*, 2015) have been proposed as valuable indicators for assessing biodiversity.

A community consists of a collection of populations of species that coexist in an environment, interacting with each other usually in reciprocal ways and forming a hierarchical structure, trophic relationships, and functionality (Fauth *et al.*, 1996; Vellend, 2010; Stroud *et al.*, 2015). The community parameters are species richness, diversity, dominance, and specific interactions are emergent attributes used to measure species parameters in a given area (Morin, 1999; Chao and Jost, 2012). Biodiversity's significance lies in its impact on the community's composition, structure and functionality, reflecting the evolutionary history of a limited space (Whittaker, 1975). For instance, woodland arthropod species assemblages can be very due to environmental factors, such as forest stands composition (Barsoum *et al.*, 2019).

Comprehensive assessment of biodiversity involving multispecies associations, functional dynamics, and environmental changes are challenging due to difficulties in

identifying many species (Baird and Hajibabaei, 2012; Hajibabaei *et al.*, 2016) and the need high quality and high-resolution survey data (Butchart *et al.*, 2010; Bourlat *et al.*, 2013; Bush *et al.*, 2020).

Multispecies associations and functional dynamics in space and time, linked to environmental changes, have provided the most comprehensive assessment of biodiversity. However, these analyses are limited by the difficulty of identifying many species due to taxonomic identification challenges that may arise due to a variety of factors, including taxonomic ambiguity, morphological variation, cryptic species, and incomplete taxonomic knowledge (Baird and Hajibabaei, 2012; Hajibabaei *et al.*, 2016), additionally, the vast community diversity, lack of resources (e.g., funding to support taxonomists' expertise), and increasing costs in terms of time and money (Yu *et al.*, 2012; Creer *et al.*, 2016). Also, high-quality, high-resolution large-scale survey data is essential for accurately detecting changes in ecosystem health, these attributes data are collected under rigorous protocols with accurate, precise, and free from errors or bias at a fine level, involving measurements taken at closely spaced intervals in time and space (Butchart *et al.*, 2010; Bourlat *et al.*, 2013; Bush *et al.*, 2020).

Monitoring approaches have evolved to address these challenges, Biomonitoring-1.0 initially relied on diagnostic indicators and direct biotic measurements but faced limitations in identifying organisms (Baird and Hajibabaei, 2012). It operated locally over extended periods, generating low-taxonomic-resolution data and binary outcomes (e.g., affected/unaffected), relying on specialized taxonomists (Baird and Hajibabaei, 2012; Woodward, Gray and Baird, 2013). To enhance information gathering, biomonitoring-1.5 involves the generation of large network databases that incorporate species lists,

biogeography, geomatics, functional traits, and species interactions enabling a more comprehensive diagnosis beyond presence/absence (Woodward, Gray and Baird, 2013).

In recent years, biomonitoring-2.0, a recent development in ecological monitoring, utilizes DNA metabarcoding techniques to significantly enhance the scalability and species identification capabilities of conventional monitoring (Baird and Hajibabaei, 2012; Woodward, Gray and Baird, 2013). This method employs high-throughput DNA/RNA sequencing and molecular-based approaches to identify numerous eukaryotic species in an ecosystem from a single bulk-sample, offering a more comprehensive assessment of biodiversity of the entire community (Taberlet *et al.*, 2012; Yu *et al.*, 2012; Ji *et al.*, 2013). Additionally, it incorporates ecological trophic networks, genomics, and high-performance computing for 'big data' analysis in large-scale monitoring efforts (Woodward, Gray and Baird, 2013).

2. Unveiling the role of arthropod communities in monitoring

Arthropoda is highly regarded for their vast species diversity and abundance of taxa, with far more species than any other phylum (Ødegaard, 2000; Budd and Telford, 2009). This phylum comprises approximately 1,300,000 described species, divided into four subphylum groups: Chelicerata (115,992 species), Myriapoda (12,010 species), Hexapoda (1,080,760 species), and Crustacea (73,141 species) (Ødegaard, 2000; Budd and Telford, 2009; Zhang, 2013). Arthropods are valuable for monitoring, serving as indicators of biological community health and richness of other taxa due to their local abundance, diversity, and ease of bulk-sampling. They possess various dispersal abilities, body sizes, short generation times, high reproduction rates, colonization history, and variable distribution ranges (Gaspar, Gaston and Borges, 2010; Samways, McGeoch and New, 2010). Moreover, arthropods are ecologically

significant, contributing to ecosystem services and serving as conservation status indicators (Obrist and Duelli, 2010; Samways, McGeoch and New, 2010; Anderson *et al.*, 2011). Arthropods play important ecological roles due to their enormous abundance and functional importance (e.g., pollinators, predators and herbivore, detritivores, seed dispersers, ecosystem engineers, parasites, parasitoids, and bioindicators), exhibiting diverse responses to the environmental factor (Smith *et al.*, 2008; Spence *et al.*, 2008). For instance, Lepidoptera richness and abundance indicates young forests ecosystems (Skórka, Settele and Woyciechowski, 2007), while Carabids (Coleoptera: Carabidae) and Syrphids (Diptera: Syrphidae) reflect invertebrate diversity and respond differently to field layer cover (Humphrey *et al.*, 1999; Koivula, 2011; Pawson *et al.*, 2011). Araneae diversity increases in maturing forests (Smith *et al.*, 2008).

Furthermore, it is widely assumed that a physically complex habitat provides a greater number of ecological niches and, therefore, greater species richness (Stein *et al.*, 2013). The co-evolutionary relationship between plants and insects has led to a great diversity of plant and insect species that continues to exist today (Lindenmayer, Franklin and Fischer, 2006; Wheat *et al.*, 2007). As a result, there is a positive correlation between arthropod diversity and plant diversity, supporting global estimates of arthropod species based on plant diversity (Stork *et al.*, 2015; Zhang *et al.*, 2016). This relationship can be leveraged through surrogate taxa, representing biodiversity patterns across different environments (Barsoum *et al.* 2019). Arthropod richness and community composition serve as indicators of forest degradation and recovery (Ji *et al.*, 2013). Plant diversity influences insect diversity through herbivory, pollination, and habitat (Haddad *et al.*, 2009; Zhang *et al.*, 2016), enabling the use of arthropods as measures of community diversity in monitoring, as substitutes for plant communities (and *vice versa*), for improved biodiversity conservation (Castagneyrol and

Jactel, 2012). Finally, arthropods are advantageous for ASV-based studies because their small size and local abundance allow for meaningful haplotype-level diversity measurement, as shown in studies by Arribas *et al.* (2020), Gálvez-Reyes *et al.* (2021), and Noguerales *et al.* (2021).

3. Methodological considerations of bulk DNA metabarcoding of arthropod communities for ASV-based analyses

DNA metabarcoding identifies multiple species in mixed samples as community DNA or environmental DNA (eDNA), including whole arthropod communities, through bulk (group of many organisms) or environmental samples (e.g., water or soil) (Yu *et al.*, 2012; Ji *et al.*, 2013; Elbrecht *et al.*, 2019). Taxonomic identification can be cross-validated with whole organism samples, enabling a comprehensive understanding of the detected biodiversity's temporal and spatial scale (Creer *et al.*, 2016; Deiner *et al.*, 2017). In this discussion, we will primarily focus on bulk samples, as they are the more commonly collected target groups in bulk. For information on eDNA sampling, refer to Deiner *et al.* (2017).

The main steps in bulk DNA metabarcoding of arthropod communities for ASV-based analyses involve: (1) Robust sampling, following a robust experimental design with a suitable number of locations and replicates. (2) Cleaning of bulk-samples using the FFS (Flotation-Filtration-Stereoscope) protocol separate arthropods from organic matter, and (3) Genomic DNA extraction from the bulk samples to obtain high-quality and -quantity DNA from small and large organisms, (4) library preparation for molecular marker identification with carefully selecting primer and PCR condition selection to avoid biases. The libraries are then purified, quantified, and pooled, (5) Sequencing on an appropriate platform with subsequent bioinformatic processing for clean, demultiplex, and quality-control of reads. (6)

Species (OTUs) identification via haplotype assignment using a reference database. Based on this, (7) creation of a community matrix from haplotypes to higher taxonomic levels, and (8) conducting ecological, diversity, and phylogenetic analyses (Figure 1).

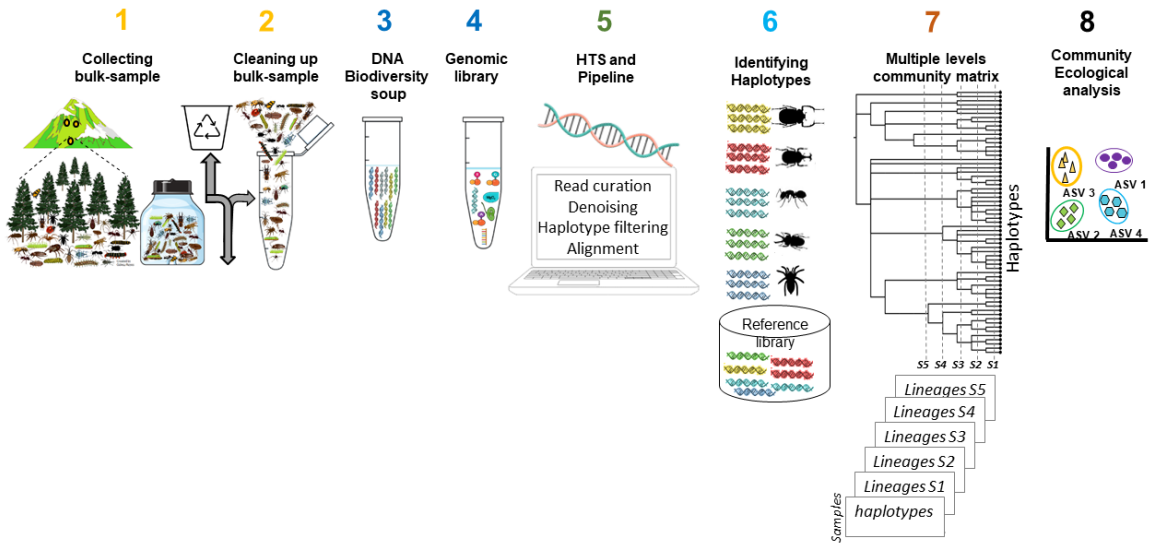

Figure 1. Main steps of bulk DNA metabarcoding of arthropod communities for ASV-based analyses. See main text for details.

DNA metabarcoding of arthropod communities involves multiple steps with crucial methodological considerations (Figure 1). Additionally, it is important to identify the detection of invertebrates, affecting both their absolute and relative abundance and addressing potential sources of error during the wet-lab stage is essential to ensure accurate under- or overestimating community diversity (Table 1). These sources of error can significantly impact the detection of invertebrates, affecting both their absolute and relative abundance (Elbrecht *et al.*, 2017). For instance, studies have shown that issues like tag-jumps

and false attribution of amplicon sequences to samples can lead to false positives, inflating diversity estimates, and ultimately resulting in species underestimation (Schnell, Bohmann and Gilbert, 2015).

Table 1. Common sources of wcMBC errors.

Source of error	Reasons	Recommendations	References
Abundance and biomass	* Specimens with low biomass may have low amplification. * Variation of composition metrics based on abundance could be due to size or taxon biases in the recovery.	(i) Samples should be divided according to body size, and (ii) an amount of DNA can be adjusted according to the sample size before amplification and sequencing, (iii) the former may not be necessary with sufficient sequencing depth.	(Elbrecht and Leese, 2015; Piñol *et al.*, 2015; Elbrecht, Peinert and Leese, 2017; Creedy *et al.*, 2019)
Primers affinity and amplification	* Due to the mismatch between the universal primer and the primer site of each species the amplification may cause a bias in the final proportion of taxa detected. * The concentration of blocking oligonucleotides could occur if species have sequences more similar to target species. * The primers could fail during the amplification so some taxa could remain undetected. * Variations in PCR conditions can affect the diversity detected.	(i) Careful primer evaluation and use of specific region/ecosystem primers), (ii) using two pairs of primers with different sizes, allows the recovery different specimens, (iii) an increase in the alignment temperature increases the specificity of the primer, however, there is a trade-off in the recovery of species, (iv) in silico proof of concept and mock community testing of newly developed primers or primers from the literature against the specific taxa of interest, (v) use the same primers, hybridization temperature and PCR cycle number across different samples and runs.	(Piñol *et al.*, 2015; Elbrecht and Leese, 2017a; Krehenwinkel, Kennedy, *et al.*, 2018)
DNA quality	* Degraded DNA produces short fragments leading to amplification and sequencing biases.	(i) Keep samples in good storage conditions and use ethanol-based preservation methods, (ii) separate target organisms from remains before extraction, (iii) compare results against mock communities containing various	(Stein *et al.*, 2013; Elbrecht and Leese, 2017a; Krehenwinkel *et al.*, 2017; Krehenwinkel,

	* PCR is negatively affected by the inhibitors, which may be present especially if DNA is extracted together with remains (leaves, gravel, mud, etc.)	known arthropods species.	Fong, *et al.*, 2018)
tag-jumping	* During the labeling of the library by short PCR cycles, sometimes sequences with different label combinations (artefacts) are generated ("tag-jumping"). This produces false labels that artificially inflate diversity.	(i) Using matching tags, (ii) using double dual tagging, (iii) performing PCR replicates of the same sample (i.e., three reactions of PCR per sample), (iv) handling labeled amplicons, incorporating negative controls into the process and sequence a subset, and (v) using unique tags per library.	(Schnell, Bohmann and Gilbert, 2015; Krehenwinkel, Kennedy, *et al.*, 2018)
label design or index design	There is a trade-off between label length and amplification efficiency.	(i) Bases of the labels must be balanced in all positions, (ii) large labels are not recommended, (iii) labels with different lengths can be used to increase the complexity at each base position in the sequence length.	(Krueger, Andrews and Osborne, 2011; Coissac, Riaz and Puillandre, 2012; Schnell, Bohmann and Gilbert, 2015)
cross-contamination	Cross-contamination can occur when handling samples, especially during cleaning and size selection.	(i) Sorting sample by specimen biomass only if specimens vary in several magnitudes, (ii) washing equipment between each sample with bleach, soapy water, distilled water, ethanol, and UV light or fire, eliminating RNases and DNases. (iii) if plastic bags or vials are not purchased sterile, they can be decontaminated by UV exposure (30 min), bleach washing, or autoclaving, (iv) include several negative controls and at least one or two positive controls, at the fieldwork and lab stages (v) if detected reads map to ASVs in the negative controls, you can exclude them from further analysis.	(Elbrecht *et al.*, 2017, 2019; Dickie *et al.*, 2018; Erdozain *et al.*, 2019; Liu *et al.*, 2020; Gálvez-Reyes *et al.*, 2021; van der Loos and Nijland, 2021; Holman, Chng and Rius, 2022)

3.1 The sampling design and methods (either from individuals or from the different parts of the environment) for monitoring studies is crucial, involving considerations such as the research question, sampling area representativeness, statistical replication, the model used,

and data interpretations. However, these studies often face limitations due to the effort required to consider multiple taxa, which can vary depending on the habitat and research method (Mueller and Geist, 2016). These limitations can create a bottleneck, as the effort, time, and cost-effectiveness of operations can vary depending on the habitat and research method used (Hajibabaei *et al.*, 2016). To ensure robustness, a well-designed study should include an adequate number of sites with proper statistical replication and collect metadata like climatic variables, soil pH, temperature, salinity, humidity, and organic matter content (Creer *et al.*, 2016). A well-designed study is essential to obtain data from diverse taxonomic groups that coexist in space and time (Schmeller *et al.*, 2015; Mueller and Geist, 2016). Nested stratified-random sampling design, which standardizes spatial and temporal sampling with statistical replications, can be an effective approach, but the sampling unit's extent should align with management site scales (Mueller and Geist, 2016). Therefore, it is beneficial to cover the area of interest with a greater number of spatial and temporal statistical replications in a stratified and randomized manner (Schmeller *et al.*, 2015).

The choice of sampling design in biodiversity studies is influenced by factors such as spatial scale, research objectives, and area size. For instance, in complex or disturbed habitats, a one-hectare site may pose logistical challenges (Emerson *et al.*, 2017). Smaller-scale sampling scenarios, like 0.25-hectare quadrants (50x50 m) distributed based on topography and logistical feasibility, have been employed (Zhang *et al.*, 2016; Emerson *et al.*, 2017). In areas with varying land use, such as forests and plantations, 1x1 m quadrants every 10 m with litter collection can be useful (Beng *et al.*, 2016). Additionally, collecting macroinvertebrates in small sampling points with an area of 0.45 m^2 has been implemented (Elbrecht, Peinert and Leese, 2017). For larger sampling scales, block designs of 100x100 m with stratified random sampling, considering different habitat types, have been employed

(Zhang *et al.*, 2016). In extensive landscapes, sites with quadrants of 500x500 m, divided into 16 area units of 100x100 m, have been utilized (Yang *et al.*, 2014). Representative sampling strategies often involve a standardized design covering 0.25 hectares (50x50 m) and employing various collection techniques like vacuum canisters, beating trays, pitfall traps, Winkler traps, and Malaise traps to enhance species richness, cryptic species, and ecological complexity (Emerson *et al.*, 2017).

Various sampling methods are used to ensure adequate coverage of arthropod diversity or targeted arthropods of interest (e.g., pollinators). Common methods used in DNA metabarcoding studies of arthropods include Malaise traps for flying insects (Yu *et al.*, 2012; Ji *et al.*, 2013; Gibson *et al.*, 2014; Yang *et al.*, 2014; Brandon-Mong *et al.*, 2015; Shokralla *et al.*, 2015), leaf litter sampling (Yang *et al.*, 2014; Beng *et al.*, 2016), Winkler traps for insects in leaf litter (Yang *et al.*, 2014), pitfall traps for ground-dwelling arthropods (Ji *et al.*, 2013; Pedro *et al.*, 2017; Gálvez-Reyes *et al.*, 2021), and light traps for nocturnal insects (Zhou *et al.*, 2013). Malaise traps are especially effective for flying insects such as Hymenoptera, Diptera, Lepidoptera, Neuroptera, Hemiptera and Coleoptera using an intercept mesh and an ethanol container. Winkler traps collect insects from leaf litter, while pitfall traps are suitable for ants, Coleoptera, and wingless micro-Hymenoptera from the soil surface. Specimens are preserved in 70% ethanol with glycerin to prevent ethanol evaporation and to attract insects. Light traps are suitable for night-active insects, while Malaise traps are used in DNA metabarcoding studies even under non-optimal storage conditions, such as two- and four-week degradation experiments (Krehenwinkel, Fong, *et al.*, 2018). Pitfall traps are cost-effective for terrestrial arthropods, suitable for extended deployment in various landscapes (Pedro *et al.*, 2017) and forests (Drummond *et al.*, 2015; Uscanga *et al.*, 2021). Malaise traps efficiently capture aerial arthropods, while pitfall traps

are effective for ground-dwelling species with limited mobility (de Kerdrel *et al.*, 2020). Although, the use of Malaise traps can affect the species richness and composition of bulk samples of terrestrial arthropods (Steinke *et al.*, 2021). These traps capture a diverse range of arthropods, and they are highly complementary, as they sample non-overlapping assemblages of arthropods. Combining Malaise and pitfall traps is common in DNA metabarcoding studies, as they capture distinct arthropod assemblages (Watts *et al.*, 2019). These methods are well-established for generating biodiversity data. For example, in alpine grasslands and forests at high altitudes, pitfall traps exhibited greater richness due to their longer field deployment duration (approximately 15 days) compared to other traps (Figure 4a; see Supplementary Material 1 for details).

3.2 Bulk-sample cleaning. The cleanup of the sample can be conducted using various methods, including decantation, individual extraction with surveillance forceps, sieving with a strainer, and filtering with filter paper. These methods optimize the timing and separation of arthropods from organic matter such as roots, litter, and soil debris. They also prevent DNA contamination and PCR inhibitors (Table 1). In our experience (Figure 2), the filtration method has shown the shortest processing time (2 hours per two samples) and highest processing quality. This method involves placing a 20×20 cm filter paper box inside a funnel in a 1-liter beaker, as described by Galvez-Reyes *et al.* (2021). The filtration method facilitates the flotation of small organisms in a liquid medium, which can then be decanted, leaving the arthropods on the filter paper. Next, water is added to the sediment and mixed for 15 seconds before being left to rest for 10 minutes to allow the supernatant to decant. Finally, the filter paper is examined under a stereoscope to capture the arthropods using a clock clamp (Figure 2). Large individuals are placed in Falcon tubes with 70% ethanol, while small

individuals are placed in 2.0 ml microtubes. An alternative method for cleaning microfauna samples from soil is the flotation-Berlese-flotation (FBF) method, as described by Arribas *et al.* (2016).

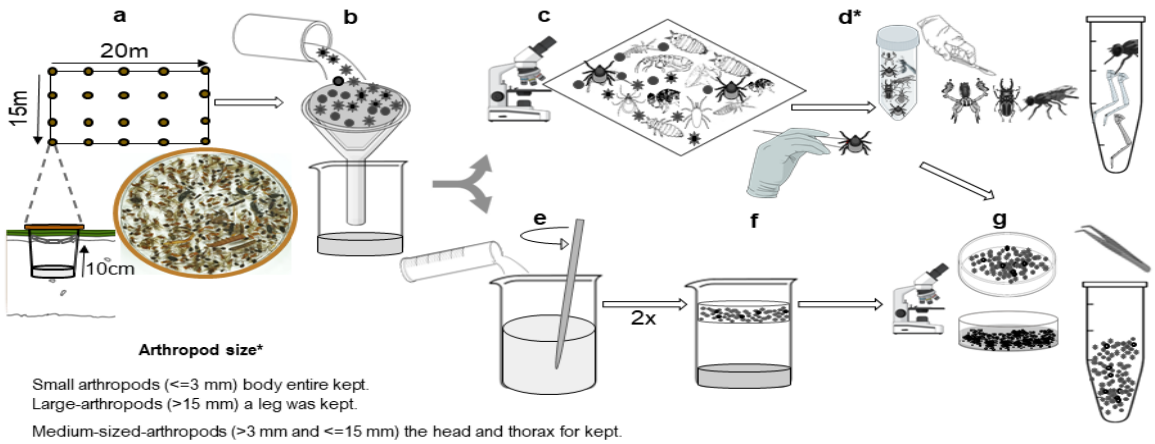

Figure 2. Sampling and Flotation-Filtration-Stereoscope (FFS) method for cleaning arthropod bulk-samples, modified from Gálvez-Reyes *et al.* (2021). Bulk-samples were collected from an *Abies religiosa* forest. The protocols involve several steps: (a) Setting up 20 pitfall-traps for each block site (15 x 20 m) with 200 ml (alcohol 70% + glycerin 15 ml solution). After collecting pitfall traps, bulk-samples were processed following an FFS protocol that allows for the 'clean' extraction of arthropods. The FFS protocol is based on arthropods flotation in water, filtration process and observation in the stereoscope. (b) Flotation and filtration steps to separate arthropods and organic matter. (c) Observation and capture of small arthropods (<=3 mm) with a small clamp. (d) Segmentation of samples by size, for large-arthropods (>15 mm) only a leg was kept, and the head and thorax for medium-sized-arthropods (>3 mm and <=15 mm). (e) Additional water flotation of the sediment portion is placed using 400 ml of water, mixing for 15 seconds, and allowing it to stand for 3 minutes. This step was done two times, depending on the sample's level of dirtiness. (g)

Decanting and observing in a stereoscope for capturing small specimens, and finally, preservation in ethanol, resulting in clean bulk-sample of specimens ready for DNA extraction.

Bulk samples may contain taxa with varying biomass, which can pose challenges for DNA metabarcoding analyses. Specimens with large body sizes can contribute to a qualitative and quantitative bias in the DNA sample, leading data set, while small samples may remain undetected (Elbrecht and Leese, 2015; Elbrecht *et al.*, 2017; Creedy *et al.*, 2019). The biomass-to-body size ratio can also have a negative impact on DNA metabarcoding approaches (Table 1). To ensure consistency in DNA extractions, it is important to separate samples by size so that the amount of tissue (biomass) used is similar. Macroinvertebrates can be classified into three size classes based on body size (<2.5 × 5, 5 × 10, and up to 10 × 20 mm). However, it has been shown that this approach can lead to bias in the abundance estimates in favor of specimens with higher biomass, while small and less abundant taxa may remain undetected. In fact, on average, about 30% fewer taxa are detected in unclassified samples (Elbrecht *et al.*, 2017). In our experience, we used three size categories for arthropods (Gálvez-Reyes *et al.*, 2021): entire body for small arthropods (*Drosophila melanogaster* <=3 mm), the head and thorax for medium-sized arthropods (adult *Apis mellifera* >3 mm and <=15 mm), and only one leg was kept for large arthropods (adult grasshopper >15 mm, as detailed in Figure 2 and Supplementary Materials 2. Recently, Lim (Lim *et al.*, 2022) sorted samples into four body size categories (0–2 mm, 2–4 mm, 4–7 mm, and >7 mm). While sorting by size can increase resolution and potential cost savings, it should be acknowledged that sorting can be time-consuming and may lead to cross-contamination between samples, particularly when sorting rare and small specimens

(Elbrecht *et al.*, 2017). However, there are methods available to avoid cross-contamination, as listed in Table 1.

3.3 DNA extractions and primer selection. For bulk-samples, the best method is to use electrical homogenizers with ceramic and/or platinum beads (such as TissueLyser and FastPrep-24) along with the Qiagen DNeasy Blood & tissue kit. The weight of the tissue in the bulk-sample can range from 52 mg to 1000 mg, as reported in previous studies (Yang *et al.*, 2014; Piñol *et al.*, 2015; Barsoum *et al.*, 2019). In an experiment conducted by us, we tested two homogenization methods (liquid nitrogen and TissueLyser) and two lysis buffers (ATL and BSA) with 500 mg of bulk-sample. Our results showed that the quality and quantity of DNA improved when we macerated the sample with BSA buffer 1x and an electric homogenizer (Qiagen TissueLyser II: two times for 1 min at 30 Hz and 30 s at 30 Hz) using platinum beads, and after adding ATL digestion buffer of the Qiagen DNeasy Blood & Tissue Kit (Gálvez-Reyes *et al.*, 2021). The BSA buffer likely prevents degradation of the sample, while the maceration process allows for minimal manipulation, reducing the risk of cross-contamination between samples. Additionally, one study showed that DNA extractions performed with TissueLyser recovered six more samples (2.31%) than those using liquid nitrogen (Elbrecht and Leese, 2015).

The successful of DNA metabarcoding relies on selecting appropriate DNA markers, balancing the ability to distinguish taxa with the availability of reference sequences and base degeneracy (Deagle *et al.*, 2014; Elbrecht and Leese, 2017; Liu *et al.*, 2020). Marker loci like 16S ribosomal RNA (16S) for bacteria, Internal Transcribed Spacer 1 (ITS1) and Internal Transcribed Spacer 2 (ITS2) for fungi, Maturase K (matK) and Ribulose-1,5-bisphosphate carboxylase (rbcL) for plants, 12S ribosomal RNA (12S) for fish and vertebrates, 18S

ribosomal RNAs (18S) for eukaryotes, diatoms, and protozoa, and Cytochrome c oxidase subunit I (COI) for eukaryotes, and metazoa are chosen based on their taxonomic identification potential (Tringe and Hugenholtz, 2008; Creer *et al.*, 2010; Gibson *et al.*, 2014; Porter and Hajibabaei, 2022). Different loci's recovery success aids in consistency tests and species detection (Evans, Gilbert and Port, 2017). Using multiple loci (e.g., 16S and 18S ribosomal gene regions) allows detection of microbe species on terrestrial arthropods (Gibson *et al.*, 2014). Markers can be used together to analyze mixed tissues for plants or fungi on arthropods (Gibson *et al.*, 2014). Employing various primer sets from a single sample has multiple applications (Gibson *et al.*, 2014). Primer design depends on the study's taxonomic groups focus and the need for broad (multiple phyla) or narrow taxonomic (single order) coverage to test study-specific hypotheses. (Deiner *et al.*, 2017). Also, the success of different primer is determined by the amplicon size because there may be a trade-off in detection with amplicon size influencing detection. Short fragments are more likely to amplify, especially in eDNA, whereas longer fragments should provide better taxonomic resolution (Deiner *et al.*, 2017; Elbrecht *et al.*, 2019; Liu *et al.*, 2020). Longer amplicons also capture more genetic variability when generating ESVs or ASVs as a surrogate for haplotypes (Callahan *et al.*, 2016; Edgar, 2016; Elbrecht *et al.*, 2019; Andújar *et al.*, 2021). Markers should balance conserved primer-binding sites for broad taxonomic coverage and variable sequences for species identification (Liu *et al.*, 2020). Degenerate bases can enhance primer coverage for non-conserved sites, benefiting taxonomic recovery (Ji *et al.*, 2013; Clarke *et al.*, 2017; Elbrecht and Leese, 2017). While primer performance prediction is challenging, degenerate primers can amplify diverse taxa (Creer *et al.*, 2016; Krehenwinkel *et al.*, 2017; Andújar *et al.*, 2017). Their use improves taxonomic coverage and quantitative recovery of

species diversity (Leray *et al.*, 2013; Elbrecht and Leese, 2017; Krehenwinkel *et al.*, 2017), enhancing primer universality (Krehenwinkel *et al.*, 2017; Elbrecht *et al.*, 2019).

Using multiple gene marker suites (multi-barcode approach) is recommended to enhance taxonomic coverage, resolution, and reduce false negatives (Deagle *et al.*, 2014; Creer *et al.*, 2016). Replicates are employed to mitigate tag switching and exclude false taxonomic assignments (Lange *et al.*, 2015), but they can be influenced by stochastic effects due to low DNA yield (Elbrecht *et al.*, 2019). Taxonomic bias from PCR primer choice has raised concerns about marker utility (Taberlet *et al.*, 2012; Deagle *et al.*, 2014). However, employing multiple primer sets for the same gene region or alternative markers doesn't significantly improve species detection (Elbrecht *et al.*, 2019). The multi-barcode approach comes with drawbacks, including increased costs for primers, sequencing, labor, and expanding reference databases (Creer *et al.*, 2016).

The cytochrome oxidase I (COI) locus is a valuable tool for animal identification (Andújar, Arribas, Yu, *et al.*, 2018), offering highly variability gene for distinguishing species, making it useful for various applications like prey detection, invasive species, cryptic species revelation, and new species discovery (Bucklin, Steinke and Blanco-Bercial, 2011; Krehenwinkel *et al.*, 2017). Unlike other DNA barcoding genes (such as 18S, 16S, 12S, and ITS), COI codes for a protein, with specific sequencing requirements facilitating bioinformatics analysis (Andújar, Arribas, Yu, *et al.*, 2018; Andújar *et al.*, 2021). COI can detect a substantial proportion of species in a bulk-sample of terrestrial arthropods (Gibson *et al.*, 2014). It has been widely used as a standard in whole-community DNA metabarcoding, with available validated primers (Elbrecht *et al.*, 2019). Multiple COI primer sets have been explored, demonstrating their effectiveness (Alberdi *et al.*, 2018; Zhang *et al.*, 2018; Hajibabaei *et al.*, 2019; Elbrecht and Leese, 2017; Braukmann *et al.*, 2019). Degenerate

primers for COI subregions have improved Metazoa DNA metabarcoding (Leray *et al.*, 2013; Arribas *et al.*, 2016; Beng *et al.*, 2016; Andújar, Arribas, Yu, *et al.*, 2018), enhancing the utility, resolution, and reliability of metabarcoding data (Yu *et al.*, 2012; Ji *et al.*, 2013; Turon *et al.*, 2020; Arribas *et al.*, 2020). COI primers find extensive use in DNA metabarcoding (Table S1).

Primer success on base degeneracy and available reference data (Elbrecht and Leese, 2017). High degeneracy of primer insect taxa amplification but requires careful evaluation against reference data due to design flaws (Elbrecht and Leese, 2017). Arribas *et al.* (2016) achieved resolution in beetles, mites, and Collembola with COI primers from Yu *et al.* (2012) though they faced a 150 bp gap issue with MiSeq Illumina (2x300). Arribas *et al.* (2020) resolved this by overlapping R1 and R2 and proposed using Fold-degen-R (Yu *et al.*, 2012) and the BF (Shokralla *et al.*, 2015) primers for a 418 bp COI with MiSeq Illumina.

3.4 Library preparation using tag by PCR. Faadrosh *et al.* (2014) employed a double indexing design for the V3-V4 hypervariable regions of the 16S rRNA gene, enabling multiplexing of numerous samples. The use 8 bp indices on both the 5’ and 3’ end adapters, differing by 4 bp between samples, processing up to 200 bulk-samples on a single MiSeq lane with 20 and 10 adapters for the 5' (index 5') and 3' (index 3') ends, respectively. Fadrosh *et al.* (2014) grouped 371 microbial samples based on their 16S sequences and performed taxonomic classification.

However, before adapter preparation, it's essential to consider potential drawbacks, including degenerate primers (Yu *et al.*, 2012), optimal barcode selection, and the risk of tag-jumping in single PCR library preparations (Schnell, Bohmann and Gilbert, 2015). To address these concerns, Arribas *et al.* (2016) recommended the use of dual labeling to

maximize unique labels per library. Additionally, it is advisable to employ primers with high-performance liquid chromatography (HPLC) purification to mitigate issues (Arribas *et al.*, 2016). HPLC is a chromatographic technique for molecule purification utilizing columns with 3 to 10-micron particles. Furthermore, Arribas *et al.* (2016, 2020) suggested implementing the double dual tagging method using the XT Illumina adapter kit in library preparation.

3.5 Sequencing, bioinformatic methods and reference databases. Coverage results in DNA metabarcoding can be influenced by sample biomass size and may reflect the sample diversity. For instance, Arribas *et al.* (2016) sequenced 96 DNA metabarcoding libraries per lane as their samples were diverse. To optimize sequencing, biomass homogenization is crucial (Elbrecht *et al.*, 2017; Ji *et al.*, 2013; Yang *et al.*, 2014). To estimate bulk sample numbers for a MiSeq lane, consider sufficient COI sequence coverage per arthropod. For example, MiSeq can process 25 million reads, allowing up to 250 bulk-samples can be placed in a single lane. Shokralla *et al.* (2015) reported obtaining a sequencing coverage of 10,480,349 reads for a bulk sample comprising 1,010 individuals, while Liu *et al.* (2013) obtained 4,258,353 reads for a bulk sample comprising 795 individuals and 6,246,714 reads for a bulk sample comprising 316 individuals. Tag jumping can introduce errors, but double-indexing adapters mitigate this issue (Kircher, Sawyer and Meyer, 2012), enabling sequencing of up to 96 samples per Illumina MiSeq Lane (Figure 3, Illumina Nextera XT with dual tagging from COI amplification).

The diverse applications of sequencing platforms in arthropod biodiversity studies involve selecting the most suitable platform based on research goals and constraints (Singer *et al.*, 2019). For scalability and accuracy, Illumina Novaseq is a preferred choice, as

demonstrated by Hebert *et al.* (2018). Johnson *et al.* (2023) successfully applied Illumina technology for environmental DNA metabarcoding in arthropod analysis. Oxford Nanopore MinION, known for portability and long-read sequencing, was used by Baloğlu *et al.* (2021) and Davidov *et al.* (2020) in arthropod research and environmental studies, respectively, showcasing its versatility. MinION was also compared with Illumina by Lemoinne *et al.* (2023), revealing its ability to elucidate finer community structures. PacBio, another platform, excels in accurate identification, as indicated by Loit *et al.* (2019) and Gueidan *et al.* (2019) in the context of fungal pathogens and lichen herbarium specimens. Comparative analysis by Anslan *et al.* (2021) suggested that different Illumina platforms, including Novaseq, yield consistent results. Additionally, nanopore metabarcoding, as demonstrated by Chang *et al.* (2023), offers accurate consensus barcodes with fewer indels, making MinION a promising option for arthropod research.

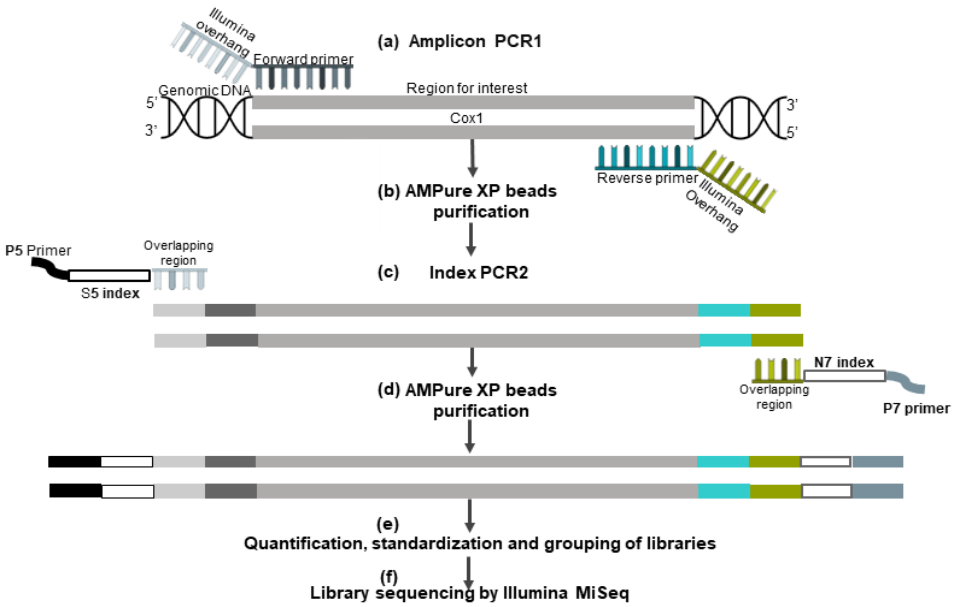

Figure 3. Library preparation of bulk-samples using Illumina MiSeq sequencing with Nextera XT adapters dual double tagging involved the following steps. (a) PCR

amplification of the COI region (PCR1): three times for each bulk-sample, using primers from the COI region with overhang adapters. (b) Pooling and purification of PCR products using Ampure XP beads (PCR clean-up). (c) Indexing of libraries via a second PCR (PCR2) with Nextera XT kit primers. (d) Another purification step using magnetic beads from the pool with Ampure XP. (e) Quantification, normalization, and library pooling, and finally. (f) Sequencing on the Illumina MiSeq platform (2×300 bp).

In an experimental validation of our recommendations were tested in a proof-of-concept experiment that involved assessing PCR bias, ligation, and sequencing. We minimized bias through various techniques including sample processing, cleanups, tissue size separation, replicates, and sequence filtering (refer to Figure 4bc and Supplementary Materials 2). Our results on diversity soup, which combine large and small specimens with different barcodes showed no significant differences in richness and compositions shown within the pools. This suggests that pools can be generated from the samples if the pooling is done up to the sequencing step. Although large (L) and small (S) specimen soups tended to have fewer ASVs, the difference with the combined pool (L+S) was not significant (refer to Figure 4bc and Supplementary Materials 2).

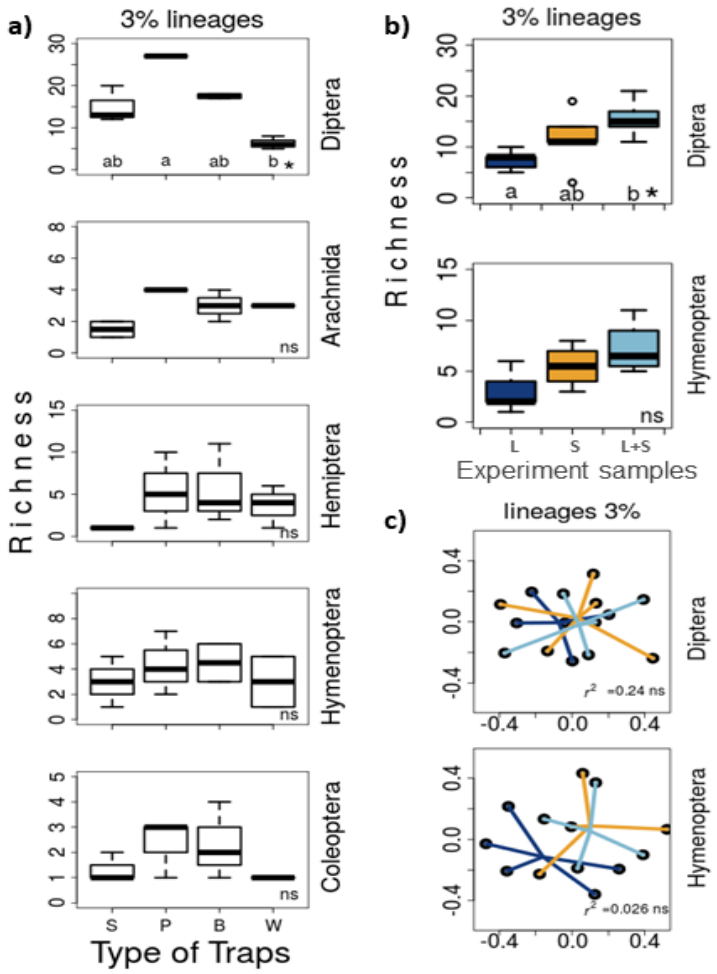

Figure 4. Arthropods richness using various sampling traps and assessed the impact of tissue size selection during DNA extraction on alpha and beta diversity. For sampling and methodological details see Supporting Material 1 and 2. **a)** Richness at 3% lineages level (similar to "species" level) across different sampling traps (S=Soapy-water, P=Pitfall traps, B=Beating tray, and W=Winkler traps). Same letters indicate that there are no significant differences among trap types. Significance codes *:p<0.05, ns: non-significant. **b)** Diptera and Hymenoptera richness at the 3% lineages level categorized by tissue sizes (L=large specimens, S=small specimens, and L+S=Soup of large and small specimens pooled together after DNA extraction). **c)** Non-Metric Multidimensional scaling (NMDS) ordinations showing community similarity (Simpson index, βsim) at 3% lineages level for Diptera and

Hymenoptera groups. Blue=large specimens, yellow=small specimens, and light-blue=Pool of large and small specimens. ns=non-significant differences of community structure estimated with ANOSIM.

Bioinformatic methods in DNA metabarcoding used in pipelines for biodiversity analysis involve multiple stages, including quality assessment, primer removal, pair merging, size filtering, denoising and dereplication. Subsequently, sequences are clustered into Operational Taxonomic Units (OTUs), independently for each library. Quality control steps involve filtering steps that are tailored to the specific experimental design and ecological question (Elbrecht *et al.*, 2018; Zinger *et al.*, 2019). However, recent pipelines, such as those described by Callahan *et al.* (2017), enable the generation of data at the haplotype level, referred to as Amplicon Sequence Variants (ASVs). Now DNA sequences can be analyzed without the need for OTU clustering (Edgar, 2016; Callahan, McMurdie and Holmes, 2017). Once an ASV matrix is generated, downstream analysis is conducted to address ecological inquiries and estimate biodiversity indices for ecosystem conservation purposes.

The identification of lineages based on various genetic similarity thresholds for OTUS clustering was proposed by Arribas *et al.* (2020). This approach encompasses quality checking, primer removal, pair merging, quality filtering, denoising, and independent library clustering. Raw reads underwent initial quality assessment using *fastqc*. Primers were trimmed, and then the reads were subsequently processed using *trimmomatic-0.36* (Bolger, Lohse and Usadel, 2014) to remove low-quality bases at the end of the reads below a specified threshold. Paired-end reads were searched using *pairfq-0.17* (Staton, 2019) and overlapping paired-end reads (R1 and R2) were merged with *usearch-9.2* (Edgar, 2013). Denoising of resulting reads generated ASVs, capable of discerning haplotypes. ASVs are

instrumental for species grouping based on DNA sequences, recognizing biological and environmental variations, and elucidating ecological patterns (Callahan, McMurdie and Holmes, 2017).

ASVs are amplified from genes that may also contain spurious sequences and nuclear mitochondrial DNA (NUMT)-driven taxonomic inflation, necessitating the filtering of such sequences (Elbrecht *et al.,* 2018; Arribas *et al.,* 2020). The removal of non-authentic ASVs requires the use of optimized filtering based on metaMATE software (Andújar *et al.,* 2021). For DNA metabarcoding studies to ensure quality and reproducibility, it is advisable to apply common filtering approaches, including quality, length, chimera, translation, frequency filtering, and denoising (Creedy *et al.,* 2022). A recent review by Creedy *et al.* (2022) on bioinformatic approaches in metabarcoding have shown a lack of harmonization, and high heterogeneity across pipelines, tasks, and tools, as well as limited adaptation of bioinformatic procedures to the COI fragment's characteristics. To ensure appropriate, comprehensive, and comparable bioinformatic analyses it is recommended to follow the standards and recommendations proposed by Creedy *et al.* (2022).

A reference database consists of DNA barcodes from well-curated morphological specimens, with DNA-barcodes being short sequences from mitochondria or chloroplast DNA that identify species (Hebert *et al.,* 2003; Ratnasingham and Hebert, 2007). To create reference databases primarily use male adult specimens for their accuracy (Elbrecht and Leese, 2017; Liu *et al.,* 2020) and play a crucial role in linking DNA sequences with taxonomic identities by associating them to physical reference collections (Liu *et al.,* 2020; Cuff *et al.,* 2021). The completeness of these databases is vital for resolution taxonomic assignments (Gibson *et al.,* 2014; Corse *et al.,* 2019), especially as DNA-based methods become more prevalent for monitoring animal communities (Leray *et al.,* 2019). Accurate

taxonomic identification depends on appropriate reference databases, which are accessible through public databases like GenBank (Benson *et al.*, 2014) and the Barcode of Life Data Systems (BOLD) (Ratnasingham and Hebert, 2007). Specialized databases for DNA metabarcoding contain millions of sequences from thousands of species, and continue to grow (Ratnasingham and Hebert, 2007), such as those provided by Porter and Hajibabaei (2020), Ji *et al.* (2013), and Magoga *et al.* (2022). BOLD contributes to GenBank, enhancing the availability of public BOLD data (Porter and Hajibabaei, 2020), and combining COI sequences from each database improves taxonomic coverage (Porter and Hajibabaei, 2020).

4. DNA metabarcoding and ASV-based approaches, status quo and future possibilities

Several studies have demonstrated the effectiveness of DNA metabarcoding in assessing arthropod biodiversity in various ecosystems (Yu *et al.*, 2012; Ji *et al.*, 2013; Bush *et al.*, 2020). It has been evaluated using amplicon and shotgun sequencing, as well as PCR-free whole mitochondrial genome sequencing of (Bista *et al.*, 2018). This approach has gained popularity in terrestrial, freshwater, and marine environments (Table S1), and has been used to analyze arthropod diversity in various contexts, such as tropical forests (Bittleston *et al.*, 2016), anthropogenic-impacted sites (Ji *et al.*, 2013; Beng *et al.*, 2016), deep-sea sediment (Sinniger *et al.*, 2016), expanding grasslands (Bowser *et al.*, 2017), Antarctic sediments (Fonseca *et al.*, 2017), and diet assessment in bat feces (Galan *et al.*, 2018). Community DNA metabarcoding can be taxonomically verified through Sanger sequencing when specimens are preserved, unlike eDNA metabarcoding with lacks whole organisms (Creer *et al.*, 2016; Deiner *et al.*, 2017). As a consequence, the difference in source material has implications for the interpretation of biodiversity over time and space (Creer *et al.*, 2016; Deiner *et al.*, 2017). Nevertheless, both community DNA and eDNA metabarcoding have successfully identifier

biodiversity drivers and responses of multiple taxa to land-use change characterized soil arthropod communities, profiled hyperdiverse mesofauna assemblages, and identifying invertebrates in freshwater and terrestrial ecosystems (Wood *et al.*, 2017; Oliverio *et al.*, 2018; Arribas *et al.*, 2016, 2021; Kirse *et al.*, 2021; Hajibabaei *et al.*, 2011; Elbrecht and Leese, 2017; Elbrecht *et al.*, 2017; Andújar *et al.*, 2018; Ji *et al.*, 2013).

The significance of assessing biodiversity response to environmental variations and comparing it with taxa- and habitat-based surrogate measures has been emphasized by Barsoum *et al.* (2019), Creedy *et al.* (2019), Wang *et al.* (2019), and Cai *et al.* (2021). These studies also highlighted the advantages of tropical forest canopy, native forests, mixed-species plantations over monoculture plantations, and insect biodiversity in pine plantations. DNA metabarcoding of terrestrial arthropods may yield varying richness levels, but provides robust beta diversity, as observed in soil arthropods (Porter *et al.*, 2019). It has the potential to evaluate wildflowers as a source of arthropod eDNA since arthropods leave DNA traces on flowers (Thomsen and Sigsgaard, 2019). Additionally, DNA metabarcoding can reveal metacommunity dynamics (Bush *et al.*, 2020).

Regarding the new opportunities offered by ASVs-based approaches, Shum and Palumbi (2021) suggest that ASVs-based approaches offer new opportunities to explore intraspecific variation using haplotype data. This can provide insights beyond species-level diversity by (i) indicating the co-occurrence of multiple individuals from the same species when multiple intraspecific haplotypes are present within a single environmental sample, (ii) revealing population structure and dispersal barriers through differences in intraspecific haplotypes between samples, and (iii) estimating demographic parameters like long-term population sizes and demographic shifts based on the distribution of haplotype differences within a species.

Various habitats have been studied using whole community metabarcoding and ASVs-based approaches, including subtropical forest, temperate woodland (Ji *et al.*, 2013), grasslands, forests (Arribas *et al.*, 2020), rainforest (Krehenwinkel, Fong, *et al.*, 2018), tropical sky-islands (Gálvez-Reyes *et al.*, 2021), and different land-uses (Noguerales *et al.*, 2021). These studies targeted on diverse taxonomic groups such as Araneae, Collembola, Myriapoda, Insecta, Nematoda, Bacteria, Fungi, and Plantae, using genetic markers like COI, 16S, 18S, ITS, and 28S. Several sequencing platforms used were with primer pairs of varying size from 127 bp to 418 bp (see Table S1). ASVs-based approaches, employing the metaMATE pipeline (Andújar *et al.*, 2021) have been used to assess alpha and beta diversity at different hierarchical levels in entire community of Acari, Collembola and Coleoptera from temperate forests (Arribas *et al.*, 2020), Collembola and Diptera communities in a tropical sky-island forest (Gálvez-Reyes *et al.*, 2021), and to understand how stochastic and niche-based processes shape structure these communities (Noguerales *et al.*, 2021).

Recent advancements in COI DNA metabarcoding datasets have allowed for the exploration, of intraspecific genetic diversity (Elbrecht *et al.*, 2018; Andújar *et al.*, 2022), enabling the use multispecies phylogeography, also known as metaphylogeography (Turon *et al.*, 2020). Carefully filtering and error removal are essential for performing metaphylogeographical analysis and obtaining ASVs (Callahan *et al.*, 2016; Callahan, McMurdie and Holmes, 2017; Andújar *et al.*, 2021). Several studies have simultaneously investigated intraspecific patterns and phylogeographic features across hundreds of species (Elbrecht *et al.*, 2018; Tsuji *et al.*, 2020; Andújar *et al.*, 2022; Antich *et al.*, 2023). For example, haplotypes from DNA metabarcoding were used to examine population genetic structure and demographic patterns in co-inhabiting kelp forest species across ecological gradients (Shum and Palumbi, 2021). DNA metabarcoding diet analysis of vampire bats

assessed predator population structure and a wide range of prey taxa (Bohmann *et al.*, 2018). Studies also evaluated the impact of environmental variables on intraspecific genetic diversity through haplotypes (Zizka, Weiss and Leese, 2020). Despite variations in laboratory procedures for DNA metabarcoding, general patterns were consistent between samples, as demonstrated in an international cross-laboratory experiment (Zaiko *et al.*, 2022). A recent study discussed the relationship between biogeographic and phylogeographic processes in marine environments using haplotype data (Antich *et al.*, 2023).

The 18S rRNA gene is a valuable tool for characterizing eukaryotic soil invertebrates in terrestrial arthropod metabarcoding. It offers comprehensive insights into soil invertebrate fauna and their responses to environmental changes. This gene has been employed in various studies to assess soil biodiversity and develop best practices for invertebrate biodiversity assessment. For instance, Singh *et al.* (2019) used it to study tropical forest conversion to rubber plantations, highlighting its ecological significance. Giebner *et al.* (2020) compared biodiversity assessment methods, emphasizing the utility of 18S rRNA and discussing its pros and cons relative to COI. Kirse *et al.* (2021) stressed its importance in accurate biodiversity assessments and conservation efforts. Smenderovac *et al.* (2022) applied 18S rRNA metabarcoding to study wood ash effects on forest soil biotic communities in Canada, demonstrating its value in sustainable forest management. Jorna *et al.* (2023) analyzed arctic tundra soil ecosystems with 18S rRNA metabarcoding, revealing diverse communities in challenging environments. These studies highlight the versatility and significance of 18S rRNA metabarcoding in terrestrial arthropods and soil invertebrate research for assessing metazoan diversity in solid substrates. The inclusion of COI and 18S sequence data gathered from arthropod-focused sampling allows the simultaneous detection of both insect and microbial taxa (Gibson *et al.*, 2014). Therefore, comparison primers that target two

frequently used genetic markers, mitochondrial cytochrome oxidase (COI) and ribosomal 18S genes can determined the utility in broad soil arthropods studies.

DNA metabarcoding holds significant potential for ecological studies, as it focuses on the current communities (Pompanon *et al.*, 2012; Taberlet *et al.*, 2012; Yu *et al.*, 2012). While DNA metabarcoding methods are evolving, there is emerging standardization in field and laboratory procedures (Andújar *et al.*, 2018; Krehenwinkel, Kennedy, *et al.*, 2018; Liu *et al.*, 2020; Creedy *et al.*, 2022; Zaiko *et al.*, 2022). Establishing a robust framework for DNA metabarcoding-based haplotype approaches is crucial to enable effective monitoring.

4.1 Insights into Insect declines: Global trends and monitoring strategies. Recent research has brought attention to concerning declines in insect populations (Butchart *et al.*, 2010; Del-Val *et al.*, 2021), yet uncertainties persist regarding the magnitude and distribution of these declines (Van Klink *et al.*, 2020). Van Klink and colleagues synthesized data from 166 extensive surveys across 1,676 sites worldwide, confirming declines in terrestrial insect populations, although not as drastic as suggested by certain earlier studies. Conversely, they noted an overall increase in freshwater insect populations, potentially attributed to initiatives promoting clean water and shifts in climate. Variations in trends suggest that localized factors likely drive many population changes, presenting opportunities for targeted conservation efforts. However, the implementation of such programs has been hindered by the time and specialized expertise required to analyze samples from entire arthropod communities (Yu *et al.*, 2012). Without extensive, minimally biased data, it is challenging to evaluate which potential drivers of long-term declines are most significant and therefore most crucial to address (Ji *et al.*, 2013; Van Klink *et al.*, 2022). Thus, high-throughput sequencing (HTS) combined with abundance measurement methods offers a pathway to practical and

comparable long-term monitoring of communities globally (Emerson *et al.*, 2023).

Moreover, sky-islands emerge as critical priorities for monitoring due to their relative biodiversity value and significance as indicators of broader global change. (Uscanga *et al.*, 2021; Mastretta-Yanes *et al.*, 2018). However, they are often overlooked in global biodiversity monitoring initiatives and biodiversity indicator frameworks (Uscanga *et al.*, 2021). The high throughput and efficiency of HTS barcoding approaches represent a viable long-term solution for monitoring and documenting changes within sky island arthropod communities (Gálvez-Reyes *et al.*, 2021). Suggestions for a coordinated approach to inventory and temporal monitoring, through spatially extensive inventories with a subset of sites subject to temporal sampling (Arribas *et al.*, 2020), can provide necessary baseline data for conservation planning (Emerson *et al.*, 2023). Range size is often used in conservation planning, prioritizing species with small ranges and frequently declining abundances. More sophisticated implementations of stratigraphically structured sampling and HTS barcoding with abundance data, such as multiplex barcoding or whole-community DNA barcoding with artificial intelligence for image recognition, also have the potential to simultaneously contribute local species records, basic niche information (stratigraphic distribution), and local abundance (Emerson *et al.*, 2023).

6. Conclusions

To ensure the relevance of biodiversity data and monitoring studies, it is crucial to integrate basic concepts and practical recommendations. The success of community metabarcoding and ASVs-based approaches depends on the conceptual framework, the hypotheses and research questions, the DNA extraction technique, the choice of genetic markers and primers

for taxonomic identification, the reference databases, and the bioinformatic methods employed. Despite the technical biases and challenges in DNA-based biodiversity analyses, molecular data can provide approximate but comparable estimates of indices to assess ecosystem health and provide promising avenues for monitoring and documenting changes within sky island arthropod communities. Also, integrating haplotype level data with ASVs-based approaches allows for testing ecological and evolutionary questions based on neutral assumptions and evaluating the effects of environmental changes at the population level, thus creating new opportunities for biodiversity monitoring, can offer valuable baseline data for conservation planning. These efforts are crucial for prioritizing species with small ranges and frequently declining abundances and addressing the broader impacts of climate change on arthropod communities.

Glossary

Amplicons. Target of DNA/RNA that is the source or product of natural or artificial events of amplification or replication. DNA/RNA amplicons of a selected marker are the results of the PCR amplification.

ASV. Amplicon Sequence Variants are DNA sequences recovered from a high-throughput marker gene analysis and distinguish sequence variation by a single nucleotide change after removal of erroneous sequences generated during PCR and sequencing.

Barcodes. A DNA barcode is one or more short gene sequences taken from standardized zones of the eukaryotic genome using a molecular tag to identify species. In a broad sense, a DNA barcode is any DNA sequence used for identification at any taxonomic level.

Monitoring. Systematic use of living organisms or their responses to determine the condition or changes in the environment.

Bulk-sample DNA. DNA derived from many individuals that represent several species.

Composition indicators. Key taxa in the trophic chain are considered representative of a broader segment of biodiversity, and they are indicators of the landscape state (Smith *et al.*, 2008).

Degenerate primer. A mixture of DNA oligonucleotides that contain different possible bases at one or several nucleotide positions, which are referred to as degenerate bases positions. The purpose of using a degenerate primer is to amplify DNA sequences that are similar but not identical. Degenerate primers are often used in amplicon sequencing, where the targeted gene(s) have some variation. Degenerate primers can be represented by IUPAC nucleotide ambiguities, or contain an inosine base, especially at 3rd codon positions for coding regions, where several different possible bases are allowed.

Ecological guild. Functional subdivision of species that plays similar roles in ecological processes and exists in the same community, in which, not necessarily show similar spatial relationships (Fauth *et al.*, 1996; Stroud *et al.*, 2015).

Ecological niche. Corresponds to the role or occupation that the species plays within a community.

Flag species. Charismatic species that serve as a symbol to attract government support, public or private donors for implementation and development of conservation programs, and it is advantageous if the species are sensitive to disturbances (Noss, 1990).

Haplotypes. Combination of alleles from different loci on a chromosome that are usually inherited together.

OTU. Operational Taxonomic Unit. An OTU could be an individual organism, a taxonomic group named as a species or genus, or a group with indeterminate evolutionary relationships that share a set of observed characters. In molecular studies molecular OTUs (MOTUs) are defined by combining reads with a certain maximum distance threshold together.

DNA metabarcoding. A rapid method of biodiversity assessment that combines two technologies: DNA taxonomy and high-throughput DNA sequencing to automated identify of several species from a single bulk sample containing whole organisms. This uses specific PCR primers to amplify DNA genes from a collection of organisms.

Phylogenetic diversity. The sum of the genetic distances that connect all the taxa in a stepped phylogeny or molecular clock. It focuses on the general evolutionary divergence between taxa rather than number of species (Simpson's Diversity) alone in a community or habitat.

Structural indicators. Features of forest that include horizontal and vertical structural complexity, age structure, stumps, natural regeneration levels, prevalence of native species and diversity of tree species (Smith *et al.*, 2008; Barsoum *et al.*, 2019).

Tag-jumping. Sequences with false label (index) combinations. It occurs when samples from one sample get mix-barcoded so it looks like they are from a different sample (or have a non-existent pair of barcodes). This happens during library preparation when there are many different labeled samples in the mix and some form chimeric sequences (annealing of similar bases in a homologous region) or when single-stranded DNA form heteroduplexes (Schnell, Bohmann and Gilbert, 2015).

Acknowledgements

We thank Brent C. Emerson and Luis David Alcaraz for their useful observations on an earlier version of the manuscript. Nancy Gálvez-Reyes is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and was supported by CONAHCYT (scholarship no. 401781). Analyses in this paper were carried out on the CONABIO computing cluster, supported by Ernesto Campos and the Subcoordinación de Soporte Informático. We are also grateful to the many people, local communities and organizations (Santiago Tlacotepec and Amanalco and Consejo Civil Mexicano para la Silvicultura Sostenible) and many people from IE-UNAM, including A. Silva, G. H. Giles, V. Reyes, M. Campos, A. Guerra, A. Villarruel, M. Garduño, and J. C. García for sampling support, T. Garrido for lab support and to F. S. Gálvez for assistance with pitfall traps.

Author Contributions

A.M.Y., N.G.R., and D.P. conceived the study. N.G.R prepared figures 1-4. Manuscript review writing was led by N.G.R and A.M.Y., with contributions from D.P. All authors approved the final version of the manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

Funding

This work was supported by CONAHCYT (178245) to AMY and DP. Also, support was obtained from IE UNAM 2016-2018 to DP.

References

Alberdi, A. *et al.* (2018) 'Scrutinizing key steps for reliable metabarcoding of environmental samples', *Methods in Ecology and Evolution*, 9(1), pp. 134–147. https://doi.org/10.1111/2041-210X.12849.

Anderson, A. *et al.* (2011) 'The potential of parasitoid Hymenoptera as bioindicators of arthropod diversity in agricultural grasslands', *Journal of Applied Ecology*, 48(2), pp. 382–390. https://doi.org/10.1111/j.1365-2664.2010.01937.x.

Anslan, S. *et al.* (2021) 'Highly comparable metabarcoding results from MGI-Tech and Illumina sequencing platforms', *PeerJ*, 9, p. e12254. https://doi.org/10.7717/peerj.12254.

Andújar, C. *et al.* (2015) 'Phylogenetic community ecology of soil biodiversity using mitochondrial metagenomics', *Molecular Ecology*, p. n/a-n/a. https://doi.org/10.1111/mec.13195.

Andújar, C., Arribas, P., Gray, C., *et al.* (2017) 'Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill', *Molecular Ecology*, 27(1), pp. 146–166. https://doi.org/10.1111/mec.14410.

Andújar, C., Arribas, P., Yu, D.W., *et al.* (2018) 'Why the COI barcode should be the community DNA metabarcode for the metazoa', *Molecular Ecology*, 27(20), pp. 3968–3975. https://doi.org/10.1111/mec.14844.

Andújar, C. *et al.* (2021) 'Validated removal of nuclear pseudogenes and sequencing artefacts from mitochondrial metabarcode data', *Molecular Ecology Resources*, 21(6), pp. 1772–1787. https://doi.org/10.1111/1755-0998.13337.

Andújar, C. *et al.* (2022) 'Community assembly and metaphylogeography of soil biodiversity: Insights from haplotype-level community DNA metabarcoding within an oceanic island', *Molecular Ecology*, 31(15), pp. 4078–4094. https://doi.org/10.1111/mec.16560.

Antich, A. *et al.* (2023) 'Metabarcoding reveals high-resolution biogeographical and metaphylogeographical patterns through marine barriers', *Journal of Biogeography*, 50(3), pp. 515–527. https://doi.org/10.1111/jbi.14548.

Arribas, P. *et al.* (2016) 'Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil', *Methods in Ecology and Evolution*, 7(9), pp. 1071–1081. https://doi.org/10.1111/2041-210X.12557.

Arribas, P. *et al.* (2020) 'The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding', *Molecular Ecology*, 30(1), pp. 48–61. https://doi.org/10.1111/mec.15591.

Bailey, J.K. *et al.* (2009) 'From Genes to Ecosystems: An Emerging Synthesis of Eco-Evolutionary Dynamics', *The New Phytologist*, 184(4), pp. 746–749. https://www.jstor.org/stable/27735828

Baloğlu, B. *et al.* (2021) 'A workflow for accurate metabarcoding using nanopore MinION sequencing', *Methods in Ecology and Evolution*, 12(5), pp. 794–804. https://doi.org/10.1111/2041-210X.13561.

Baird, D.J. and Hajibabaei, M. (2012) 'Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing.', *Molecular ecology*, 21(8), pp. 2039–44. http://www.ncbi.nlm.nih.gov/pubmed/22590728.

Barsoum, N. *et al.* (2019) 'The devil is in the detail: Metabarcoding of arthropods provides a sensitive measure of biodiversity response to forest stand composition compared with surrogate measures of biodiversity', *Ecological Indicators*, 101, pp. 313–323. https://doi.org/10.1016/j.ecolind.2019.01.023.

Baselga, A., Gómez-Rodríguez, C. and Vogler, A.P. (2015) 'Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes', *Global Ecology and Biogeography*, 24(8), pp. 873–882. https://doi.org/10.1111/geb.12322.

Baselga, A. *et al.* (2013) 'Whole-community DNA barcoding reveals a spatio-temporal continuum of biodiversity at species and genetic levels', *Nature Communications*, 4(1), p. 1892. https://doi.org/10.1038/ncomms2881.

Basset, Y. *et al.* (2012) 'Arthropod diversity in a tropical forest', *Science*, 338(6113), pp. 1481–1484. https://doi.org/10.1126/science.1226727.

Bell, G. (2001) 'Neutral Macroecology', *Science*, 293(5539), pp. 2413–2418. https://doi.org/10.1126/science.293.5539.2413.

Beng, K.C. *et al.* (2016) 'The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics.', *Scientific reports*, 6(October 2015), p. 24965. https://doi.org/10.1038/srep24965.

Benson, D.A. *et al.* (2014) 'GenBank', *Nucleic Acids Research*, 42(D1), pp. D32–D37. https://doi.org/10.1093/nar/gkt1030.

Bista, I. *et al.* (2018) 'Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples', *Molecular Ecology Resources*, 18(5), pp. 1020–1034. https://doi.org/10.1111/1755-0998.12888.

Bittleston, L.S. *et al.* (2016) 'Metabarcoding as a tool for investigating arthropod diversity in Nepenthes pitcher plants', *Austral Ecology*, 41(2), pp. 120–132. https://doi.org/10.1111/aec.12271.

Bohmann, K. *et al.* (2018) 'Using DNA metabarcoding for simultaneous inference of common vampire bat diet and population structure', *Molecular Ecology Resources*, 18(5), pp. 1050–1063. https://doi.org/10.1111/1755-0998.12891.

Bohmann, K. *et al.* (2021) 'Strategies for sample labelling and library preparation in DNA metabarcoding studies', *Molecular Ecology Resources*, 22(4), pp. 1231–1246. https://doi.org/10.1111/1755-0998.13512.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: a flexible trimmer for Illumina sequence data', *Bioinformatics*, 30(15), pp. 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bourlat, S.J. *et al.* (2013) 'Genomics in marine monitoring: new opportunities for assessing marine health status.', *Marine pollution bulletin*, 74(1), pp. 19–31. https://doi.org/10.1016/j.marpolbul.2013.05.042.

Bowser, M. *et al.* (2017) 'Arthropod and oligochaete assemblages from grasslands of the southern Kenai Peninsula, Alaska', *Biodiversity Data Journal*, 5, p. e10792. https://doi.org/10.3897/BDJ.5.e10792.

Brandon-mong, G.-J. *et al.* (2015) 'DNA metabarcoding of insects and allies: an evaluation of primers and pipelines', *Bulletin of Entomologica Research* [Preprint]. https://doi.org/10.1017/S0007485315000681.

Braukmann, T.W.A. *et al.* (2019) 'Metabarcoding a diverse arthropod mock community', *Molecular Ecology Resources*, 19(3), pp. 711–727. https://doi.org/10.1111/1755-0998.13008.

Bucklin, A., Steinke, D. and Blanco-Bercial, L. (2011) 'DNA Barcoding of Marine Metazoa', *Annual Review of Marine Science*, 3(1), pp. 471–508. https://doi.org/10.1146/annurev-marine-120308-080950.

Budd, G.E. and Telford, M.J. (2009) 'The origin and evolution of arthropods', *Nature*, 457(7231), pp. 812–817. https://doi.org/10.1038/nature07890.

Bush, A. *et al.* (2020) 'DNA metabarcoding reveals metacommunity dynamics in a threatened boreal wetland wilderness', *Proceedings of the National Academy of Sciences*, 117(15), pp. 8539–8545. https://doi.org/10.1073/pnas.1918741117.

Buss, D.F. *et al.* (2014) 'Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs', *Environmental Monitoring and Assessment*, 187(1), p. 4132. https://doi.org/10.1007/s10661-014-4132-8.

Butchart, S.H.M. *et al.* (2010) 'Global Biodiversity: Indicators of Recent Declines', *Science*, 328(5982), pp. 1164–1168. https://doi.org/10.1126/science.1187512.

Callahan, B.J. *et al.* (2016) 'DADA2: High-resolution sample inference from Illumina amplicon data', *Nature Methods*, 13(7), pp. 581–583. https://doi.org/10.1038/nmeth.3869.

Callahan, B.J., McMurdie, P.J. and Holmes, S.P. (2017) 'Exact sequence variants should replace operational taxonomic units in marker-gene data analysis', *The ISME Journal*, 11(12), pp. 2639–2643. https://doi.org/10.1038/ismej.2017.119.

Calvignac-Spencer, S. *et al.* (2013) 'Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity.', *Molecular ecology*, 22(4), pp. 915–24. https://doi.org/10.1111/mec.12183.

Castagneyrol, B. and Jactel, H. (2012) 'Unraveling plant–animal diversity relationships: a meta-regression analysis', *Ecology*, 93(9), pp. 2115–2124. https://doi.org/10.1890/11-1300.1.

Challenger, A. *et al.* (2009) 'Factores de cambio y estado de la biodiversidad', *Capital natural de México*, 2, pp. 37–73.

Chang, J.J.M. *et al.* (2023) 'Primed and ready: Nanopore metabarcoding can now recover highly accurate consensus barcodes that are generally indel-free'. bioRxiv, p. 2023.08.04.552069. https://doi.org/10.1101/2023.08.04.552069.

Chao, A. and Jost, L. (2012) 'Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size', *Ecology*, 93(12), pp. 2533–2547. https://doi.org/10.1890/11-1952.1.

Chua, P.Y.S. *et al.* (2023) 'Future of DNA-based insect monitoring', *Trends in Genetics*, 39(7), pp. 531–544. https://doi.org/10.1016/j.tig.2023.02.012.

Clarke, L.J. *et al.* (2017) 'Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies', *Ecology and Evolution*, 7(3), pp. 873–883. https://doi.org/10.1002/ece3.2667.

Coissac, E., Riaz, T. and Puillandre, N. (2012) 'Bioinformatic challenges for DNA metabarcoding of plants and animals.', *Molecular ecology*, 21(8), pp. 1834–47. https://doi.org/10.1111/j.1365-294X.2012.05550.x.

Corse, E. *et al.* (2019) 'One-locus-several-primers: A strategy to improve the taxonomic and haplotypic coverage in diet metabarcoding studies', *Ecology and Evolution*, 9(8), pp. 4603–4620. https://doi.org/10.1002/ece3.5063.

Creedy, T.J. *et al.* (2022) 'Coming of age for COI metabarcoding of whole organism community DNA: Towards bioinformatic harmonisation', *Molecular Ecology Resources*, 22(3), pp. 847–861. https://doi.org/10.1111/1755-0998.13502.

Creedy, T.J., Ng, W.S. and Vogler, A.P. (2019) 'Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy', *Ecology and Evolution*, 9(6), pp. 3105–3116. https://doi.org/10.1002/ece3.4839.

Creer, S. *et al.* (2010) 'Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises', *Molecular Ecology*, 19(s1), pp. 4–20. https://doi.org/10.1111/j.1365-294X.2009.04473.x.

Creer, S. *et al.* (2016) 'The ecologist's field guide to sequence-based identification of biodiversity', *Methods in Ecology and Evolution* [Preprint]. https://doi.org/10.1111/2041-210X.12574.

Cuff, J.P. *et al.* (2021) 'Overcoming the pitfalls of merging dietary metabarcoding into ecological networks', *Methods in Ecology and Evolution*, 13(3), pp. 545–559. https://doi.org/10.1111/2041-210X.13796.

Davidov, K. *et al.* (2020) 'Identification of plastic-associated species in the Mediterranean Sea using DNA metabarcoding with Nanopore MinION', *Scientific Reports*, 10(1), p. 17533. https://doi.org/10.1038/s41598-020-74180-z.

Deagle, B.E. *et al.* (2014) 'DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match', *Biology Letters*, 10(9), p. 20140562. https://doi.org/10.1098/rsbl.2014.0562.

Deiner, K. *et al.* (2017) 'Environmental DNA metabarcoding: Transforming how we survey animal and plant communities', *Molecular Ecology*, 26(21), pp. 5872–5895. https://doi.org/10.1111/mec.14350.

Del-Val, E. *et al.* (2021) 'Comparison of arthropod communities between high and low input maize farms in Mexico', *CABI Agriculture and Bioscience*, *2*, 1-10. https://doi.org/10.1186/s43170-021-00060-9

Dickie, I.A. *et al.* (2018) 'Towards robust and repeatable sampling methods in eDNA-based studies', *Molecular Ecology Resources*, 18(5), pp. 940–952. https://doi.org/10.1111/1755-0998.12907.

Dobson, A. *et al.* (2021) 'Biodiversity loss due to more than climate change', *Science*, 374(6568), pp. 699–700. https://doi.org/10.1126/science.abm6216.

Drummond, A.J. *et al.* (2015) 'Evaluating a multigene environmental DNA approach for biodiversity assessment', *GigaScience*, 4(1), pp. s13742-015-0086–1. https://doi.org/10.1186/s13742-015-0086-1.

Edgar, R.C. (2013) 'UPARSE: highly accurate OTU sequences from microbial amplicon reads', *Nature Methods*, 10(10), pp. 996–998. https://doi.org/10.1038/nmeth.2604.

Edgar, R.C. (2016) 'UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing', *bioRxiv*, p. 81257. https://doi.org/10.1101/081257.

Elbrecht, V. *et al.* (2016) 'Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects', *PeerJ*, 4, p. e1966. https://doi.org/10.7717/peerj.1966.

Elbrecht, V. *et al.* (2017) 'Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring', *Methods in Ecology and Evolution*, 8(10), pp. 1265–1275. https://doi.org/10.1111/2041-210X.12789.

Elbrecht, V. *et al.* (2018) 'Estimating intraspecific genetic diversity from community DNA metabarcoding data', *PeerJ*, 6, p. e4644. https://doi.org/10.7717/peerj.4644.

Elbrecht, V. *et al.* (2019) 'Validation of COI metabarcoding primers for terrestrial arthropods', *PeerJ*, 7, p. e7745. https://doi.org/10.7717/peerj.7745.

Elbrecht, V. and Leese, F. (2015) 'Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass - sequence relationships with an innovative metabarcoding protocol', *Peer J* [Preprint].

Elbrecht, V. and Leese, F. (2017) 'Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment', *Frontiers in Environmental Science*, 5. https://www.frontiersin.org/articles/10.3389/fenvs.2017.00011

Elbrecht, V., Peinert, B. and Leese, F. (2017) 'Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding', *Ecology and Evolution*, 7(17), pp. 6918–6926. https://doi.org/10.1002/ece3.3192.

Emerson, B.C. *et al.* (2017) 'A combined field survey and molecular identification protocol for comparing forest arthropod biodiversity across spatial scales', *Molecular Ecology Resources*, 17(4), pp. 694–707. https://doi.org/10.1111/1755-0998.12617.

Emerson, B.C. *et al.* (2023) 'Collective and harmonized high throughput barcoding of insular arthropod biodiversity: Toward a Genomic Observatories Network for islands', *Molecular Ecology, 32*(23), 6161-6176. https://doi.org/10.1111/mec.16683

Erdozain, M. *et al.* (2019) 'Metabarcoding of storage ethanol vs. conventional morphometric identification in relation to the use of stream macroinvertebrates as ecological indicators in forest management', *Ecological Indicators*, 101, pp. 173–184. https://doi.org/10.1016/j.ecolind.2019.01.014.

Evans, D.M., Gilbert, J.D.J. and Port, G.R. (2017) 'Everything is connected: network thinking in entomology', *Ecological Entomology*, 42(S1), pp. 1–3. https://doi.org/10.1111/een.12449.

Fadrosh, D.W. *et al.* (2014) 'An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform', *Microbiome*, 2(1), p. 6. https://doi.org/10.1186/2049-2618-2-6.

Fauth, J.E. *et al.* (1996) 'Simplifying the Jargon of Community Ecology: A Conceptual Approach', *The American Naturalist*, 147(2), pp. 282–286. https://doi.org/10.1086/285850.

Favreau, J.M. *et al.* (2006) 'Recommendations for Assessing the Effectiveness of Surrogate Species Approaches', *Biodiversity and Conservation*, 15(12), pp. 3949–3969. https://doi.org/10.1007/s10531-005-2631-1.

Fonseca, V.G. *et al.* (2017) 'Revealing higher than expected meiofaunal diversity in Antarctic sediments: A metabarcoding approach', *Scientific Reports*, 7(1), pp. 1–11. https://doi.org/10.1038/s41598-017-06687-x.

Galan, M. *et al.* (2018) 'Metabarcoding for the parallel identification of several hundred predators and their prey: Application to bat species diet analysis', *Molecular Ecology Resources*, 18(3), pp. 474–489. https://doi.org/10.1111/1755-0998.12749.

Gálvez-Reyes, N. *et al.* (2021) 'Dispersal limitations and long-term persistence drive differentiation from haplotypes to communities within a tropical sky-island: Evidence from community metabarcoding', *Molecular Ecology*, 30(24), pp. 6611–6626. https://doi.org/10.1111/mec.16195.

Gaspar, C., Gaston, K.J. and Borges, P. a.V. (2010) 'Arthropods as surrogates of diversity at different spatial scales', *Biological Conservation*, 143(5), pp. 1287–1294. https://doi.org/10.1016/j.biocon.2010.03.007.

Gaston, K.J. and Spicer, J.I. (2013) *Biodiversity: An Introduction*. John Wiley & Sons.

Gibson, J. *et al.* (2014) 'Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics', *Proceedings of the National Academy of Sciences*, 111(22), pp. 8007–8012. https://doi.org/10.1073/pnas.1406468111.

Giebner, H. *et al.* (2020) 'Comparing diversity levels in environmental samples: DNA sequence capture and metabarcoding approaches using 18S and COI genes', *Molecular Ecology Resources*, 20(5), pp. 1333–1345. https://doi.org/10.1111/1755-0998.13201.

Gómez-Rodríguez, C. *et al.* (2019) 'Understanding dispersal limitation through the assessment of diversity patterns across phylogenetic scales below the species level', *Global Ecology and Biogeography*, 28(3), pp. 353–364. https://doi.org/10.1111/geb.12857.

Gómez-Rodríguez, C. and Baselga, A. (2018) 'Variation among European beetle taxa in patterns of distance decay of similarity suggests a major role of dispersal processes', *Ecography*, 41(11), pp. 1825–1834. https://doi.org/10.1111/ecog.03693.

Gueidan, C. *et al.* (2019) 'PacBio amplicon sequencing for metabarcoding of mixed DNA samples from lichen herbarium specimens', *MycoKeys*, 53, pp. 73–91. https://doi.org/10.3897/mycokeys.53.34761.

Haddad, N.M. *et al.* (2009) 'Plant species loss decreases arthropod diversity and shifts trophic structure', *Ecology Letters*, 12(10), pp. 1029–1039. https://doi.org/10.1111/j.1461-0248.2009.01356.x.

Hajibabaei, M. *et al.* (2011) 'Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos.', *PloS one*, 6(4), p. e17497. https://doi.org/10.1371/journal.pone.0017497.

Hajibabaei, M. *et al.* (2016) 'A new way to contemplate Darwin's tangled bank: how DNA barcodes are reconnecting biodiversity science and biomonitoring', *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 371(1702). https://doi.org/10.1098/rstb.2015.0330.

Hajibabaei, M. *et al.* (2019) 'COI metabarcoding primer choice affects richness and recovery of indicator taxa in freshwater systems', *bioRxiv*, p. 572628. https://doi.org/10.1101/572628.

Hebert, P.D.N. *et al.* (2003) 'Biological identifications through DNA barcodes.', *Proceedings of the Royal Society B: Biological Sciences*, 270(1512), pp. 313–321. https://doi.org/10.1098/rspb.2002.2218.

Hebert, P.D.N. *et al.* (2018) 'A Sequel to Sanger: amplicon sequencing that scales', *BMC Genomics*, 19(1), p. 219. https://doi.org/10.1186/s12864-018-4611-3.

Hoban, S. *et al.* (2022) 'Global genetic diversity status and trends: towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition', *Biological Reviews*, 97(4), pp. 1511–1538. https://doi.org/10.1111/brv.12852.

Holman, L.E., Chng, Y. and Rius, M. (2022) 'How does eDNA decay affect metabarcoding experiments?', *Environmental DNA*, 4(1), pp. 108–116. https://doi.org/10.1002/edn3.201.

Hubbell, S. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press. https://doi.org/10.2307/j.ctt7rj8w.

Humphrey, J.W. *et al.* (1999) 'Relationships between insect diversity and habitat characteristics in plantation forests', *Forest Ecology and Management*, 113(1), pp. 11–21. https://doi.org/10.1016/S0378-1127(98)00413-7.

Ji, Y. *et al.* (2013) 'Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding', *Ecology Letters*, 16(10), pp. 1245–1257. https://doi.org/10.1111/ele.12162.

Johnson, M.D. *et al.* (2023) 'Environmental DNA metabarcoding from flowers reveals arthropod pollinators, plant pests, parasites, and potential predator–prey interactions while revealing more arthropod diversity than camera traps', *Environmental DNA*, 5(3), pp. 551–569. https://doi.org/10.1002/edn3.411.

Jorna, J. *et al.* (2023) 'Metabarcoding inventory of an arctic tundra soil ecosystem reveals highly heterogeneous communities at a small scale', *Polar Biology*, 46(5), pp. 461–471. https://doi.org/10.1007/s00300-023-03131-x.

de Kerdrel, G.A. *et al.* (2020) 'Rapid and cost-effective generation of single specimen multilocus barcoding data from whole arthropod communities by multiple levels of multiplexing', *Scientific Reports*, 10(1), p. 78. https://doi.org/10.1038/s41598-019-54927-z.

Kircher, M., Sawyer, S. and Meyer, M. (2012) 'Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform', *Nucleic Acids Research*, 40(1), p. e3. https://doi.org/10.1093/nar/gkr771.

Kirse, A. *et al.* (2021) 'Unearthing the Potential of Soil eDNA Metabarcoding—Towards Best Practice Advice for Invertebrate Biodiversity Assessment', *Frontiers in Ecology and Evolution*, 9. https://www.frontiersin.org/articles/10.3389/fevo.2021.630560 (Accessed: 23 December 2022).

Kimura, M. (1968) ‘Evolutionary rate at the molecular level’, *Nature*, 217(5129), pp. 624–626. https://doi.org/10.1038/217624a0.

Kimura, M. (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Koivula, M.J. (2011) ‘Useful model organisms, indicators, or both? Ground beetles (Coleoptera, Carabidae) reflecting environmental conditions’, *ZooKeys*, (100), pp. 287–317. https://doi.org/10.3897/zookeys.100.1533.

Krehenwinkel, H. *et al.* (2017) ‘Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding’, *Scientific Reports*, 7(1), pp. 1–12. https://doi.org/10.1038/s41598-017-17333-x.

Krehenwinkel, H., Kennedy, S.R., *et al.* (2018) ‘Scaling up DNA barcoding – Primer sets for simple and cost efficient arthropod systematics by multiplex PCR and Illumina amplicon sequencing’, *Methods in Ecology and Evolution*, 9(11), pp. 2181–2193. https://doi.org/10.1111/2041-210X.13064.

Krehenwinkel, H., Fong, M., *et al.* (2018) ‘The effect of DNA degradation bias in passive sampling devices on metabarcoding studies of arthropod communities and their associated microbiota’, *PLoS ONE*, 13(1), pp. 1–14. https://doi.org/10.1371/journal.pone.0189188.

Krueger, F., Andrews, S.R. and Osborne, C.S. (2011) ‘Large Scale Loss of Data in Low-Diversity Illumina Sequencing Libraries Can Be Recovered by Deferred Cluster Calling’, *PLOS ONE*, 6(1), p. e16607. https://doi.org/10.1371/journal.pone.0016607.

Lange, A. *et al.* (2015) ‘AmpliconDuo: A Split-Sample Filtering Protocol for High-Throughput Amplicon Sequencing of Microbial Communities’, *PLOS ONE*, 10(11), p. e0141590. https://doi.org/10.1371/journal.pone.0141590.

Lemoinne, A. *et al.* (2023) ‘Fine-scale congruence in bacterial community structure from marine sediments sequenced by short-reads on Illumina and long-reads on Nanopore’. bioRxiv, p. 2023.06.06.541006. https://doi.org/10.1101/2023.06.06.541006.

Loit, K. *et al.* (2019) ‘Relative Performance of MinION (Oxford Nanopore Technologies) versus Sequel (Pacific Biosciences) Third-Generation Sequencing Instruments in Identification of Agricultural and Forest Fungal Pathogens’, *Applied and Environmental Microbiology*, 85(21), pp. e01368-19. https://doi.org/10.1128/AEM.01368-19.

Leray, M. *et al.* (2013) ‘A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents.’, *Frontiers in zoology*, 10(1), p. 34. https://doi.org/10.1186/1742-9994-10-34.

Leray, M. *et al.* (2019) ‘GenBank is a reliable resource for 21st century biodiversity research’, *Proceedings of the National Academy of Sciences*, 116(45), pp. 22651–22656. https://doi.org/10.1073/pnas.1911714116.

Li, L., Zheng, B. and Liu, L. (2010) ‘Biomonitoring and Bioindicators Used for River Ecosystems: Definitions, Approaches and Trends’, *Procedia Environmental Sciences*, 2, pp. 1510–1524. https://doi.org/10.1016/j.proenv.2010.10.164.

Lim, J.Y. *et al.* (2022) ‘Semi-quantitative metabarcoding reveals how climate shapes arthropod community assembly along elevation gradients on Hawaii Island’, *Molecular Ecology*, 31(5), pp. 1416–1429. https://doi.org/10.1111/mec.16323.

Lindenmayer, D.B., Franklin, J.F. and Fischer, J. (2006) 'General management principles and a checklist of strategies to guide forest biodiversity conservation', *Biological Conservation*, 131(3), pp. 433–445. https://doi.org/10.1016/j.biocon.2006.02.019.

Liu, M. *et al.* (2020) 'A practical guide to DNA metabarcoding for entomological ecologists', *Ecological Entomology*, 45(3), pp. 373–385. https://doi.org/10.1111/een.12831.

Liu, S. *et al.* (2013) 'SOAP B arcode: revealing arthropod biodiversity through assembly of Illumina shotgun sequences of PCR amplicons', *Methods in Ecology and Evolution*. Edited by D. Orme, 4(12), pp. 1142–1150. https://doi.org/10.1111/2041-210X.12120.

van der Loos, L.M. and Nijland, R. (2021) 'Biases in bulk: DNA metabarcoding of marine communities and the methodology involved', *Molecular Ecology*, 30(13), pp. 3270–3288. https://doi.org/10.1111/mec.15592.

McCravy, K.W. (2018) 'A Review of Sampling and Monitoring Methods for Beneficial Arthropods in Agroecosystems', *Insects*, 9(4), p. 170. https://doi.org/10.3390/insects9040170.

Mace, G.M. *et al.* (2014) 'Approaches to defining a planetary boundary for biodiversity', *Global Environmental Change*, 28, pp. 289–297. https://doi.org/10.1016/j.gloenvcha.2014.07.009.

McGill, B.J. *et al.* (2007) 'Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework', *Ecology Letters*, 10(10), pp. 995–1015. https://doi.org/10.1111/j.1461-0248.2007.01094.x.

McGill, B.J. *et al.* (2019) 'Unifying macroecology and macroevolution to answer fundamental questions about biodiversity', *Global Ecology and Biogeography*, 28(12), pp. 1925–1936. https://doi.org/10.1111/geb.13020.

Magoga, G. *et al.* (2022) 'Curation of a reference database of COI sequences for insect identification through DNA metabarcoding: COins', *Database*, 2022, p. baac055. https://doi.org/10.1093/database/baac055.

Markert, B. *et al.* (1999) 'The use of bioindicators for monitoring the heavy-metal status of the environment', *Journal of Radioanalytical and Nuclear Chemistry*, 240(2), pp. 425–429. https://doi.org/10.1007/BF02349387.

Meyer, A. *et al.* (2021) 'Morphological vs. DNA metabarcoding approaches for the evaluation of stream ecological status with benthic invertebrates: Testing different combinations of markers and strategies of data filtering', *Molecular Ecology*, 30(13), pp. 3203–3220. https://doi.org/10.1111/mec.15723.

Morin, P. (1999) 'Productivity, Intraguild Predation, and Population Dynamics in Experimental Food Webs', *Ecology*, 80(3), pp. 752–760. https://doi.org/10.1890/0012-9658(1999)080[0752:PIPAPD]2.0.CO;2.

Mueller, M. and Geist, J. (2016) 'Conceptual guidelines for the implementation of the ecosystem approach in biodiversity monitoring', *Ecosphere*, 7(5), p. e01305. https://doi.org/10.1002/ecs2.1305.

Miyata, K. *et al.* (2022) 'Comparative environmental RNA and DNA metabarcoding analysis of river algae and arthropods for ecological surveys and water quality assessment', *Scientific Reports*, 12(1), p. 19828. https://doi.org/10.1038/s41598-022-23888-1.

Múrria, C. *et al.* (2017) 'Local environment rather than past climate determines community composition of mountain stream macroinvertebrates across Europe', *Molecular Ecology*, 26(21), pp. 6085–6099. https://doi.org/10.1111/mec.14346.

Niemelä, J. (2000) 'Biodiversity monitoring for decision-making', *Annales Zoologici Fennici*, 37(4), pp. 307–317. https://www.jstor.org/stable/23735723 (Accessed: 12 May 2023).

Noguerales, V. *et al.* (2021) 'Community metabarcoding reveals the relative role of environmental filtering and spatial processes in metacommunity dynamics of soil microarthropods across a mosaic of montane forests', *Molecular Ecology*, n/a(n/a), pp. 1–19. https://doi.org/10.1111/mec.16275.

Noss, R.F. (1990) 'Indicators for Monitoring Biodiversity: A Hierarchical Approach', *Conservation Biology*, 4, pp. 355–364. https://doi.org/10.1111/j.1523-1739.1990.tb00309.x.

Noss, R.F. (1999) 'Assessing and monitoring forest biodiversity: A suggested framework and indicators', *Forest Ecology and Management*, 115, pp. 135–146. https://doi.org/10.1016/S0378-1127(98)00394-6.

Obrist, M.K. and Duelli, P. (2010) 'Rapid biodiversity assessment of arthropods for monitoring average local species richness and related ecosystem services', *Biodiversity and Conservation*, 19(8), pp. 2201–2220. https://doi.org/10.1007/s10531-010-9832-y.

Ødegaard, F. (2000) 'How many species of arthropods? Erwin's estimate revised', *Biological Journal of the Linnean Society*, 71(4), pp. 583–597. https://doi.org/10.1111/j.1095-8312.2000.tb01279.x.

Oliverio, A.M. *et al.* (2018) 'A DNA metabarcoding approach to characterize soil arthropod communities', *Soil Biology and Biochemistry*, 125, pp. 37–43. https://doi.org/10.1016/j.soilbio.2018.06.026.

Paliy, O. and Shankar, V. (2016) 'Application of multivariate statistical techniques in microbial ecology', *Molecular Ecology*, 25(5), pp. 1032–1057. https://doi.org/10.1111/mec.13536.

Papadopoulou, A. *et al.* (2011) 'Testing the Species–Genetic Diversity Correlation in the Aegean Archipelago: Toward a Haplotype-Based Macroecology?', *The American Naturalist*, 178(2), pp. 241–255. https://doi.org/10.1086/660828.

Parmesan, C. and Yohe, G. (2003) 'A globally coherent fingerprint of climate change impacts across natural systems', *Nature*, 421(6918), pp. 37–42. https://doi.org/10.1038/nature01286.

Pawson, S.M. *et al.* (2011) 'Maximising biodiversity in plantation forests: Insights from long-term changes in clearfell-sensitive beetles in a Pinus radiata plantation', *Biological Conservation*, 144(12), pp. 2842–2850. https://doi.org/10.1016/j.biocon.2011.08.001.

Pedro, P.M. *et al.* (2017) 'Metabarcoding Analyses Enable Differentiation of Both Interspecific Assemblages and Intraspecific Divergence in Habitats With Differing Management Practices', *Environmental Entomology*, 46(6), pp. 1381–1389. https://doi.org/10.1093/ee/nvx166.

Pimm, S.L. *et al.* (1995) 'The Future of Biodiversity', *Science*, 269(5222), pp. 347–350. https://doi.org/10.1126/science.269.5222.347.

Piñol, J. *et al.* (2015) 'Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods', *Molecular Ecology Resources*, 15(4), pp. 819–830. https://doi.org/10.1111/1755-0998.12355.

Pompanon, F. *et al.* (2012) 'Who is eating what: diet assessment using next generation sequencing.', *Molecular ecology*, 21(8), pp. 1931–50. https://doi.org/10.1111/j.1365-294X.2011.05403.x.

Porter, T.M. *et al.* (2019) 'Variations in terrestrial arthropod DNA metabarcoding methods recovers robust beta diversity but variable richness and site indicators', *Scientific Reports*, 9(1), p. 18218. https://doi.org/10.1038/s41598-019-54532-0.

Porter, T.M. and Hajibabaei, M. (2018) 'Automated high throughput animal CO1 metabarcode classification', *Scientific Reports*, 8(1), p. 4226. https://doi.org/10.1038/s41598-018-22505-4.

Porter, T.M. and Hajibabaei, M. (2020) 'Putting COI Metabarcoding in Context: The Utility of Exact Sequence Variants (ESVs) in Biodiversity Analysis', *Frontiers in Ecology and Evolution*, 8. https://www.frontiersin.org/articles/10.3389/fevo.2020.00248 (Accessed: 15 May 2023).

Porter, T.M. and Hajibabaei, M. (2022) 'MetaWorks: A flexible, scalable bioinformatic pipeline for high-throughput multi-marker biodiversity assessments', *PLOS ONE*, 17(9), p. e0274260. https://doi.org/10.1371/journal.pone.0274260.

Ratnasingham, S. and Hebert, P.D.N. (2007) 'bold: The Barcode of Life Data System (http://www.barcodinglife.org)', *Molecular Ecology Notes*, 7(3), pp. 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x.

Rockström, J., W. Steffen, K. Noone, Å. Persson, F. S. Chapin, III, E. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. Schellnhuber, B. Nykvist, C. A. De Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P. K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. W. Corell, V. J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J. Foley (2009) Planetary boundaries: exploring the safe operating space for humanity, *Ecology and society*, 14(2). https://www.jstor.org/stable/26268316

Rosindell, J., Hubbell, S.P. and Etienne, R.S. (2011) 'The Unified Neutral Theory of Biodiversity and Biogeography at Age Ten', *Trends in Ecology and Evolution*, 26(7), pp. 340–348. https://doi.org/10.1016/j.tree.2011.03.024.

Roslin, T. (2001) 'Large-scale spatial ecology of dung beetles', *Ecography*, 24(5), pp. 511–524. https://doi.org/10.1111/j.1600-0587.2001.tb00486.x.

Ruppert, K.M., Kline, R.J. and Rahman, M.S. (2019) 'Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA', *Global Ecology and Conservation*, 17, p. e00547. https://doi.org/10.1016/j.gecco.2019.e00547.

Sala, O.E. *et al.* (2000) 'Global Biodiversity Scenarios for the Year 2100', *Science*, 287(5459), pp. 1770–1774. https://doi.org/10.1126/science.287.5459.1770.

Samways, M.J., McGeoch, M.A. and New, T.R. (2010) *Insect Conservation: A Handbook of Approaches and Methods*. Oxford University Press.

Schmeller, D.S. *et al.* (2015) 'Towards a global terrestrial species monitoring program', *Journal for Nature Conservation*, 25, pp. 51–57. https://doi.org/10.1016/j.jnc.2015.03.003.

Schmidt, T.M., DeLong, E.F. and Pace, N.R. (1991) 'Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing', *Journal of Bacteriology*, 173(14), pp. 4371–4378. https://doi.org/10.1128/jb.173.14.4371-4378.1991.

Schnell, I.B., Bohmann, K. and Gilbert, M.T.P. (2015) 'Tag jumps illuminated--reducing sequence-to-sample misidentifications in metabarcoding studies.', *Molecular ecology resources*, 15(6), pp. 1289–303. https://doi.org/10.1111/1755-0998.12402.

Shokralla, S. *et al.* (2015) 'Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform', *Scientific Reports*, 5(1), p. 9687. https://doi.org/10.1038/srep09687.

Shum, P. and Palumbi, S.R. (2021) 'Testing small-scale ecological gradients and intraspecific differentiation for hundreds of kelp forest species using haplotypes from metabarcoding', *Molecular Ecology*, 30(13), pp. 3355–3373. https://doi.org/10.1111/mec.15851.

Singer, G. a. C. *et al.* (2019) 'Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater', *Scientific Reports*, 9(1), p. 5991. https://doi.org/10.1038/s41598-019-42455-9.

Singh, D. *et al.* (2019) 'Tropical forest conversion to rubber plantation affects soil micro- & mesofaunal community & diversity', *Scientific Reports*, 9(1), p. 5893. https://doi.org/10.1038/s41598-019-42333-4.

Sinniger, F. *et al.* (2016) 'Worldwide Analysis of Sedimentary DNA Reveals Major Gaps in Taxonomic Knowledge of Deep-Sea Benthos', *Frontiers in Marine Science*, 3. https://www.frontiersin.org/articles/10.3389/fmars.2016.00092 (Accessed: 15 May 2023).

Skórka, P., Settele, J. and Woyciechowski, M. (2007) 'Effects of management cessation on grassland butterflies in southern Poland', *Agriculture, Ecosystems & Environment*, 121(4), pp. 319–324. https://doi.org/10.1016/j.agee.2006.11.001.

Smenderovac, E. *et al.* (2022) 'Forest soil biotic communities show few responses to wood ash applications at multiple sites across Canada', *Scientific Reports*, 12(1), p. 4171. https://doi.org/10.1038/s41598-022-07670-x.

Smith, G.F. *et al.* (2008) 'Identifying practical indicators of biodiversity for stand-level management of plantation forests', *Biodiversity and Conservation*, 17(5), pp. 991–1015. https://doi.org/10.1007/s10531-007-9274-3.

Spence, J.R. *et al.* (2008) 'Conservation of forest-dwelling arthropod species: Simultaneous management of many small and heterogeneous risks', *Canadian Entomologist*, 140(4), pp. 510–525. https://doi.org/10.4039/N07-Ls05.

Staton, E. (2019) 'sestaton/Pairfq'. https://github.com/sestaton/Pairfq (Accessed: 30 June 2020).

Stein, E.D. *et al.* (2013) 'Evaluating Ethanol-based Sample Preservation to Facilitate Use of DNA Barcoding in Routine Freshwater Biomonitoring Programs Using Benthic Macroinvertebrates', *PLOS ONE*, 8(1), p. e51273. https://doi.org/10.1371/journal.pone.0051273.

Steinke, D. *et al.* (2021) 'Effects of Malaise trap spacing on species richness and composition of terrestrial arthropod bulk samples', *Metabarcoding and Metagenomics*, 5, p. e59201. https://doi.org/10.3897/mbmg.5.59201.

Stork, N.E. *et al.* (2015) 'New approaches narrow global species estimates for beetles, insects, and terrestrial arthropods', *Proceedings of the National Academy of Sciences*, 112(24), pp. 7519–7523. https://doi.org/10.1073/pnas.1502408112.

Stroud, J.T. *et al.* (2015) 'Is a community still a community? Reviewing definitions of key terms in community ecology', *Ecology and Evolution*, 5(21), pp. 4757–4765. https://doi.org/10.1002/ece3.1651.

Taberlet, P. *et al.* (2012) 'Environmental DNA', *Molecular Ecology*, 21(8), pp. 1789–1793.

Taberlet, P. *et al.* (2018) *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford University Press.

Thomsen, P.F. and Sigsgaard, E.E. (2019) 'Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods', *Ecology and Evolution*, 9(4), pp. 1665–1679. https://doi.org/10.1002/ece3.4809.

Tringe, S.G. and Hugenholtz, P. (2008) 'A renaissance for the pioneering 16S rRNA gene', *Current Opinion in Microbiology*, 11(5), pp. 442–446. https://doi.org/10.1016/j.mib.2008.09.011.

Tsuji, S. *et al.* (2020) 'Quantitative evaluation of intraspecific genetic diversity in a natural fish population using environmental DNA analysis', *Molecular Ecology Resources*, 20(5), pp. 1323–1332. https://doi.org/10.1111/1755-0998.13200.

Turon, X. *et al.* (2020) 'From metabarcoding to metaphylogeography: separating the wheat from the chaff', *Ecological Applications*, 30(2), p. e02036. https://doi.org/10.1002/eap.2036.

Uscanga, A. *et al.* (2021) 'Evaluating species origins within tropical sky-islands arthropod communities', *Journal of Biogeography*, 48(9), pp. 2199–2210. https://doi.org/10.1111/jbi.14144.

Van Klink, R. *et al.* (2020) Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances, *Science*, *368*(6489), 417-420. https://doi.org/10.1126/science.aax9931

Vellend, M. (2010) '2010 Community Ecology', 85(2), pp. 183–206. https://doi.org/10.1086/652373.

Vellend, M. and Geber, M.A. (2005) 'Connections between species diversity and genetic diversity', *Ecology Letters*, 8(7), pp. 767–781. https://doi.org/10.1111/j.1461-0248.2005.00775.x.

Watts, C. *et al.* (2019) 'DNA metabarcoding as a tool for invertebrate community monitoring: a case study comparison with conventional techniques', *Austral Entomology*, 58(3), pp. 675–686. https://doi.org/10.1111/aen.12384.

Weber, M.G. *et al.* (2017) 'Evolution in a Community Context: On Integrating Ecological Interactions and Macroevolution', *Trends in Ecology and Evolution*, 32(4), pp. 291–304. https://doi.org/10.1016/j.tree.2017.01.003.

Wheat, C.W. *et al.* (2007) 'The genetic basis of a plant–insect coevolutionary key innovation', *Proceedings of the National Academy of Sciences*, 104(51), pp. 20427–20431. https://doi.org/10.1073/pnas.0706229104.

Whittaker, R.H. (1975) 'Communities and ecosystems, 2nd edn Macmillan Publishing Company', *New York, NY* [Preprint].

Wilson, M.C. *et al.* (2016) 'Habitat fragmentation and biodiversity conservation: key findings and future challenges', *Landscape Ecology*, 31(2), pp. 219–227. https://doi.org/10.1007/s10980-015-0312-3.

Wood, J.R. *et al.* (2017) 'No single driver of biodiversity: divergent responses of multiple taxa across land use types', *Ecosphere*, 8(11), p. e01997. https://doi.org/10.1002/ecs2.1997.

Woodward, G., Gray, C. and Baird, D.J. (2013) 'Biomonitoring for the 21st Century: new perspectives in an age of globalisation and emerging environmental threats', *Limnetica,32*, 2(2), pp. 159–174.

Yang, Chenxue *et al.* (2014) 'Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator', *Ecological Indicators*, 46, pp. 379–389. https://doi.org/10.1016/j.ecolind.2014.06.028.

Yu, D.W. *et al.* (2012) 'Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring', *Methods in Ecology and Evolution*, 3(4), pp. 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x.

Zaiko, A. *et al.* (2022) 'Towards reproducible metabarcoding data: Lessons from an international cross-laboratory experiment', *Molecular Ecology Resources*, 22(2), pp. 519–538. https://doi.org/10.1111/1755-0998.13485.

Zhang, G.K. *et al.* (2018) 'Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities', *Evolutionary Applications*, 11(10), pp. 1901–1914. https://doi.org/10.1111/eva.12694.

Zhang, K. *et al.* (2016) 'Plant diversity accurately predicts insect diversity in two tropical landscapes', *Molecular ecology*, 68125438. https://doi.org/10.1111/MEC.13770.

Zhang, Z.-Q. (2013) *Animal biodiversity: An update of classification and diversity in 2013. In: Zhang, Z.-Q. (Ed.) Animal Biodiversity: An Outline of Higher-level Classification and Survey of Taxonomic Richness (Addenda 2013)*. https://www.biotaxa.org/Zootaxa/article/view/zootaxa.3703.1.3 (Accessed: 12 May 2023).

Zhou, X. *et al.* (2013) 'Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification.', *GigaScience*, 2(1), p. 4. https://doi.org/10.1186/2047-217X-2-4.

Zimmermann, J. *et al.* (2015) 'Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies', *Molecular Ecology Resources*, 15(3), pp. 526–542. https://doi.org/10.1111/1755-0998.12336.

Zinger, L. *et al.* (2019) 'DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions', *Molecular Ecology*, 28(8), pp. 1857–1862. https://doi.org/10.1111/mec.15060.

Zizka, V.M.A., Weiss, M. and Leese, F. (2020) 'Can metabarcoding resolve intraspecific genetic diversity changes to environmental stressors? A test case using river macrozoobenthos', *Metabarcoding and Metagenomics*, 4, p. e51925. https://doi.org/10.3897/mbmg.4.51925.

CAPÍTULO II

COMUNIDAD DE ARTRÓPODOS GENÉTICAMENTE DIFERENCIADOS ENTRE LOCALIDADES DEL NEVADO DE TOLUCA

Fotografía de Miriam Campo: Bosque de *Abies religiosa* bajo conservación

PREFACIO

La evaluación integradora de los patrones de diversidad a diferentes escalas biológicas puede arrojar luz sobre la identidad y la fuerza de los procesos ecológicos y evolutivos que impulsan la distribución de la biodiversidad (Gómez-Rodríguez et al., 2019; McGill et al., 2019; Weber et al., 2017). Sin embargo, los patrones de diversidad normalmente se describen de manera aislada para cada uno de los niveles de biodiversidad, desde los genes hasta los ecosistemas. La evaluación integrada de patrones macroecológicos en varios niveles, que abarca genotipos, genealogías y especies (Baselga et al., 2013; Baselga, Gómez-Rodríguez, y Vogler, 2015) proporciona una nueva vía para explorar en un marco ecoevolutivo los impulsores potenciales de los patrones de biodiversidad. Este enfoque se basa en la teoría ecológica neutral, que postula que la estocasticidad de la dispersión es uno de los principales impulsores de la diversidad local y debería actuar de manera análoga en todas las escalas taxonómicas (Hubbell, 2001; Bell, 2001; Vellend, 2010; Rosindel et al., 2011). Los estudios empíricos que utilizan este enfoque hasta la fecha (por ejemplo, Baselga et al., 2013; 2015; Gómez-Rodríguez y Baselga, 2018; Arribas et al., 2020) han encontrado apoyo para la generalidad de los procesos neutrales, siendo las restricciones de dispersión un factor dominante impulsor de la variación espacial en la estructura de la comunidad en múltiples niveles jerárquicos (Baselga et al., 2013; 2015). Sin embargo, la generalidad de estos patrones observados requiere más datos empíricos, incluidas comunidades enteras (en lugar de unos pocos linajes seleccionados) muestreadas a través de escalas geográficas, taxones y regiones del mundo (Baselga et al., 2013; Arribas et al., 2020).

En este capítulo, nos enfocamos en una montaña tropical dentro de la Faja Volcánica Transmexicana para evaluar patrones espaciales en el ensamblaje de la comunidad de artrópodos a escala fina (<19 km) con el fin de comprender el papel de la limitación de la dispersión y las características del paisaje como impulsores de la diversidad. Recolectamos muestras de comunidades enteras de artrópodos para ocho órdenes a una escala espacial que va desde 50 m hasta 19 km, utilizando ADN de wcMBCm. Exploramos múltiples niveles jerárquicos, desde haplotipos individuales hasta linajes en umbrales de similitud de 0.5, 1.5, 3, 5, 7.5%, para evaluar patrones de riqueza, rotación y disminución de la distancia de similitud con respecto al aislamiento por distancia y aislamiento por resistencia (costos de dispersión dados por las características del paisaje).

Nuestros resultados revelaron que la distancia y la altitud influyen en la disminución de la distancia de similitud en todos los niveles jerárquicos. Este fenómeno es válido para grupos de artrópodos de capacidades de dispersión contrastantes, aunque la fuerza varía según la escala espacial. Nuestros hallazgos respaldan un modelo en el que la diferenciación a escala local mediada por restricciones de dispersión y combinada con la persistencia a largo plazo de los linajes, emerge como un impulsor significativo de la diversidad dentro de las islas de cielo de altas montañas.

Artículo: Dispersal limitations and long-term persistence drive differentiation from haplotypes to communities within a tropical sky-island: evidence from community metabarcoding. ***(Publicado)***

Received: 3 October 2020 | Revised: 9 September 2021 | Accepted: 13 September 2021

DOI: 10.1111/mec.16195

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Dispersal limitations and long-term persistence drive differentiation from haplotypes to communities within a tropical sky-island: Evidence from community metabarcoding

Nancy Gálvez-Reyes[1,2] | Paula Arribas[3] | Carmelo Andújar[3] | Brent C. Emerson[3] | Daniel Piñero[1] | Alicia Mastretta-Yanes[4,5]

[1]Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, CDMX, Mexico

[2]Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México, CDMX, Mexico

[3]Island Ecology and Evolution Research Group, Instituto de Productos Naturales y Agrobiología (IPNA-CSIC), Santa Cruz de Tenerife, Spain

[4]Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), CDMX, Mexico

[5]Consejo Nacional de Ciencia y Tecnología, Benito Juárez (CONACYT), CDMX, Mexico

Correspondence
Nancy Gálvez-Reyes and Daniel Piñero, Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad Universitaria, CDMX, México.
Email: nancygalvez@ecologia.unam.mx; pinero@unam.mx

Alicia Mastretta-Yanes, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), Liga Periférico Insurgentes Sur 4903, Col. Parques del Pedregal, Tlalpan, CDMX, México.
Email: amastretta@conabio.gob.mx

Funding information
Consejo Nacional de Ciencia y Tecnología, Grant/Award Number: 178245; Ministerio de Economía y Competitividad, Grant/Award Number: CGL2015-74178-JIN; Ministerio de Economía, Industria y Competitividad., Grant/Award Number: CGL2017-85718-P

Abstract

Neutral theory proposes that dispersal stochasticity is one of the main drivers of local diversity. Haplotypes-level genetic variation can now be efficiently sampled from across whole communities, thus making it possible to test neutral predictions from the genetic to species-level diversity, and higher. However, empirical data is still limited, with the few studies to date coming from temperate latitudes. Here, we focus on a tropical mountain within the Transmexican Volcanic Belt to evaluate spatially fine-scale patterns of arthropod community assembly to understand the role of dispersal limitation and landscape features as drivers of diversity. We sampled whole-communities of arthropods for eight orders at a spatial scale ranging from 50 m to 19 km, using whole community metabarcoding. We explored multiple hierarchical levels, from individual haplotypes to lineages at 0.5, 1.5, 3, 5, and 7.5% similarity thresholds, to evaluate patterns of richness, turnover, and distance decay of similarity with isolation-by-distance and isolation-by-resistance (costs to dispersal given by landscape features) approaches. Our results showed that distance and altitude influence distance decay of similarity at all hierarchical levels. This holds for arthropod groups of contrasting dispersal abilities, but with different strength depending on the spatial scale. Our results support a model where local-scale differentiation mediated by dispersal constraints, combined with long-term persistence of lineages, is an important driver of diversity within tropical sky islands.

KEYWORDS
arthropods, community metabarcoding, neutral theory, Nevado de Toluca, tropical mountains

[Correction added on 26 October 2022, after first online publication: Corresponding author 'Daniel Piñero' has been added to this version.]

in sampling blocks separated from 50 m to 19 km, and evaluated patterns of richness, turnover, and distance decay in community similarity at multiple hierarchical levels (haplotype, putative species and supra-specific levels). As we are interested in the effect of ecological and topographic features on dispersal limitation, in addition to testing for the effect of isolation-by-distance (IBD), we also performed isolation-by-resistance (IBR) analyses, incorporating costs to dispersal for landscape features such as altitude or habitat type (McRae, 2006). This method yields biologically more informative distance decay relationships than Euclidean distance alone (McRae, 2006). Our analyses allow us to test with empirical data and within a neutral framework, the suggestion that global patterns of hyperdiverse tropical mountains can be reinforced by differentiation of small isolated populations, combined with their long-term persistence (Bray & Bocak, 2016; Rahbek, Borregaard, Colwell, et al., 2019).

2 | MATERIALS AND METHODS

2.1 | Study area and bulk sampling

The study area comprises *Abies religiosa* forests which grow at around 2800–3500 m.a.s.l. in the Nevado de Toluca volcano, which is a sky-island in the TMBV (Mastretta-Yanes et al., 2015; Rzedowski, 2006). We sampled arthropods from 14 sampling points distributed across four sites: Tlacotepec (TLC; three sampling points), San Juan de las Huertas (SJH; three sampling points), San Bartolo (ASB; five sampling points) and Agua Bendita (AAB; three sampling points; Figure 1), all within the conservation zone of the natural protected area of Nevado de Toluca (Figure 1). Sampling was performed during the rainy season of 2015 during mid-August and September. Pitfall traps were active for 15 days, and both the placement and collection of all traps was undertaken within a 2-day period to reduce stochastic heterogeneity among samples due to temporal climatic variation. At each of the 14 sampling points, three sampling blocks of 20 × 15 m were established, resulting in a total of 42 community samples (Figure 1). Each sample consisted of all the specimens collected by 20 pitfall traps distributed equidistantly inside each sampling totalling 840 pitfall traps. Sampling blocks were separated by at least 50 m within each sampling point, and maximum distance among sampling points was 19 km (Figure 1). Pitfall traps consisted of plastic cups of 13 cm height and 10 cm diameter, with five 2 × 1.5 cm perforations, 2 cm apart from each other, at a height of 10 cm above the base. Cups were buried to the height of the window and lids were added to prevent the entry of rainwater. Cups were filled with a mixture of 185 ml of ethanol 70% and 15 ml of glycerin (Figure 1b; Figure S1a), which we previously tested was enough to prevent evaporation and thus DNA degradation. Pitfall traps of each block site (20 traps) were collected after 15 days, as suggested by Cardoso et al. (2008, 2009) and Emerson et al. (2017) and pooled in a single bulk-sample in a bottle containing ethanol at 96%. Sampling was performed with SEMARNAT permit No. SGPA/DGVS/02641/15.

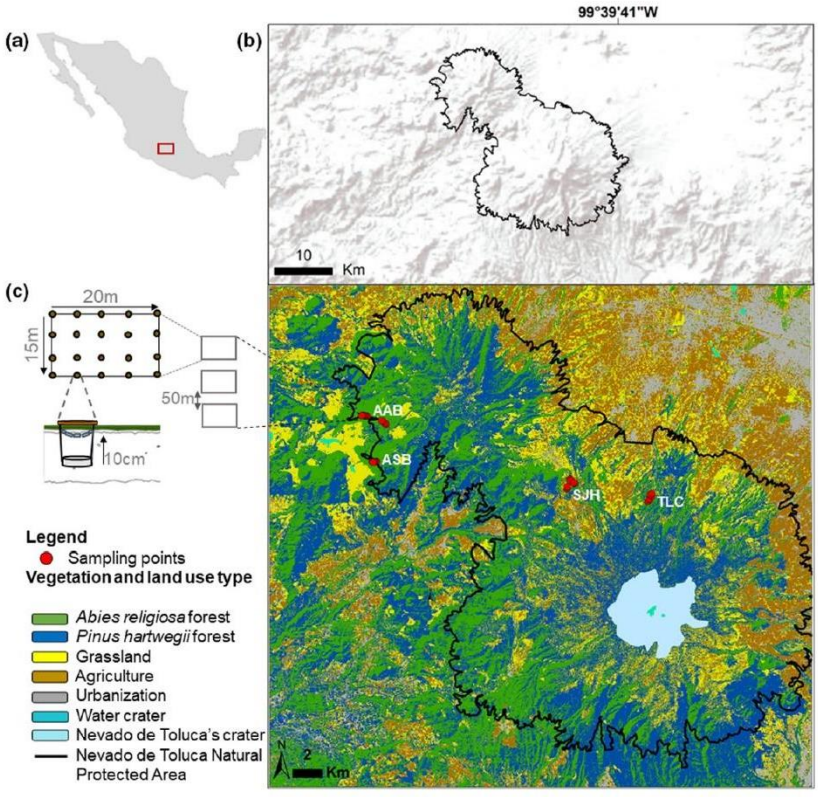

FIGURE 1 Study site and sampling design. (a) Location of study area within Mexico. (b) Nevado de Toluca volcano natural protected area over a topographic map. (c) At each sampling point (red circle) we sampled three sampling blocks of 20 × 15 m using 20 pitfall traps equidistantly distributed. Sampling blocks were separated by 50 m (within sampling points) to 19 km (between sampling points). Sampling points were distributed in four sites to the West (ABB: Agua Bendita and ASB: San Bartolo) and East (SJH: San Juan de las Huertas, and TCL: Tlacotepec) hillsides of Nevado de Toluca. Vegetation types are from González-Fernández et al. (2019), see text for details [Colour figure can be viewed at wileyonlinelibrary.com]

2.2 | Molecular laboratory processes

We cleaned up each sample (comprising 20 pooled pitfalls) following a Flotation-Filtration-Stereoscope protocol (FFS) that allowed us to have "clean" bulk specimens ready for DNA extraction (see Figure S1b–g). The samples had considerably different body sizes, so to prevent differences in biomass between specimens from causing biases at the DNA extraction and PCR steps (Elbrecht & Leese, 2015), we divided specimens by size class boundaries into: small (e.g., size *Drosophila melanogaster* <=3 mm), medium (e.g., size adult *Apis mellifera* >3 mm and <=15 mm), or large specimens (e.g., adult grasshopper >15 mm). We then divided each sample as follows: complete bodies for small arthropods, the thorax (head included) from medium-sized arthropods and two legs from large arthropods. Subsets of a few samples were weighed to determine that the head + thorax of medium sized individuals were comparable in weight to two legs of the larger size class. Samples ($n = 84$) were processed independently for DNA extraction and library construction but using the same library barcode identifier (thus, the final number of libraries retrieved was 42, one library per sample). For DNA extractions, we treated each bulk-sample with an electric homogenizer (Qiagen TissueLyser II: two times for 1 min at 30 Hz and 30 s at 30 Hz) using platinum beads with Qiagen DNeasy Blood & Tissue Kit and BSA buffer 1x.

For metabarcoding library construction, we used the double dual tagging method with the XT Illumina Adapter Kit, as in Arribas et al. (2016). We amplified a 418-base pair region from the 5′ end of mitochondrial COI gene (within the standard barcode region for metazoa) with the primers B_F 5′ CCIGAYATRGCITTYCCICG 3′ (Shokralla et al., 2015) and Fol-degen-R 5′ TANACYTCNGGRTGNCCRAARAAYCA 3′ (Yu et al., 2012) modified to include Illumina overhang adaptors for subsequent nested PCR. For each subsample ($n = 84$), we performed and pooled together equimolar amounts of three independent PCR replicates for each arthropod-size sample. We included a negative control reaction with no DNA template in all experiments. All information regarding PCR reagents and conditions is available in Appendix S1. Each pool of PCR amplicons was cleaned of remaining primers and primer dimers with Agencourt AMPure XP beads (Beckman Coulter) to purify the COI amplicon away from remaining primers and primer dimers. Then, amplicons were used as template for a limited-cycle PCR amplification to add dual-indices barcodes (N7 index and S5 index) and the P5 and P7 Illumina sequencing adapters using the Nextera XT Index Kit from Illumina. The 42 resulting metabarcoding libraries and a negative control were sequenced on a lane of Illumina MiSeq 2 × 300 bp at the Cornell Institute of Biotechnology, Cornell University, USA (Figures S2a–f).

2.3 | Bioinformatics processing to identify lineages different thresholds of genetic similarity

The resulting paired-end reads of the 42 samples were quality filtered following procedures described by Arribas et al. (2020). Briefly, processing included quality checking, primer removal, pair merging, quality filtering, denoising, and clustering each library independently. Raw reads were quality checked with fastqc (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Primers were trimmed using the fastx_trimmer option of fastx-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), trimming 21 and 27 bases for R1 and R2, respectively. Reads were then processed in trimmomatic-0.36 (Bolger et al., 2014) using TRAILING:20 to remove bases at the end of a read, if below a threshold quality of 20. R1 and R2 reads were used to search paired sequences with PAIRFQ-0.17 (Staton, 2019) and overlapping paired reads were merged with the fastq_mergepairs command, -fastq_minovlen 50 and -fastq_maxdiffs 15 in USEARCH-9.2 (Edgar, 2013). Quality-filtered (Maxee = 1), dereplicated (-fastx_uniques) and sorted (-sortbylength) options were used to keep only reads of 418 bp in USEARCH-10 (Edgar & Flyvbjerg, 2015). Surviving reads were denoised to generate amplicon sequence variants (ASVs; sensu Callahan et al., 2016), which are suggested to be predicted biological sequences to be used for direct analysis without the need of OTU clustering (Callahan et al., 2017; Edgar, 2016).

High-level taxonomic categories (order level) were then assigned to ASVs for each of the following orders: Diptera, Collembola, Arachnida, Coleoptera, Hymenoptera, Hemiptera, Myriapoda, Lepidoptera, using the lowest common ancestor (LCA) algorithm of MEGAN-6 (Huson et al., 2016). Taxonomic identification of each read was done using BLAST against the *nucleotide* NCBI *nt* database (6 June 2018; blastn -outfmt 5 -evalue 0.001). Remaining sequences were not further considered. The tree was exported, visualised and edited using FIGTREE-1.4.3 (Rambaut, 2012). Each ASV data set was aligned in Geneious-8.0.2 with MAFFT, using the FFT-NS-1 algorithm, a scoring matrix of 200/PAM/K=2, GAP open penalty of 3, and the translation align option. All sequences were inspected for insertions, deletions or stop codons disrupting the reading frame, which were then excluded.

A community table was then generated with read-counts (haplotype abundance) of each retained ASV for the eight orders by matching ASVs against the complete collection of reads (i.e., reads before the dereplicating and denoising steps) using the -search_exact command with USEARCH-10 (Edgar & Flyvbjerg, 2015). Additional filtering according to ASV abundances in community tables (one ASV per taxa) was performed as described by Arribas et al. (2020). To summarise, we first removed from each library those ASVs with abundances of four or fewer reads (same criteria as denoising). Second, within each library we removed ASVs that comprised less than 1% of reads within the taxonomic group that they belonged to. Although the recently released METAMATE software (Andújar et al., 2021) can provide for a more robust evaluation of filtering strategies, its application here was compromised by the limited availability of reference sequences matching the obtained ASVs (only five ASVs were classified as verified-authentic ASV). However, we were able to compare the efficiency of the 1% threshold used here with METAMATE, but only indirectly as METAMATE excludes ASVs only if they fall below a given threshold in all libraries (or binning groups) where they are found. Consequently, we used METAMATE to explore the effect of filtering by taxonomic group and library with thresholds of 0.8%, 1%, and 1.5%. These three thresholds removed respectively 79%, 83%, and 88% of

ASVs identified as verified non-authentic by METAMATE, indicating that our filtering appropriately reduces the prevalence of NUMTs and sequencing noise for the estimation of community-wide geographical patterns within the multihierarchical approach. Community tables of filtered haplotypes were then transformed into incidence (presence/absence) data and used for downstream analyses.

To identify lineages within each of the orders we use a range of clustering levels. Each of the ASV filtered data sets was used to generate an UPGMA tree with corrected distances under a F84 model, and based on this tree, filtered ASVs were considered haplotypes, which were then nested into lineages following the genetic similarity at different thresholds of genetic similarity (0.5%, 1.5%, 3%, 5% and 7.5%), plus an additional threshold derived from a species delimitation analyses conducted with the generalized mixed Yule-coalescent (GMYC) model in R (Pons et al., 2006). Analyses were performed using VEGAN (Oksanen et al., 2019), cluster, PMCMR, hier.part, ECODIST, and BETAPART (Baselga & Orme, 2012).

2.4 | Community diversity and composition

The eight arthropod orders subsets of ASVs were used to conduct analyses of community diversity and composition among sampling sites, considering haplotypes (raw ASVs) and all lineages (0.5%, 1.5%, 3%, 5% and 7.5% lineages, plus that corresponding to the GMYC species delimitation). Total accumulation richness curves were first estimated for multiple levels of genetic similarity (haplotypes, lineages at 3% and 5%) using the R package BETAPART (Baselga & Orme, 2012). ANOVAs with post-hoc.kruskal (Bonferroni method) were then used to test for significant differences in alpha diversity among communities.

Total beta diversity (Sorensen index, β_{sor}), additive turnover (Simpson index, β_{sim}; species replacement, without the effect of variation in richness) and nestedness (Sorensen–Simpson index, β_{sne}; pure richness effect) components were then estimated based on community compositions at different hierarchical levels. We used the R package VEGAN (Oksanen et al., 2019) and community composition matrices to perform nonparametric multidimensional scaling ordination (NMDS) based on Sorensen similarity. Plots were created with the ordispider option to visualise the compositional ordination of the communities among sites. An analysis of similarity (ANOSIM) was performed for each taxonomic group to compare arthropod community composition among sites. ANOSIM is a nonparametric method for analysing variance and testing multivariate differences between groups, based on a resemblance matrix and rank dissimilarity (Clarke, 1993). Plots were visualised with the R package GGPLOT2 (Wickham et al., 2020).

2.5 | Distance-decay of similarity and landscape connectivity

To examine the effect of landscape features on the variation of arthropods community composition, but considering the multihierarchical approach of Baselga et al. (2013), we tested for distance decay of similarity across the different lineages. We considered the following as independent variables: (i) geographic distance (i.e., IBD), and (ii) effective distance, that is, resistance to dispersion considering landscape features of slope, altitude, and vegetation type (i.e., IBR). We focused on Collembola, Diptera, Arachnida, Coleoptera, Hemiptera, and Hymenoptera assemblages for IBD, because they were the six sampled orders showing the highest completeness values, and in Collembola and Diptera for IBR because they demonstrate strong differences in dispersal potential (Figure S3). For response variables, pairwise similarity of assemblages at the haplotype and higher lineages (0.5%, 1.5%, 3%, 5%, 7.5% and GMYC) among all sites were used. The analysis was done with all six orders for the entire sampling area, but only with Diptera and Collembola at finer geographical distances within the West (AAB and ASB), and East (SJH, TLC) sampling points of our study area (Figure 1a). These orders were chosen because they have contrasting dispersal abilities. At each lineage threshold, patterns of similarity between pairs of sites were calculated using the means of the Simpson's similarity index (1 − pairwise β_{sim} = a/[a + min(b,c)] diversity) using the R package BETAPART (Baselga & Orme, 2012), where *a* is the number of species present in both territories, and *b* and *c* the number of species unique to one or another, respectively.

The program CIRCUITSCAPE-4.0 (McRae et al., 2013) was used to calculate "resistance distances" between pairs of nodes on a raster grid. Rasters of the study area were used as input files for circuitscape, in which each cell is assigned a conductance value corresponding to the relative probability of movement through it, together with a list of focal nodes - the geographic coordinates of the sampling blocks. To assess the extent to which community similarity in different arthropod groups can be explained by landscape features, conductance grids were used, representing different levels of resistance to dispersion depending on vegetation heterogeneity, altitude and slope. We developed 30 alternative "hypotheses" of resistance surfaces computed from a range of conductance matrices comprising: 10 alternative hypotheses (A-J) based on conductance values assigned to vegetation type, 12 hypotheses based on binary altitude thresholds, three hypotheses based on distribution of conductances around average sampling altitude and four hypotheses based on slope (Table S2). This approach has been extensively used to test how landscape features (present or past) drive genetic differentiation within species (e.g., Lee-Yaw et al., 2009; Mastretta-Yanes et al., 2018; Roffler et al., 2016). To compare the effect of these landscape variables against the effect of distance alone, we also generated a "flat" landscape; that is, a landscape in which all cells have equal conductance. This is equivalent to Euclidean distance but accounts for the finite size of the input landscape being analysed, and is thus more appropriate for comparison (Lee-Yaw et al., 2009).

For the vegetation type rasters, a high resolution digital land cover map of the Nevado of Toluca was used, based on satellite images (SPOT 6/7) of 2015 defining the *Abies religiosa* forest (sampled vegetation), *Pinus hartwegii* forest, alpine grassland, agriculture, urbanization, water in the crater and Nevado de Toluca's crater

(González-Fernández et al., 2019). After cross-validation with our sampling records, minor modifications were performed to adjust *Abies* forest distribution in areas of high topographic complexity (see Appendix S2 for details). Conductance grids were then constructed for the vegetation heterogeneity analyses assigning different conductance values to each vegetation type (maximum value of conductance is 1, meaning no resistance to dispersion) as detailed in Table S2.

To build the conductance grids for altitude and slope, a similar approach was followed, assigning different conductance values to altitudinal and slope ranges. In total 30 conductance grids were tested (Table S2; Figure S5). Additionally, distance decay of similarity was tested for at smaller geographic distances, for which West (AAB, ASB) and East (SJH, TLC) sites were analysed separately, but using only the "flat" raster and the raster from the previous analysis with the highest explanatory power.

A negative exponential function "decay.model" in the R package BETAPART (Baselga & Orme, 2012) was used to adjust a negative exponential function to a generalized linear model (GML). We used Simpson similarity ($1 - \beta_{sim}$; Baselga, 2010) as a response variable, pairwise effective distances of each resistance surface as predictors, log link and Gaussian error (Arribas et al., 2020; Gómez-Rodríguez & Baselga, 2018). Finally, we evaluated the fractal pattern (i.e., self-similar systems; Baselga et al., 2015) by a log–log Pearson correlation of the haplotype level and, independently: (1) number of lineages, (2) initial similarity (i.e., intercept), and (3) mean similarity. This was undertaken for each of the six orders across the sampling area, and for Collembola and Diptera at the East and West sections. High correlation values are indicative of self-similarity in lineage branching (i.e., number of lineages) and/or spatial geometry of lineage distributional ranges (i.e., initial and mean similarity; Baselga et al., 2015), supporting the presence of fractal patterns, which are predicted under a neutral process of community evolution (Arribas et al., 2020). Analyses and graphical representations of data were performed with the R packages VEGAN (Oksanen et al., 2019) BETAPART (Baselga et al., 2018) and ECODIST (Goslee & Urban, 2020).

3 | RESULTS

3.1 | Phylogenetic groups and ASVs recovered by COI

Across the 42 libraries, MiSeq sequencing generated a total of 11,639,999 paired reads. We obtained from 108,583 to 419,903 reads in each direction for each sample. Of these, from 39,484 to 178,720 sequences remained after quality filtering (totalling 4,990,334). After read merging and sequence filtering to a length of 418 bp, each sample comprised between 33,609 and 155,634 sequences (totalling 4,305,390). Taxonomic assignments with *usearch* showed high similarity to a broad range of arthropod species (Figure 2a). The 4,305,390 sequences comprised 1,277 ASVs (unique variants: Diptera, 385; Collembola, 270; Arachnida, 155; Coleoptera, 136; Hymenoptera, 133; Hemiptera, 116; Myriapoda, 51; Lepidoptera, 31) (Figure 2b). The GMYC threshold values obtained were: Diptera, 0.9%; Collembola, 2.9%; Arachnida, 1.3%; Coleoptera, 0.7%; Hymenoptera, 1%; Hemiptera, 1.8%; Myriapoda, 1.5%; Lepidoptera 1.5% (Table S1).

3.2 | Arthropod ASVs richness at multihierarchical levels

Rarefaction curves indicate that sampling effort was sufficient to achieve from 60 to 91% completeness at the levels of individual haplotypes and lineages at 3%, for four of the eight taxa (Diptera, Collembola, Arachnida and Hymenoptera Figure S3). The groups that yielded higher richness within communities (alpha diversity by sample) were Diptera (mean = 87 SD ± 23.1 haplotypes and mean = 21 ± 5.15 lineages at 3%), Collembola (mean = 38 ± 9.44 haplotypes and mean = 16 ± 2.91 lineages at 3%), and Arachnida (mean = 12 ± 6.28 haplotypes and mean = 9 ± 3.64 lineages at 3%). Diptera, Collembola, Arachnida, Hymenoptera and Coleoptera presented significant differences in richness among sampling sites but patterns were not consistent across groups (Figure 3).

3.3 | Community composition at multihierarchical levels

The overall turnover (β_{sim}) among sampling points was high for all the arthropod groups across all levels analysed (close to 1, β_{sim} haplotypes from 0.904 to 0.957; β_{sim} 3% from 0.896 to 0.945; Figure S4) and average pairwise β_{sne} was close to zero, β_{sne} haplotypes from 0.010 to 0.026; β_{sne} 3% from 0.014 to 0.033. The mean value for β_{sim} for haplotypes of Coleoptera (n = 136) was 0.957, the average pairwise β_{sne} was 0.014, and β_{total} was 0.970 (β_{sor}). For Diptera (n = 385), the mean value for β_{sim} was 0.904, the average pairwise β_{sne} was 0.024, and the average pairwise β_{sor} was 0.928.

NMDS ordination plots revealed differences in community composition among sampling sites, particularly among those located in the West (AAB and ASB) and East (SJH and TLC) sampling points within Nevado de Toluca (Figure 4). ANOSIM revealed that differences of dissimilarity were significant for the different groups and lineages at 3 and 5%, except on Myriapoda (Figure 4). The orders Diptera, Collembola, Hemiptera and Coleoptera each formed two geographic groups: (i) AAB-ASB (West) and (ii) SJH-TLC (East). In contrast, three groups were recovered within Arachnida: (i) AAB, (ii) ASB and (iii) SJH-TLC (Figure 4). The variation in Collembola, Arachnida and Diptera among sites largely contributed to the ordinations of community similarity among sites. The highest values were observed in Collembola (Haplotype $r^2 = 0.914$, $p < .001$; lineages at 3% $r^2 = 0.826$, $p < .001$), followed by Diptera (Haplotype $r^2 = 0.595$, $p < .001$; lineages at 3% $r^2 = 0.372$, $p < .001$) and Arachnida (Haplotype $r^2 = 0.569$, $p < .001$; lineages at 3% $r^2 = 0.452$, $p < .001$; Figure 4).

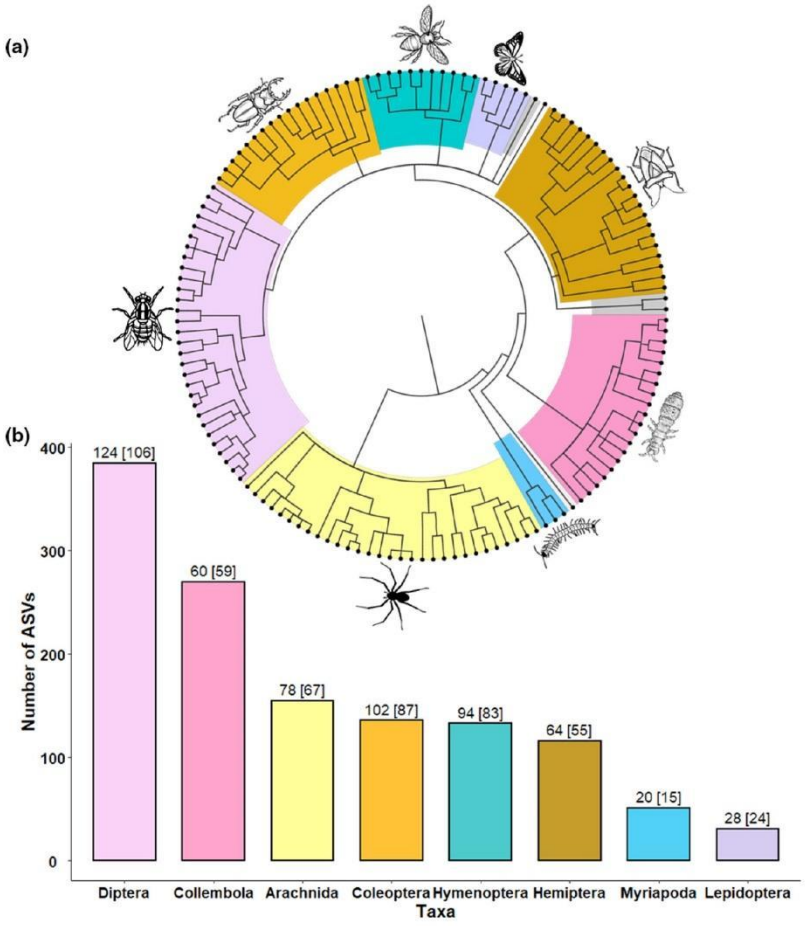

FIGURE 2 Amplicon sequence variant (ASV) diversity at Nevado de Toluca *Abies* forests. (a) Phylogenetic tree constructed by taxonomic composition of the arthropods estimated with the lowest common ancestor (LCA) algorithm on MEGAN. (b) Number of ASVs recovered at haplotype level for each order. Numbers above the bar show putative species by GMYC and values inside of brackets are lineages at 3% [Colour figure can be viewed at wileyonlinelibrary.com]

3.4 | Landscape community connectivity

Across all sampling sites (max distance 19 km), pairwise similarity within communities at all levels of clustering decreased with Euclidean ("flat") distance (Figure 5ac, Figure S7). The fit of IBD was higher in Collembola (from $r^2 = 0.704$, $b = -2.07$, $p < .001$ at the haplotype level, to $r^2 = 0.580$, $b = -0.92$, $p < .001$ at 7.5% lineages; Table S3; Figure 5a) than Diptera (from $r^2 = 0.293$, $b = -0.56$, $p < 0.001$ at the haplotype level, to $r^2 = 0.036$, $b = -0.14$, $p < .001$ lineages at 7.5%; Table S4; Figure 5c). Similar IBD results were found for Arachnida, Coleoptera, Hemiptera and Hymenoptera (Figure S7, Table S7). However, variation was better explained by IBR (Figure 5bd; Figure S5). For Collembola, the explanatory power of the resistance surface "Altitude 3,000" was slightly higher than the "flat surface" at different lineages (from $r^2 = 0.723$, $b = -1.63$, $p < .001$ at the haplotype level, $b = -0.71$, $p < .001$ at 7.5% lineages; Table S3; Figure 5b). In Diptera the highest explanatory power was given by the resistance surface "Mean-elevation peak (symmetrical)" (from $r^2 = 0.286$, $b = -0.15$, $p < .001$ at the haplotype level, $b = -0.05$, $p < .001$ at 7.5% lineages; Table S4, Figure 5d).

Distance-decay relationships (DDR) at a finer geographic scale within the West and East regions remained strong in Collembola, but were weaker or nonsignificant in Diptera (Figure 6). East sampling points (<5 km), also showed distance decay of similarity in Collembola (Table S5, Figure 6a), and Diptera, but coefficients were lower in the second (Table S5, Figure 6c). When considering IBR analysis, very similar results were found (Table S5, Figure S6a,c). In the West section of our sampling (<2 km), a similar pattern to the East was found for Collembola (Table S5, Figure 6b) but Diptera showed nonsignificant results both for the IBD and IBR analyses (Table S5, Figure 6d, Figure S6d). Exponential decay curves yielded higher slopes for Collembola than Diptera. All lineages of Diptera and Collembola show the same pattern of distance decay, but each lineage shows equal or greater pairwise similarity than at the previous one (Figures 5 and 6; Table S3 and S4). Regarding the multihierarchical analysis, the distance decay at the haplotype level showed a significant log–log correlation with the number of lineages, initial similarity, and mean similarity of communities. This is true for both our entire sampling area and focusing on the West and East areas (Table S6). The log-log linear correlations suggest that the patterns of assemblage variation across hierarchical levels can be described by a fractal geometry (Baselga et al., 2013, 2015). Thus, we found that community variation across genetic similarity levels can be described by a fractal geometry in each group in all the geographic scales sampled.

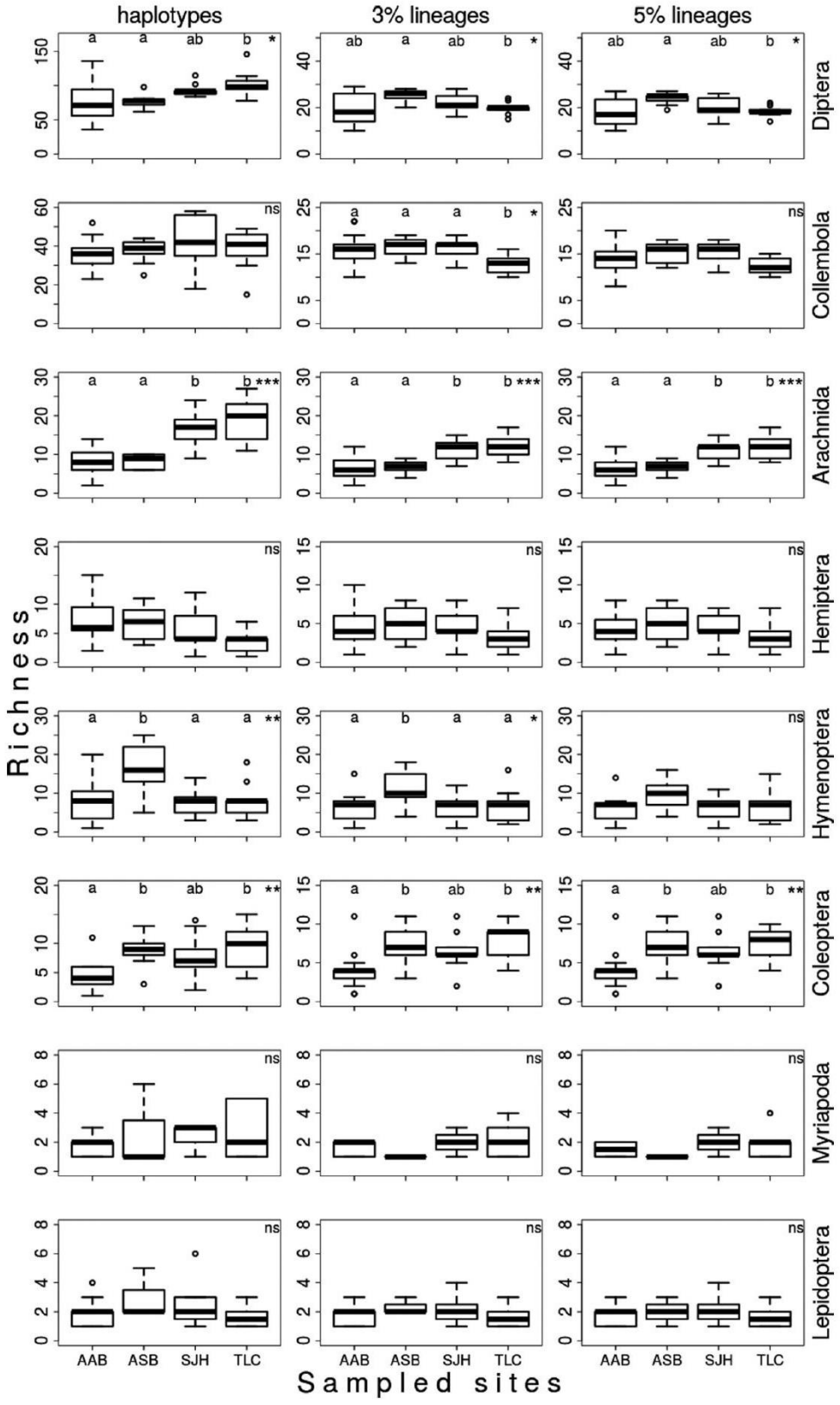

FIGURE 3 Arthropods richness at different clustering levels by sampling sites. AAB, Agua Bendita; ASB, San Bartolo; SJH, San Juan de las Huertas; TLC, Santiago Tlacotepec. Same letters indicate that there are no significative differences among those sites $p < .05$. Significance codes *$p < .05$, **$p < .01$, ***$p < .001$, ns, nonsignificant

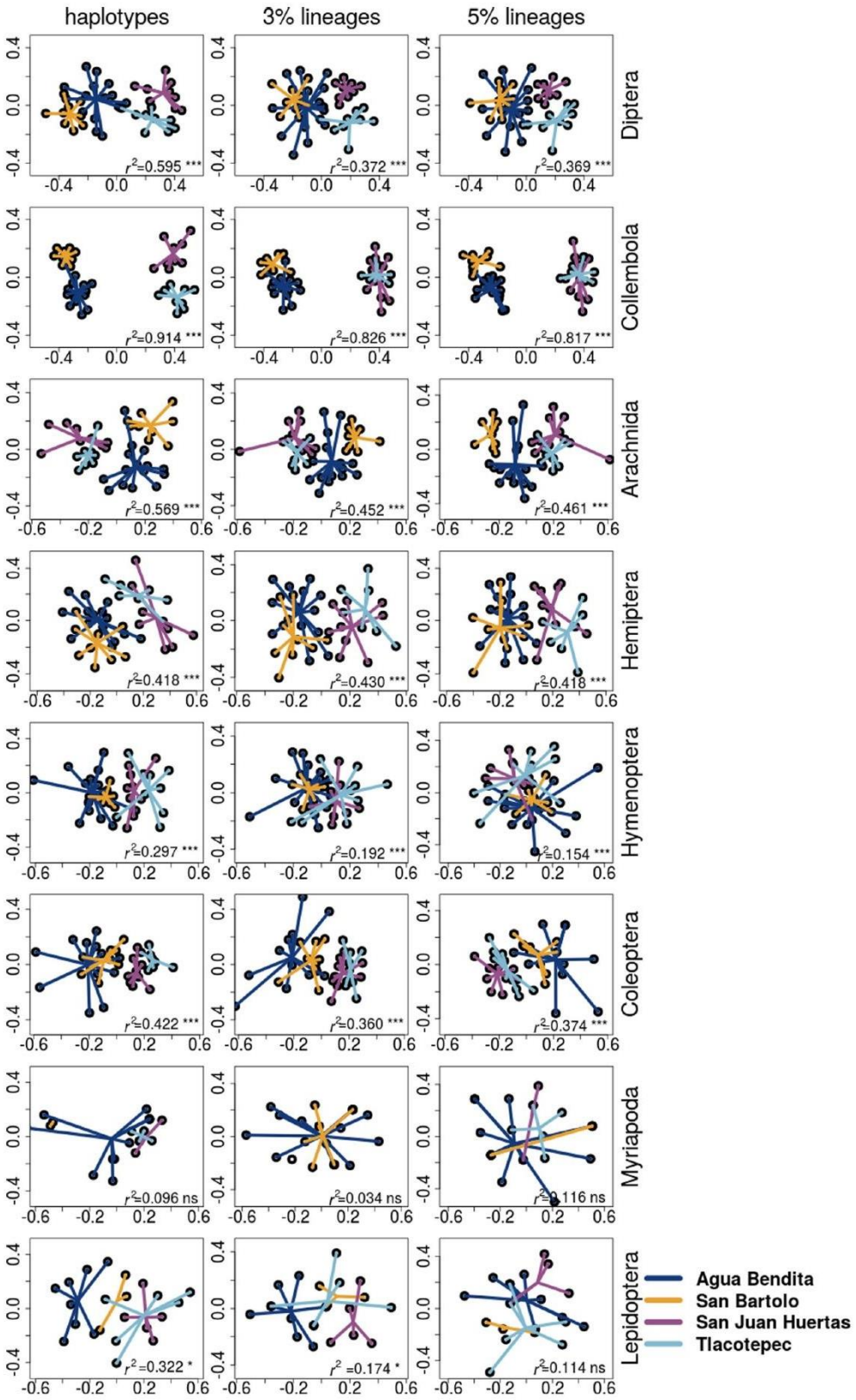
haplotypes
3% lineages
5% lineages
Diptera
Collembola
Arachnida
Hemiptera
Hymenoptera
Coleoptera
Myriapoda
Lepidoptera
r²=0.595 ***
r²=0.372 ***
r²=0.369 ***
r²=0.914 ***
r²=0.826 ***
r²=0.817 ***
r²=0.569 ***
r²=0.452 ***
r²=0.461 ***
r²=0.418 ***
r²=0.430 ***
r²=0.418 ***
r²=0.297 ***
r²=0.192 ***
r²=0.154 ***
r²=0.422 ***
r²=0.360 ***
r²=0.374 ***
r²=0.096 ns
r²=0.034 ns
r²=0.116 ns
r²=0.322 *
r² =0.174 *
r²=0.114 ns
Agua Bendita
San Bartolo
San Juan Huertas
Tlacotepec

FIGURE 4 Nonmetric multidimensional scaling (NMDS) ordinations of community similarity (Simpson index, βsim) at the haplotype, 3% and 5% lineages for each of the eight taxonomic orders studied. The asterisk represents significant differences of community structure among sites estimated with ANOSIM at $*p < .05$, $**p < .01$ $***p < .001$ and ns, nonsignificant values [Colour figure can be viewed at wileyonlinelibrary.com]

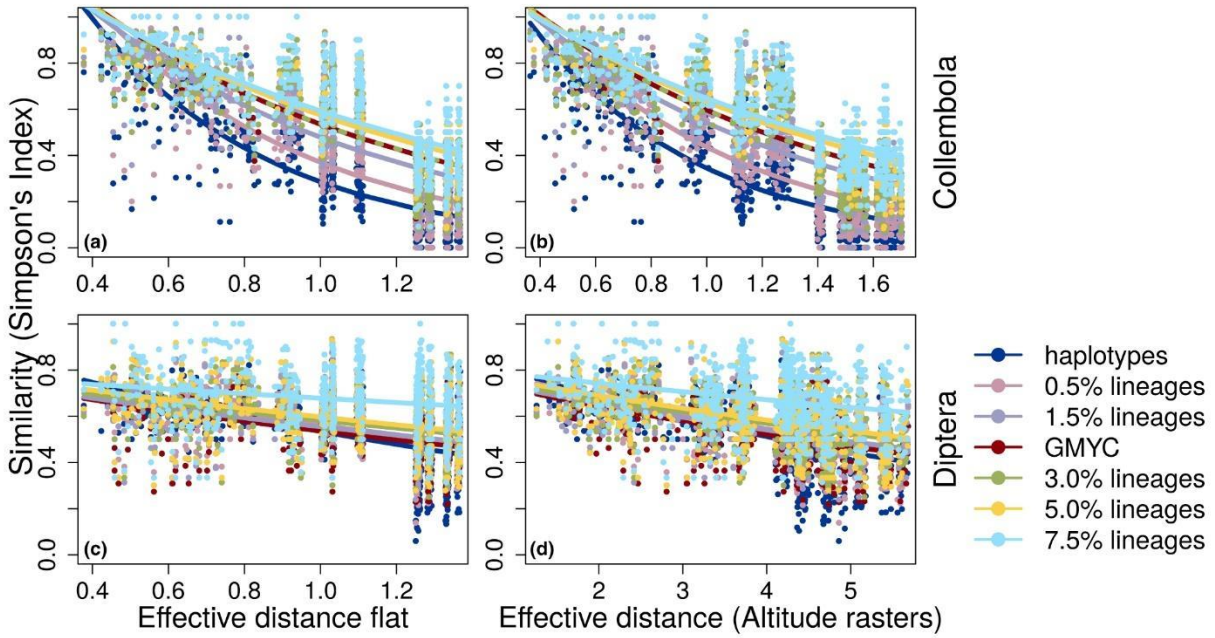

FIGURE 5 Distance decay of community similarity at Nevado de Toluca for Collembola (a, b) and Diptera (c, d) at multiple levels of genetic similarity. Decay of similarity is shown against flat effective distance ("flat" surface; a, c) and against effective distance using the resistance surface "Altitude 3,000" for Collembola (b) and "Mean-elevation peak (symmetrical)" for Diptera (c), which were the ones showing the highest explanatory power (significance levels in Table S3). Dots represent similarity between pairs of sampling sites and lines are the fitted model for each lineage. Altitude rasters as in Figure S5. Significance levels and r^2 are in Table S3 and S4. Notice that effective distances are resistances and hence are not comparable to km [Colour figure can be viewed at wileyonlinelibrary.com]

4 | DISCUSSION

We recovered 1,277 ASVs, 496 lineages at 3%, 441 at 5% and 570 using GMYC across the eight arthropod orders. Composition of arthropod communities exhibited high turnover among sampling blocks and ANOSIM tests also showed significant differences among sites within blocks. Across all sampling points, spanning a maximum distance of only 19 km, we found that distance (pure geographical distance and corrected by elevation) is a key variable explaining community structure, from the level of haplotypes through to lineages, suggesting that dispersal limitations drive community structure across multihierarchical levels within a single sky-island. This pattern holds at finer geographic distances (<2 km), but only on it up with considerably low dispersal ability (nonwinged Collembola).

4.1 | Arthropod communities recovered by metabarcoding and sampling blocks

The estimation of species richness in any ecological setting, and especially in forested environments, can be challenging due to the rarity of some species, differences in detection probabilities, and the field effort necessary to collect enough samples or species to ensure meaningful coverage (Andújar et al., 2017; Arribas et al., 2016; Creedy et al., 2019). In our study, we used pitfall traps in sampling blocks, maximizing the probability of detecting arthropod species by sampling intensively at multiple sites in one mountain, covering eight orders of arthropods, and characterising these from haplotypes to communities. Our sampling method and size sorting step allowed the recovery of eight major arthropod orders, congruently with other metabarcoding analyses (Creedy et al., 2019; Elbrecht et al., 2017, 2018). According to our rarefaction curves, our sampling detected different taxa per sample (Figure S3), demonstrating the utility of cMBC and our sampling design to study a region of high biological diversity and ecological complexity. It is difficult to compare our results against morphological studies in Nevado de Toluca because there are no complete checklists of arthropods for Mexican highlands. For instance, out of 29 dung beetle species found by a recent survey of four sky-islands, more than 10% were new species (Arriaga-Jiménez et al., 2018).

The metabarcoding pipeline we implemented (Arribas et al., 2020) allowed us to analyse diversity patterns for each order

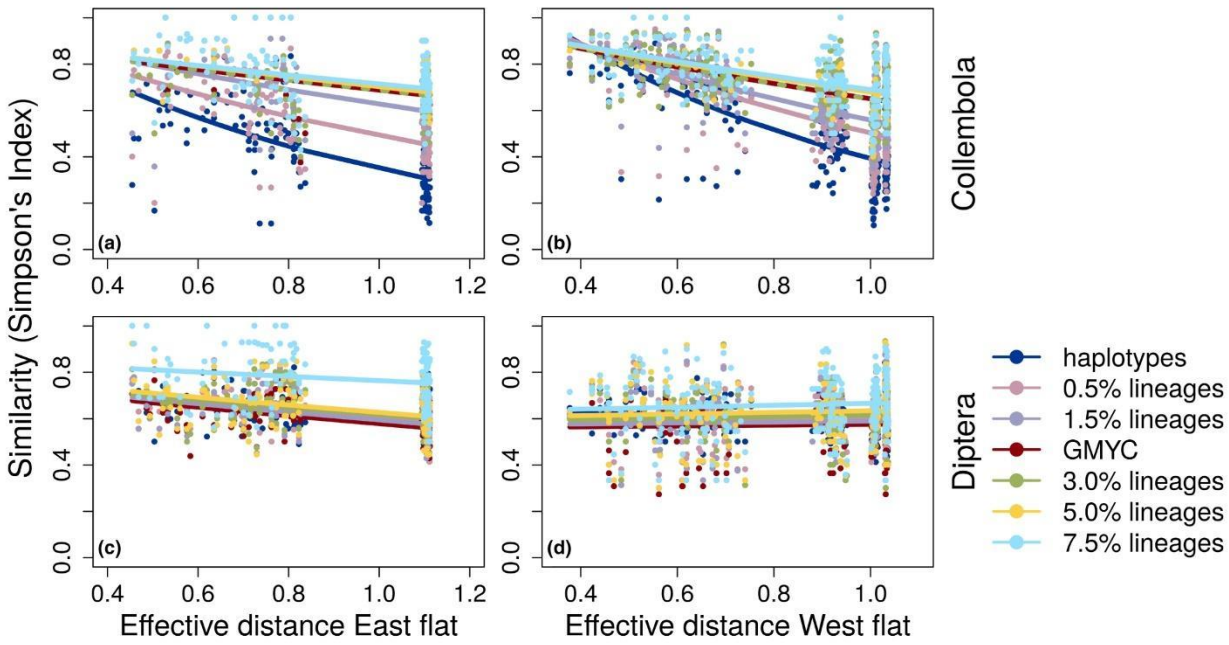

FIGURE 6 Distance decay of community similarity at fine geographic distances within Nevado de Toluca for Collembola (a, b) and Diptera (c, d). Analysis at <5 km (a, c) using the Eastern subset of sampling sites (SJH and TLC) and at <2 km (b, d) using the Western subset (AAB and ASB). Decay of similarity is shown against geographic distance ("flat" surface). Dots represent pairs of sampling sites for each clustering level and lines are the fitted model. Flat rasters are in Figure S5. Significance levels and r^2 are in Table S5 [Colour figure can be viewed at wileyonlinelibrary.com]

separately and to analyse community assembly within a multihierarchical framework, from individual haplotypes to lineage divergence at 7.5%. Our approach identified 1,277 ASVs from 42 bulk samples, allowing us to estimate community composition and turnover without the bias introduced by traditional taxonomy (Creedy et al., 2019), including small taxa (<0.5 mm), such as Collembola, and specimens that break easily, such as Diptera (Figure 3). Although the cMBC method allowed us to explore a large data set of hundreds of pitfall traps, there are potential limitations which may introduce bias such as the use of universal primers for PCR (Piñol et al., 2015), biomass differences (Elbrecht et al., 2017), and/or tag jumping (Schnell et al., 2015). We have attempted to reduce potential bias by conducting PCRs in triplicate, dividing our specimens by size classes, and by using HLD desalted primers to decrease tag jumping.

4.2 | Strong turnover of arthropods communities at multihierarchical levels

Our study provides haplotype-level and higher lineages data of entire communities (Figure 4, Figure S4). High beta diversity was found across communities in all orders, and was dominated by lineages turnover (β_{sim}) instead of nestedness (β_{sne}), from haplotypes to higher hierarchical levels (Figure S4). We found significant differentiation among sites and vegetation type (*Abies* forests) according to site haplotype composition, and this differentiation persisted when molecular entities that conservatively represent species were considered. Diptera and Collembola presented the highest differentiation among sites, specifically dividing the East (SJH-TLC) and the West (AAB-ASB) sampling points, as shown by the NMDS and ANOSIM analyses (Figure 4). Again, this occurs from the level of haplotype to higher lineages (Figure 4). The West-East division in community structure within Nevado de Toluca (Figure 4) corresponds to opposing hillsides. In mountain landscapes, East-facing slopes with morning sun may provide different conditions from cold and foggy West-facing slopes (Rahbek, Borregaard, Colwell, et al., 2019), which could influence community assembly. While we cannot discard the role of environmental heterogeneity (environmental distance) driving community differences between western-eastern sampling points, it should be emphasized that all sampling points were located within *A. religiosa* forests with no apparent differentiation within the Nevado de Toluca. In fact, *A. religiosa* forests are expected to grow under similar conditions within a single mountain, having a quite restricted environmental niche associated with moist and cold sites (Rzendowski, 2006).

4.3 | Dispersal limitations drive community structure across multihierarchical levels at fine geographical scales

To investigate whether dispersal limitations shape distance-decay within a mountain, we examined DDR at multihierarchical levels, including haplotypes which are expected to behave neutrally (but see

Gerber et al., 2001), even across environmental gradients. IBD analyses were conducted for six orders, and IBR analyses were conducted for Diptera and Collembola (which have contrasting dispersal abilities) using 30 conductance matrices. For Diptera and Collembola, we considered our entire sampling in Nevado de Toluca (19 km max distance among sampling points) and finer geographic scales within the East (<5 km) and West (<2 km) subsets of our sampling. We found that decay of similarity of communities decreases with spatial distance at the level of haplotypes, at all levels of lineage divergence (from 0.5 to 7.5%), and when haplotypes were segregated to putative molecular species using GMYC (Figure 5). Distance decay was high in all orders (Figure 6, Figure S7), but it is more marked in groups with more limited dispersal abilities (e.g., Collembola and Arachnida) than in groups with higher dispersal abilities (e.g., Diptera, Coleoptera, Hemiptera and Hymenoptera). Interestingly, for Collembola our results hold both considering our entire sampling as well as finer (<2 km) geographic distances (Figures 5 and 6), which is consistent with genetic studies within Collembola showing genetic differentiation over very short geographic distances (Arribas et al., 2020; Cicconardi et al., 2013).

High dispersal ability is expected to enhance community similarity (Baselga et al., 2012). Our results support this, as distance decay is higher in the wingless Collembola ($r^2 = 0.704$ and $r^2 = 0.599$ at the haplotype and GMYC levels, respectively; Table S3; Figure 5a) than in the winged Diptera ($r^2 = 0.293$ and $r^2 = 0.195$ at the haplotype and GMYC levels, respectively; Table S4; Figure 5c). Similar patterns of higher distance decay relationships at multihierarchical levels in poorly dispersing organisms than in better dispersers have been found in European water beetles (Baselga et al., 2013), Iberian leaf beetles (Baselga et al., 2015) and European beetles (Gómez-Rodríguez & Baselga, 2018) at much larger (hundreds of km) geographical scales than here (but see Gómez-Rodríguez et al., 2019 where the pattern was not clear for terrestrial molluscs), and only recently even at the scale of a few kilometres, in Collembola, Acari and Coleoptera (Arribas et al., 2020). Communities of good dispersers are more homogeneous not only because they can disperse larger distances, but also because they can more easily overcome geographical barriers between suitable habitat (Thompson & Townsend, 2006; Vellend, 2010). Thus, as dispersal abilities become limited, landscape features impeding dispersal are expected to play a more important role in structuring diversity, which can be explicitly tested including landscape features in analyses such as IBR.

IBR quantifies "effective distances" between communities that may yield more biologically informative DDR than Euclidean distance (McRae, 2006; McRae et al., 2008). We found that altitudinal differences better explain similarity decay than distance alone ("flat" landscape), slope or vegetation type. The resistance surface "flat" (i.e., IBD) has slightly less explanatory power for Collembola ($r^2 = 0.704$, $p < .001$ at the haplotype level; Table S3; Figure 5a) than "Altitude 3,000" ($r^2 = 0.723$, $p < .001$ at the haplotype level; Table S3; Figure 5b), the best fitting resistance surface. Altitude 3,000 resistance surface corresponds to the elevation at which the Nevado de Toluca volcano massif begins (Figure S5; Table S2), suggesting that Collembola followed a pattern of IBD and that within their limited dispersal ability, landscape features do not represent an impediment. For Diptera, the highest explanatory power was provided by the resistance surface "mean-elevation peak (symmetrical)" ($r^2 = 0.286$, $p < .001$ at the haplotype level; Table S4; Figure 5d). This resistance surface assumes maximum conductance at the mean altitude of our sampling and a gradual decrease until reaching altitudes outside of our sampling range, but still where *Abies* forest can be found (Table S2). This suggests that even though they are good dispersers, the efficiency of Diptera movement through relatively unsuitable conditions (different altitudes) is compromised. We thus find that for Diptera, it is not distance alone that drives community structure, but also landscape features. Although our sampling blocks are separated by short distances from 50 m to 19 km, effective distances among sites depends upon the elevation model used to set the conductance values (Table S2; Figure S5). This is consistent with Janzen's prediction of "mountain passes being higher in the tropics" (Janzen, 1967), and adds to the recent empirical data corroborating it (Polato et al., 2018). However, although our results show that landscape connectivity contributes to dispersal limitation, geographic distance seems to play a more dominant role both for both orders. This is consistent with dispersal limitation acting over evolutionary time, as has been suggested to explain the small spatial scale diversification of *Scarelus* beetles within tropical mountains (Bray & Bocak, 2016).

Distance decay patterns at the species level could reflect spatially correlated environmental heterogeneity (i.e., between western and eastern sides of the Nevado de Toluca). While some degree of environmental distance could explain biodiversity patterns (but see above on the homogeneity of the sampling study habitat), the following findings point to dispersal limitation within this single single sky-island as a major driver of community assembly: (i) spatial patterns of community dissimilarity are recurrently found for multiple hierarchical levels; (ii) high values of turnover and local endemicity occur at multiple spatial scales; and (iii) there is a consistent multihierarchical pattern of distance decay in community similarity at even small spatial scales for less dispersive species, and to a lesser extent for more dispersive species.

Multihierarchical approaches are useful to assess whether variation in biological assemblages driven by dispersion follow a fractal geometry where clumping arises at all spatial scales and hence the distribution of haplotypes and higher lineages is self-similar (Baselga et al., 2013, 2015; Gómez-Rodríguez et al., 2019). Our results also reveal the existence of a fractal pattern for DDR, with similarity decreasing with spatial distance from the haplotype level to 7.5% lineages (Table S6). These patterns represent a considerably finer geographic scale than that reported in previous studies (from 820 km to 4,500 km as in Baselga et al., 2013, 2015). DDR decreased with distance at even finer geographic scales (<5 and <2 km) in Collembola (Table S5, Figure 6 and S6), revealing that for arthropods with low dispersal ability, DDR can occur at very fine geographic distances and at all multihierarchical levels, within the geographic confines of a sky island. Given these short distances, our findings

are important not only for understanding evolution, but also for biomonitoring efforts aiming to detect changes in community assembly, even in relatively short distances among sampling points.

4.4 | Implications for evolution in insular systems and tropical mountains

We found that dispersal limitations drive community structure across multihierarchical levels within a single habitat in a geographically limited sky-island. This is congruent with analyses in oceanic islands, showing that when dispersal ability and climate tolerance are restricted, strong geographic isolation within an island can occur even in a few kms but extending back even millions of years (Salces-Castellano et al., 2020). Our results thus represent an additional source of evidence of how topography and dispersal limitations can interact to promote community wide diversification, even at a very limited spatial scale.

Dispersal constraints are expected to be more pronounced in the tropics than temperate areas, as tropical species have typically narrower thermal tolerances and lower dispersal than temperate species, leading to higher isolation-by-distance and isolation-by-elevation (e.g., Polato et al., 2018). Nevado de Toluca is part of a tropical sky-islands archipelago, and considering this broader spatial context has interesting implications for why tropical mountains are biodiversity hotspots. It has been hypothesized that the global pattern of hyperdiverse tropical mountains probably reflects the differentiation of small, spatially isolated populations combined with the long-term maintenance of these populations, leading to speciation (Rahbek, Borregaard, Colwell, et al., 2019). In this context, spatial isolation normally refers to habitat fragments distributed across different mountain peaks (Fjeldså et al., 2012; Rahbek, Borregaard, Antonelli, et al., 2019; Rahbek, Borregaard, Colwell, et al., 2019). However, our results suggest that the processes of differentiation and long-term persistence of small populations may also hold at local scales, within sky islands. First, our data shows that a single sky-island harbours arthropod communities that are spatially structured, at the haplotype and lineage levels, even within a single type of forest with presumably similar environmental conditions, even at short geographic distances. Second, it has previously been shown that montane ecosystems within the TMVB (including Nevado de Toluca) are able to persist within the same mountain during climate fluctuations (Mastretta-Yanes et al., 2018). Coupling these results together points to a model where a single sky-island can act as a cradle for population differentiation and that this differentiation can persist, and accumulate, relatively in situ over evolutionary time scales. Previous single-species case studies on poor dispersing beetle taxa have reached similar conclusions (Bray & Bocak, 2016), but our multiorder community-level data (Figure 5, Figure S7) shows that rather than a particular case restricted to few poor dispersing taxa, the phenomenon could be widespread among montane tropical arthropods.

5 | CONCLUSIONS

Our results provide strong empirical support for the suggestion that global patterns of hyperdiverse tropical mountains can be reinforced by differentiation of small isolated populations, combined with their long-term persistence (Bray & Bocak, 2016; Rahbek, Borregaard, Colwell, et al., 2019). Our work shifts the focus from population isolation among different mountains, as we find that arthropod communities can have strong turnover at the intraspecific level within a sky-island, and at a limited geographic scale (<20 km). Distance and elevation are found to be strong drivers of biodiversity structure, from the level of individual haplotypes through to presumed species and higher levels. This pattern holds even at a very fine spatial scale (<2 km) for taxa for which inherent dispersal ability is low, being exemplified by Collembola. Thus, our results support a general model where dispersal limitations act as a source of local-scale genetic differentiation within tropical mountains, which may translate to species-level diversification over time.

ACKNOWLEDGEMENTS

We are thankful to A. Sunny and A. González for providing the raster of vegetation types, to D. Ortiz and W. Tobon for help with spatial data transformations and to E. Campos for IT support. Nancy Gálvez-Reyes is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and was supported by CONACYT (scholarship no. 401781). CONACYT also provided financial support with the project 178245 granted to D. Piñero. B. Emerson was supported by the Spanish Agencia Estatal de Investigación (CGL2017-85718-P), co-financed by FEDER. P. Arribas was supported by a postdoctoral grant by the Spanish Ministry of Economy and Competitiveness (MINECO, Spain) within the Juan de la Cierva Formación Program. C. Andújar was supported by the Spanish Ministry of Economy and Competitiveness (MINECO, Spain) (CGL2015-74178-JIN MINECO/FEDER, UE). Analyses were carried out using CONABIO's computing cluster with the assistance of E. Campos. Thanks to Santiago Tlacotepec and Amanalco communities for allowing us to sample in their lands, to C.C.M.S.S., A. Silva, M. Campos, V. Reyes, A. Guerra, I. Rojas, G. H. Giles, A. Villarruel, M. Garduño, J. Jiménez and J. C. García for sampling support, T. Garrido, M. Soto and A. Domíngez for lab support and to F. S. Gálvez for assistance with pitfall traps.

AUTHOR CONTRIBUTIONS

Alicia Mastretta-Yanes, Nancy Gálvez-Reyes, Brent C. Emerson, and Daniel Piñero conceived and designed the study. Nancy Gálvez-Reyes, Alicia Mastretta-Yanes and Daniel Piñero performed fieldwork, Nancy Gálvez-Reyes performed laboratory work, processed metabarcoding data and made analyses. Paula Arribas, Carmelo Andújar, and Brent C. Emerson supervised analyses and contributed to the discussion. Manuscript writing was led by Nancy Gálvez-Reyes and Alicia Mastretta-Yanes, with contributions from all authors. All authors approved the final version of the manuscript.

DATA AVAILABILITY STATEMENT

ASVs tables and aligned sequences of this project have been deposited at Dryad repository https://doi.org/10.5061/dryad.wh70rxwkw. Scripts and sampling metadata used for analyses are available at https://github.com/AliciaMstt/Multihierarchical_NevadoToluca. Raw sequences are at SRA project: PRJNA743572. Laboratory protocol, primers and PCR conditions are in Appendix S1.

ORCID

Nancy Gálvez-Reyes https://orcid.org/0000-0003-2712-2377
Paula Arribas https://orcid.org/0000-0002-0358-8271
Carmelo Andújar https://orcid.org/0000-0001-9759-7402
Brent C. Emerson https://orcid.org/0000-0003-4067-9858
Daniel Piñero https://orcid.org/0000-0002-2509-2445
Alicia Mastretta-Yanes https://orcid.org/0000-0003-2951-6353

REFERENCES

Andújar, C., Arribas, P., Gray, C., Bruce, C., Woodward, G., Yu, D. W., & Vogler, A. P. (2018). Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Molecular Ecology*, *27*(1), 146–166. https://doi.org/10.1111/mec.14410

Andújar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology*, *27*(20), 3968–3975. https://doi.org/10.1111/mec.14844

Andújar, C., Creedy, T. J., Arribas, P., López, H., Salces-Castellano, A., Pérez-Delgado, A. J., Vogler, A. P., & Emerson, B. C. (2021). Validated removal of nuclear pseudogenes and sequencing artefacts from mitochondrial metabarcode data. *Molecular Ecology Resources*, *21*(6), 1772–1787. https://doi.org/10.1111/1755-0998.13337

Andújar, C., Pérez-González, S., Arribas, P., Zaballos, J. P., Vogler, A. P., & Ribera, I. (2017). Speciation below ground: Tempo and mode of diversification in a radiation of endogean ground beetles. *Molecular Ecology*, *26*(21), 6053–6070. https://doi.org/10.1111/mec.14358

Arriaga-Jiménez, A., Rös, M., & Halffter, G. (2018). High variability of dung beetle diversity patterns at four mountains of the Trans-Mexican Volcanic Belt. *PeerJ*, *6*, e4468. https://doi.org/10.7717/peerj.4468

Arribas, P., Andújar, C., Hopkins, K., Shepherd, M., & Vogler, A. P. (2016). Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. *Methods in Ecology and Evolution*, *7*(9), 1071–1081. https://doi.org/10.1111/2041-210X.12557

Arribas, P., Andújar, C., Salces-Castellano, A., Emerson, B. C., & Vogler, A. P. (2020). The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding. *Molecular Ecology*, *30*(1), 1–14. https://doi.org/10.1111/mec.15591

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, *19*, 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Baselga, A., Fujisawa, T., Crampton-Platt, A., Bergsten, J., Foster, P. G., Monaghan, M. T., & Vogler, A. P. (2013). Whole-community DNA barcoding reveals a spatio-temporal continuum of biodiversity at species and genetic levels. *Nature Communications*, *4*, 1892. https://doi.org/10.1038/ncomms2881

Baselga, A., Gómez-Rodríguez, C., & Vogler, A. P. (2015). Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes. *Global Ecology and Biogeography*, *24*(8), 873–882. https://doi.org/10.1111/geb.12322

Baselga, A., Lobo, J. M., Svenning, J. C., Aragón, P., & Araújo, M. B. (2012). Dispersal ability modulates the strength of the latitudinal richness gradient in European beetles. *Global Ecology and Biogeography*, *21*, 1106–1113. https://doi.org/10.1111/j.1466-8238.2011.00753.x

Baselga, A., & Orme, C. D. L. (2012). betapart: An R package for the study of beta diversity: Betapart package. *Methods in Ecology and Evolution*, *3*(5), 808–812. https://doi.org/10.1111/j.2041-210X.2012.00224.x

Baselga, A., Orme, D., Villeger, S., Bortoli, J. D., Leprieur, F., Logez, M., & Henriques-Silva, R. (2018). *betapart: Partitioning Beta Diversity into Turnover and Nestedness Components (1.5.1)* [Computer software]. Retrieved from https://CRAN.R-project.org/package=betapart

Bell, G. (2001). Neutral macroecology. *Science*, *293*(5539), 2413–2418. https://doi.org/10.1126/science.293.5539.2413

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bray, T. C., & Bocak, L. (2016). Slowly dispersing neotenic beetles can speciate on a penny coin and generate space-limited diversity in the tropical mountains. *Scientific Reports*, *6*(1), 33579. https://doi.org/10.1038/srep33579

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, *11*(12), 2639–2643. https://doi.org/10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Cardoso, P., Gaspar, C., Pereira, L. C., Silva, I., Henriques, S. S., da Silva, R. R., & Sousa, P. (2008). Assessing spider species richness and composition in Mediterranean cork oak forests. *Acta Oecologica*, *33*(1), 114–127. https://doi.org/10.1016/j.actao.2007.10.003

Cardoso, P., Henriques, S. S., Gaspar, C., Crespo, L. C., Carvalho, R., Schmidt, J. B., Sousa, P., & Szűts, T. (2009). Species richness and composition assessment of spiders in a Mediterranean scrubland. *Journal of Insect Conservation*, *13*(1), 45–55. https://doi.org/10.1007/s10841-007-9116-3

Cicconardi, F., Fanciulli, P. P., & Emerson, B. C. (2013). Collembola, the biological species concept and the underestimation of global species richness. *Molecular Ecology*, *22*(21), 5382–5396. https://doi.org/10.1111/mec.12472

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, *18*(1), 117–143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

Creedy, T. J., Ng, W. S., & Vogler, A. P. (2019). Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy. *Ecology and Evolution*, *9*(6), 3105–3116. https://doi.org/10.1002/ece3.4839

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*(10), 996–998. https://doi.org/10.1038/nmeth.2604

Edgar, R. C. (2016). UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv*, *81257*. https://doi.org/10.1101/081257

Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics (Oxford, England)*, *31*(21), 3476–3482. https://doi.org/10.1093/bioinformatics/btv401

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—sequence relationships with an innovative metabarcoding protocol. *PLoS One*, *10*(7), e0130324. https://doi.org/10.1371/journal.pone.0130324

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, *7*(17), 6918–6926. https://doi.org/10.1002/ece3.3192

Elbrecht, V., Vamos, E. E., Steinke, D., & Leese, F. (2018). Estimating intraspecific genetic diversity from community DNA metabarcoding data. *PeerJ*, *6*, e4644. https://doi.org/10.7717/peerj.4644

Emerson, B. C., Casquet, J., López, H., Cardoso, P., Borges, P. A. V., Mollaret, N., Oromí, P., Strasberg, D., & Thébaud, C. (2017). A combined field survey and molecular identification protocol for comparing forest arthropod biodiversity across spatial scales. *Molecular Ecology Resources*, *17*(4), 694–707. https://doi.org/10.1111/1755-0998.12617

Fjeldså, J., Bowie, R. C. K., & Rahbek, C. (2012). The role of mountain ranges in the diversification of birds. *Annual Review of Ecology, Evolution, and Systematics*, *43*(1), 249–265. https://doi.org/10.1146/annurev-ecolsys-102710-145113

García-Olivares, V., Patiño, J., Overcast, I., Salces-Castellano, A., López de Heredia, U., Mora-Márquez, F., Machado, A., Hickerson, M. J., & Emerson, B. C. (2019). A topoclimate model for Quaternary insular speciation. *Journal of Biogeography*, *46*(12), 2769–2786. https://doi.org/10.1111/jbi.13689

Gerber, A. S., Loggins, R., Kumar, S., & Dowling, T. E. (2001). Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes? *Annual Review of Genetics*, *35*(1), 539–566. https://doi.org/10.1146/annurev.genet.35.102401.091106

Gómez-Rodríguez, C., & Baselga, A. (2018). Variation among European beetle taxa in patterns of distance decay of similarity suggests a major role of dispersal processes. *Ecography*, *41*(11), 1825–1834. https://doi.org/10.1111/ecog.03693

Gómez-Rodríguez, C., Miller, K. E., Castillejo, J., Iglesias-Piñeiro, J., & Baselga, A. (2019). Understanding dispersal limitation through the assessment of diversity patterns across phylogenetic scales below the species level. *Global Ecology and Biogeography*, *28*(3), 353–364. https://doi.org/10.1111/geb.12857

González-Fernández, A., Arroyo-Rodríguez, V., Ramírez-Corona, F., Manjarrez, J., Aguilera-Hernández, A., & Sunny, A. (2019). Local and landscape drivers of the number of individuals and genetic diversity of a microendemic and critically endangered salamander. *Landscape Ecology*, *34*(8), 1989–2000. https://doi.org/10.1007/s10980-019-00871-2

Goodman, K. R., Welter, S. C., & Roderick, G. K. (2012). Genetic divergence is decoupled from ecological diversification in the Hawaiian Nesoydne planthoppers. *Evolution: International Journal of Organic Evolution*, *66*(9), 2798–2814. https://doi.org/10.1111/j.1558-5646.2012.01643.x

Goslee, S., & Urban, D. (2020). *ecodist: Dissimilarity-Based Functions for Ecological Analysis (2.0.5)* [Computer software]. Retrieved from https://CRAN.R-project.org/package=ecodist

He, K., Gutiérrez, E. E., Heming, N. M., Koepfli, K.-P., Wan, T., He, S., Jin, W., Liu, S.-Y., & Jiang, X.-L. (2019). Cryptic phylogeographic history sheds light on the generation of species diversity in sky-island mountains. *Journal of Biogeography*, *46*(10), 2232–2247. https://doi.org/10.1111/jbi.13664

Hubbell, S. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton University Press; JSTOR. https://doi.org/10.1515/9781400837526

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN community edition—interactive exploration and analysis of large-scale microbiome sequencing data. *PLOS Computational Biology*, *12*(6), e1004957. https://doi.org/10.1371/journal.pcbi.1004957

Janzen, D. H. (1967). Why Mountain Passes are Higher in the Tropics Author (s): Daniel H. Janzen Source: The American Naturalist , Vol. 101 , No. 919 (May—Jun ., 1967), pp. 233-249 Published by: The University of Chicago Press for The American Society of Naturalist. *The American Naturalist*, *101*(919), 233–249.

Knowles, L. L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshopprers. *Molecular Ecology*, *10*(3), 691–701. https://doi.org/10.1046/j.1365-294x.2001.01206.x

Lee-Yaw, J. A., Davidson, A., McRae, B. H., & Green, D. M. (2009). Do landscape processes predict phylogeographic patterns in the wood frog? *Molecular Ecology*, *18*(9), 1863–1874. https://doi.org/10.1111/j.1365-294X.2009.04152.x

Mastretta-Yanes, A., Moreno-Letelier, A., Piñero, D., Jorgensen, T. H., & Emerson, B. C. (2015). Biodiversity in the Mexican highlands and the interaction of geology, geography and climate within the Trans-Mexican Volcanic Belt.c. *Journal of Biogeography*, *42*(9), 1586–1600. https://doi.org/10.1111/jbi.12546

Mastretta-Yanes, A., Xue, A. T., Moreno-Letelier, A., Jorgensen, T. H., Alvarez, N., Pinero, D., & Emerson, B. C. (2018). Long-term insitu persistence of biodiversity in tropical sky islands revealed by landscape genomics. *Molecular Ecology*, *27*(2), 432–448. https://doi.org/10.1111/mec.14461

McCormack, J. E., Huang, H., Knowles, L. L., Gillespie, R., & Clague, D. (2009). Sky islands. In *Encyclopedia of islands* (pp. 841–843).

McGill, B. J., Chase, J. M., Hortal, J., Overcast, I., Rominger, A. J., Rosindell, J., Borges, P. A. V., Emerson, B. C., Etienne, R. S., Hickerson, M. J., Mahler, D. L., Massol, F., McGaughran, A., Neves, P., Parent, C., Patiño, J., Ruffley, M., Wagner, C. E., & Gillespie, R. (2019). Unifying macroecology and macroevolution to answer fundamental questions about biodiversity. *Global Ecology and Biogeography*, *28*(12), 1925–1936. https://doi.org/10.1111/geb.13020

McRae, B. H. (2006). Isolation by resistance. *Evolution*, *60*(8), 1551–1561. https://doi.org/10.1554/05-321.1

McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, *89*(10), 2712–2724. https://doi.org/10.1890/07-1861.1

McRae, B. H., Shah, V. B., & Mohapatra, T. K. (2013). *Circuitscape 4 User Guides (4.0)* [Computer software]. The Nature Conservancy. Retrieved from http://www.circuitscape.org

Múrria, C., Bonada, N., Vellend, M., Zamora-Muñoz, C., Alba-Tercedor, J., Sainz-Cantero, C. E., Garrido, J., Acosta, R., El Alami, M., Barquín, J., Derka, T., Álvarez-Cabria, M., Sáinz-Bariain, M., Filipe, A. F., & Vogler, A. P. (2017). Local environment rather than past climate determines community composition of mountain stream macroinvertebrates across Europe. *Molecular Ecology*, *26*(21), 6085–6099. https://doi.org/10.1111/mec.14346

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package (2.5-6)* [Computer software]. Retrieved from https://CRAN.R-project.org/package=vegan

Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, *15*(4), 819–830. https://doi.org/10.1111/1755-0998.12355

Polato, N. R., Gill, B. A., Shah, A. A., Gray, M. M., Casner, K. L., Barthelet, A., Messer, P. W., Simmons, M. P., Guayasamin, J. M., Encalada, A. C., Kondratieff, B. C., Flecker, A. S., Thomas, S. A., Ghalambor, C. K., Poff, N. L., Funk, W. C., & Zamudio, K. R. (2018). Narrow thermal tolerance and low dispersal drive higher speciation in tropical mountains. *Proceedings of the National Academy of Sciences*, *115*(49), 12471–12476. https://doi.org/10.1073/pnas.1809326115

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595–609. https://doi.org/10.1080/10635150600852011

Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-Bravo, D., Rasmussen, C. M. Ø., Richardson, K., Rosing, M. T., Whittaker, R. J., & Fjeldså, J. (2019). Building mountain

biodiversity: Geological and evolutionary processes. *Science*, *365*(6458), 1114–1119. https://doi.org/10.1126/science.aax0151

Rahbek, C., Borregaard, M. K., Colwell, R. K., Dalsgaard, B., Holt, B. G., Morueta-Holme, N., Nogues-Bravo, D., Whittaker, R. J., & Fjeldså, J. (2019). Humboldt's enigma: What causes global patterns of mountain biodiversity? *Science*, *365*(6458), 1108–1113. https://doi.org/10.1126/science.aax0149

Rambaut, A. (2012). *FigTree v.1.4.2*. Retrieved from http://tree.bio.ed.ac.uk/software/figtree/

Roffler, G. H., Schwartz, M. K., Pilgrim, K. L., Talbot, S. L., Sage, G. K., Adams, L. G., & Luikart, G. (2016). Identification of landscape features influencing gene flow: How useful are habitat selection models? *Evolutionary Applications*, *9*(6), 805–817. https://doi.org/10.1111/eva.12389

Rosindell, J., Hubbell, S. P., & Etienne, R. S. (2011). The unified neutral theory of biodiversity and biogeography at age ten. *Trends in Ecology & Evolution*, *26*(7), 340–348. https://doi.org/10.1016/j.tree.2011.03.024

Rzedowski, J. (2006). Bosque de coníferas. *Vegetación De México*, Comisión Nacional para el Conocimientoy Uso de la Biodiversidad, *1*, 1, 295–327. México.

Salces-Castellano, A., Patiño, J., Alvarez, N., Andújar, C., Arribas, P., Braojos-Ruiz, J. J., Arco-Aguilar, M., García-Olivares, V., Karger, D. N., López, H., Manolopoulou, I., Oromí, P., Pérez-Delgado, A. J., Peterman, W. E., Rijsdijk, K. F., & Emerson, B. C. (2020). Climate drives community-wide divergence within species over a limited spatial scale: Evidence from an oceanic island. *Ecology Letters*, *23*(2), 305–315. https://doi.org/10.1111/ele.13433

Schnell, I. B., Bohmann, K., & Gilbert, M. T. P. (2015). Tag jumps illuminated–reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*, *15*(6), 1289–1303. https://doi.org/10.1111/1755-0998.12402

Sheldon, K. S., Huey, R. B., Kaspari, M., & Sanders, N. J. (2018). Fifty years of mountain passes: A perspective on Dan Janzen's classic article. *The American Naturalist*, *191*(5), 553–565. https://doi.org/10.1086/697046

Shokralla, S., Porter, T. M., Gibson, J. F., Dobosz, R., Janzen, D. H., Hallwachs, W., Golding, G. B., & Hajibabaei, M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Scientific Reports*, *5*(1), 9687. https://doi.org/10.1038/srep09687

Staton, E. (2019). *Sestaton/Pairfq* [Perl]. Retrieved from https://github.com/sestaton/Pairfq (Original work published 2013)

Thompson, R., & Townsend, C. (2006). A truce with neutral theory: Local deterministic factors, species traits and dispersal limitation together determine patterns of diversity in stream invertebrates. *Journal of Animal Ecology*, *75*(2), 476–484. https://doi.org/10.1111/j.1365-2656.2006.01068.x

Uscanga, A., López, H., Piñero, D., Emerson, B. C., Mastretta-Yanes, A., & Parmakelis, A. (2021). Evaluating species origins within tropical sky-islands arthropod communities. *Journal of Biogeography*, *48*(9), 2199–2210. https://doi.org/10.1111/jbi.14144

Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, *85*(2), 183–206. https://doi.org/10.1086/652373

Weber, M. G., Wagner, C. E., Best, R. J., Harmon, L. J., & Matthews, B. (2017). Evolution in a community context: On integrating ecological interactions and macroevolution. *Trends in Ecology & Evolution*, *32*(4), 291–304. https://doi.org/10.1016/j.tree.2017.01.003

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & RStudio. (2020). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics (3.3.2)* [Computer software]. Retrieved from https://CRAN.R-project.org/package=ggplot2

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, *3*(4), 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Gálvez-Reyes, N., Arribas, P., Andújar, C., Emerson, B. C., Piñero, D., & Mastretta-Yanes, A. (2021). Dispersal limitations and long-term persistence drive differentiation from haplotypes to communities within a tropical sky-island: Evidence from community metabarcoding. *Molecular Ecology*, 30, 6611–6626. https://doi.org/10.1111/mec.16195

CAPÍTULO III

RIQUEZA Y COMPOSICIÓN EN LAS COMUNIDADES DE ARTRÓPODOS EN BOSQUES BAJO MANEJO FORESTAL, DISTURBIO Y REFORESTACIÓN

Fotografía de Giezi Anthony Gálvez: Cráter y bosque del Nevado de Toluca.

PREFACIO

Actualmente, la amenaza pérdida de la biodiversidad y servicios ecosistémicos causada por la deforestación y la fragmentación de los bosques es recurrente. La solución consiste en la recuperación de áreas deforestadas mediante plantaciones forestales y silvicultura sustentable (Back et al., 2020; Torres-Rojo et al., 2016). Sin embargo, es crucial monitorear el impacto de estas estrategias en la biodiversidad para garantizar el cumplimiento de los objetivos de conservación. El ADN de wcMBC de artrópodos se presenta como una metodología capaz de proporcionar datos de alta resolución para monitorear el estado de conservación. Los artrópodos constituyen un taxón apropiado para evaluar las diferencias entre comunidades, y su composición puede indicar los efectos del manejo forestal, los disturbios y la reforestación sobre la biodiversidad.

Aquí evaluamos el papel de las limitaciones de dispersión, el manejo forestal, el disturbio, y la reforestación como impulsores de la diferenciación de haplotipos a nivel comunitario dentro de la Isla del cielo tropical, ubicada en el Nevado de Toluca, la cual también es un área protegida con bosques bajo conservación y manejo forestal de algunas áreas. Para llevar a cabo de este análisis, recolectamos muestras de artrópodos de seis órdenes utilizando ADN wcMBC, con un rendimiento de secuenciación de variantes de secuencias de amplicones COI (ASV). Exploramos múltiples niveles jerárquicos, desde haplotipos hasta linajes, estableciendo umbrales de similitud del 3% y 5%, con el fin de evaluar patrones de riqueza, rotación y disminución de la distancia de similitud, utilizando un enfoque de aislamiento por distancia. Nuestros resultados revelaron que los bosques

bajo conservación exhiben la mayor riqueza, y se observa un cambio significativo en la estructura y composición de las comunidades, el cual varía según los diferentes tipos de condiciones de uso del bosque. Finalmente, encontramos que la distancia, el manejo forestal, el disturbio y las actividades de reforestación influyen en el deterioro de la similitud en todos los niveles jerárquicos, lo que sugiere que son impulsores importantes de la diversidad dentro de las islas tropicales.

Artículo: Strong turnover in the composition on forests subjected to management and reforested activities revealed by metabarcoding of arthropod communities. ***(En preparación)***

Revista: Conservation biology

Gálvez-Reyes Nancy*[1,2], Piñero Daniel[1], and Mastretta-Yanes Alicia*[3,4].

[1] Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad Universitaria, Coyoacán, 04510 CDMX, México.

[2] Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México, Ciudad Universitaria, CDMX, México.

[3] Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), Liga Periférico Insurgentes Sur 4903, Col. Parques del Pedregal, Tlalpan, CDMX, México.

[4] Consejo Nacional de Ciencia y Tecnología, Benito Juárez (CONACYT), CDMX, México, Avenida Insurgentes Sur 1582, Crédito Constructor, Benito Juárez, CDMX, México.

*nancygalvez@ecologia.unam.mx , pinero@unam.mx , amastretta@conabio.gob.mx

Abstract (150-250 words)

Recovering deforested areas through forest plantations and sustainable forestry has been proposed as a strategy to combat biodiversity loss, arguing that plantations and sustainable forestry ensure the persistence of the tree cover. However, they also involve anthropogenic changes and disturbances, whose effects on fauna communities and biodiversity evolution are poorly known. Besides, fine-scale community differences can be due to neutral processes, confounding the effect of human activities. Here, we evaluate the role of dispersal limitations, management and reforested activities as drivers of differentiation from the haplotype to the community levels within the tropical sky-island Nevado de Toluca, which is also a protected area with forests under both conservation or sustainable forestry. For this, we sampled arthropods of six orders using whole-communities metabarcoding with COI sequencing throughput Amplicons Sequences Variant (ASV). We explored multiple hierarchical levels, from haplotypes to lineages at 3% and 5% similarity thresholds, to evaluate patterns of richness, turnover, and distance decay of similarity with an isolation-by-distance approach. Our results showed that forests under conservation have the highest richness and that there is strong turnover in the structure and composition of communities, varying according to the different types of forest use conditions. Finally, distance and human activities influence on the decay of similarity at all hierarchical levels. The patterns vary by arthropod group due to each having different dispersal abilities, depending on spatial scale and the type of anthropogenic change.

Keywords: Arthropods, management forest, reforestation, under conservation, whole-community metabarcoding, Nevado de Toluca.

Introduction

Deforestation and fragmentation of forests are main threats to biodiversity and ecosystem services. As a strategy to combat this biodiversity decline, it has been proposed to recover deforested areas through forest plantations and to perform sustainable forestry as an alternative to illegal logging (Back et al., 2020; Torres-Rojo et al., 2016). For example, these strategies were part of the arguments used to change the type of natural protected area of the Nevado de Toluca, México, which was downgraded from "National Park" to "Area of Protection for Flora and Fauna" (Depraz et al., 2017; Mastretta-Yanes et al., 2014), thus allowing sustainable forestry in territories previously devoted to conservation. However, to ensure that conservation objectives are fulfilled locally, it is necessary to have a baseline of local diversity and perform periodic monitoring. From studying plantations and other intensively managed forests we know that they often support less biodiversity than native forests (Wang et al., 2019; Zhang et al., 2016), for this reason it has been proposed to divide the landscape into reserves separated from areas of extensive management (Betts et al., 2021), but forest under sustainable management had been less studies. However, plantations with native species and sustainable forestry indeed ensure the persistence of tree cover (Nghiem & Tran, 2016), so it is necessary to understand with detail the effects of forestry on fauna communities and the evolution of biodiversity. In addition to that, fine-scale community differences can be due to neutral processes, related to dispersal limitations (Arribas et al., 2020; Baselga et al., 2013, 2015), confounding the effect of human activities. Therefore, changes in the richness and composition of arthropod communities could be either due to the differences between the different management treatments (e.g., plantations, sustainable forestry) or due to neutral processes through the dispersive limitation of the

individuals conforming the communities. To study these effects from an ecological perspective and to provide effective management recommendations, data on the distribution of richness and composition of whole-communities over fine and large geographic scales and management types are needed, but this has been hampered by the absence of sufficient empirical data and monitoring schemes.

Biodiversity monitoring documents temporal changes in the distribution and existence of species (Hajibabaei et al., 2016; Li et al., 2010). However, the information, technical and financial capacity to monitor and analyze trends of change are limited to allow large-scale comparisons (Buss et al., 2015). Whole-communities metabarcoding (wcMBC) is a methodology that involves bulk sampling sequencing of mixed communities and subsequent clustering of DNA reads into operational taxonomic units (OTUs) to approximate community profiles at the species level (Andújar et al., 2017, 2021; Arribas et al., 2020; Ji et al., 2013).

The wcMBC of arthropods could be applied to monitor the conservation status of Mexican forests, because it has shown to produce high resolution data within a single Mexican mountain (Gálvez-Reyes et al., 2021). Additionally, arthropods are a suitable taxon for evaluating if community differences are due to neutral processes or management (treatments) differences because arthropod diversity is a good structural indicator of ecosystems (Obrist & Duelli, 2010), correlating with plant communities (Basset et al., 2012; Lynggaard et al., 2020; Zhang et al., 2016). Arthropods constitute important components of the mountains of Mexico due to their richness, abundance, diversity of niches they occupy and for presenting species with both Neartic and Neotropical biogeographic histories (Morrone & Márquez, 2001). In addition, they have different dispersive limitations with

different effects on the differentiation between localities, as expected by the theory of neutral biodiversity (Baselga et al., 2013, 2015).

The composition of the arthropod communities in montane regions is shaped by the height and isolation of the mountain, the habitat, the degree of conservation and age of the forest, complexity of the canopy, undergrowth, spatial distance and human activities (Barsoum et al., 2019; Colwell et al., 2008; Lynggaard et al., 2020; Uscanga et al., 2021; Zhang et al., 2016). For example, different nearby mountains may have different species composition since climates are not continuous between them, leading to isolation and increased potential for speciation (Graham & Fine, 2008; Kozak & Wiens, 2006; Uscanga et al., 2021; Wiens, 2004). Topoclimatic variables can be a driver of metacommunity structure within habitat (Noguerales et al., 2021). As for the effect of land-use, conserved forests are often richer than managed forests (Lynggaard et al., 2020; Paillet et al., 2010; Wang et al., 2019; Zhang et al., 2016), since this type of habitat can offer diverse ecological niches (Graham & Fine, 2008). The composition of arthropods has been explained by the different land-uses in the forest according to stands under forest management over time (Barsoum et al., 2019). Forest arthropod monitoring can detect the biodiversity benefit of native forests and mixed plantations over monocultures (Wang et al., 2019). As well as detecting changes in the composition of arthropods in ecosystems under restoration (Edwards et al., 2014; Fernandes et al., 2019; Lynggaard et al., 2020). It has also been suggested that the dense canopy cover of forests allows for greater diversity (Lange et al., 2014; Zou et al., 2015). In addition, the undergrowth harbors a greater diversity of insects compared to its thinning (Rambo et al., 2014); in such a way that the ground vegetation (undergrowth) could harbor greater diversity and richness of arthropods (Lange et al., 2014). Thus, it is possible that some groups of arthropods are correlated to mature forests. For example, a study found that insect

richness and diversity increased in locations with greater structural complexity and environmental heterogeneity, particularly with increasing canopy cover, litter, and dead wood (Fuller et al., 2008). Likewise, Beng et al. observed that different types of land-use showed different communities of unique arthropods, this is reflected in the marked differences in the structure of the communities between the different types of forests (Beng et al., 2016). In parallel to the former drivers of differentiation, the neutral dispersal limitation, which is the process that causes the location of an individual to be restricted in some sense by the location of its parent, has been highlighted by the neutral theory as the main driver of local diversity (Hubbell, 2001; Rosindell et al., 2011). This means that dispersal limitations of individuals give rise to spatial patterns at large (McGill, 2010), and that the limited spatial scale may influence dispersal restrictions in soil arthropods (Arribas et al., 2020; Noguerales et al., 2021). Just as community connectivity occurs due to height and distance within a forest under conservation (Gálvez-Reyes et al., 2021), and dispersal capacity and climatic tolerance are restricted, strong geographic isolation can be found over distances of only a few kilometers for multiple arthropod species coexisting on an oceanic island (Salces-Castellano et al., 2020). Therefore, before testing whether forest management and plantations are being carried out in a way that impacts local biodiversity, arthropod's dispersal limitations should be considered.

The tropical sky islands are useful for testing dispersal constraints of arthropods because they provide an environment of high elevation gradients and steep valleys, inhabited by highly isolated communities or communities from other mountains (Mastretta-Yanes et al., 2015; Rahbek et al., 2019; Uscanga et al., 2021). These factors have been used to explain the phylogeographic and biogeographic patterns of structure found among the Mexican tropical sky islands (Mastretta-Yanes et al., 2015; Uscanga et al., 2021). We now know that

the changes in the composition, from the genetic to the community level, can occur due to neutral processes at a fine geographical scale inside the sky-island Nevado de Toluca (Gálvez-Reyes et al., 2021). However, changes can also be due to differences between types of management, therefore it is necessary to determine if forest management and plantations are being carried out in a way that impacts the local structure and biodiversity and the confounding effect of both, isolation by distance and treatment. We hypothesized that the diversity and structure of the arthropod community varies according to the distance locality and the different human activities (forest management, disturbed and reforested) in the Nevado de Toluca. Here, we evaluate fine-scale spatial patterns of diversity and composition of arthropod fauna to test the role of dispersal limitation and different types of forest management and forest conservation status.

Materials and methods

Study sites and sampling design

The study sites include *Abies religiosa* and *Pinus hartwegii* forests which grow at around 2,500-3,500 m.a.s.l. in the Nevado de Toluca mountain (Figure 1a), which is a sky-island within the TMBV (Mastretta-Yanes et al., 2015; Rzendowski, 2006). We sampled arthropods from treatments that include forests under conservation, forestry managed forest, *Abies religiosa* disturbed forests and reforested forests of *Pinus hartwegii* as detailed below and in Table S1. Sampling was performed in the rainy season of 2015 during mid-August and September with pitfall traps as detailed in Gálvez-Reyes et al. (2021). At each sampling point arthropods were sampled with pitfall traps distributed in three blocks of 15 x 20 m with 20 pitfalls each (Figure 1c). Each sample (biodiversity soup) consisted of all the specimens

collected by 20 pitfall traps, totalling 2,880 pitfall traps and 115 biodiversity soups (Figure 1; Table S1). Sampling blocks were separated by at least 50 m and maximum distance among sampling points was 20 km (Figure 1). Pitfall traps of each block site (20 traps) were collected after 15 days, as suggested by Cardoso et al. (2008, 2009) and Emerson et al. (2017) and pooled in a single bulk-sample in a bottle containing ethanol at 96% as suggested by Gálvez-Reyes et al. (2021). Sampling was performed with SEMARNAT permit No. SGPA/DGVS/02641/15.

Our experimental design included the following comparisons: (i) *A. religiosa forests under management versus forest under conservation*. Forest under conservation (CON) were inside the protected area Nevado de Toluca. Forestry management included chaponeo (cutting herbs and bushes to increase light availability to young trees and decrease fuel material for forest fires) activity chaponeo, cut 10% of the wood volume one year ago (Y1), seven years ago (Y7), and 10 years ago (Y10). The former treatments were sampled with 63 biodiversity soups (1,260 pitfall traps in total) distributed in the Amanalco locality. (ii) *Disturbed forest (DIS) vs forest under conservation.* The disturbed forest was an *A. religiosa* forest where cows were allowed to pasture freely. For this, 18 biodiversity soups were used (360 pitfall traps) in the Santiago Tlacotepec locality. (iii) *Reforested plantation vs forest under conservation.* This is a *Pinus hartwegii* forest that had been reforested in 2013 (R13) and another which was reforested in 1980 (R80). In total 27 biodiversity soups were used (540 pitfall traps) distributed across Santiago Tlacotepec and Zinacantepec localities. (iv) *A. religiosa vs P. hartwegii forests under conservation.* For these 36 biodiversity soups were used (720 pitfall traps), distributed across San Juan de las Huertas, Santiago Tlacotepec and Zinacantepec localities.

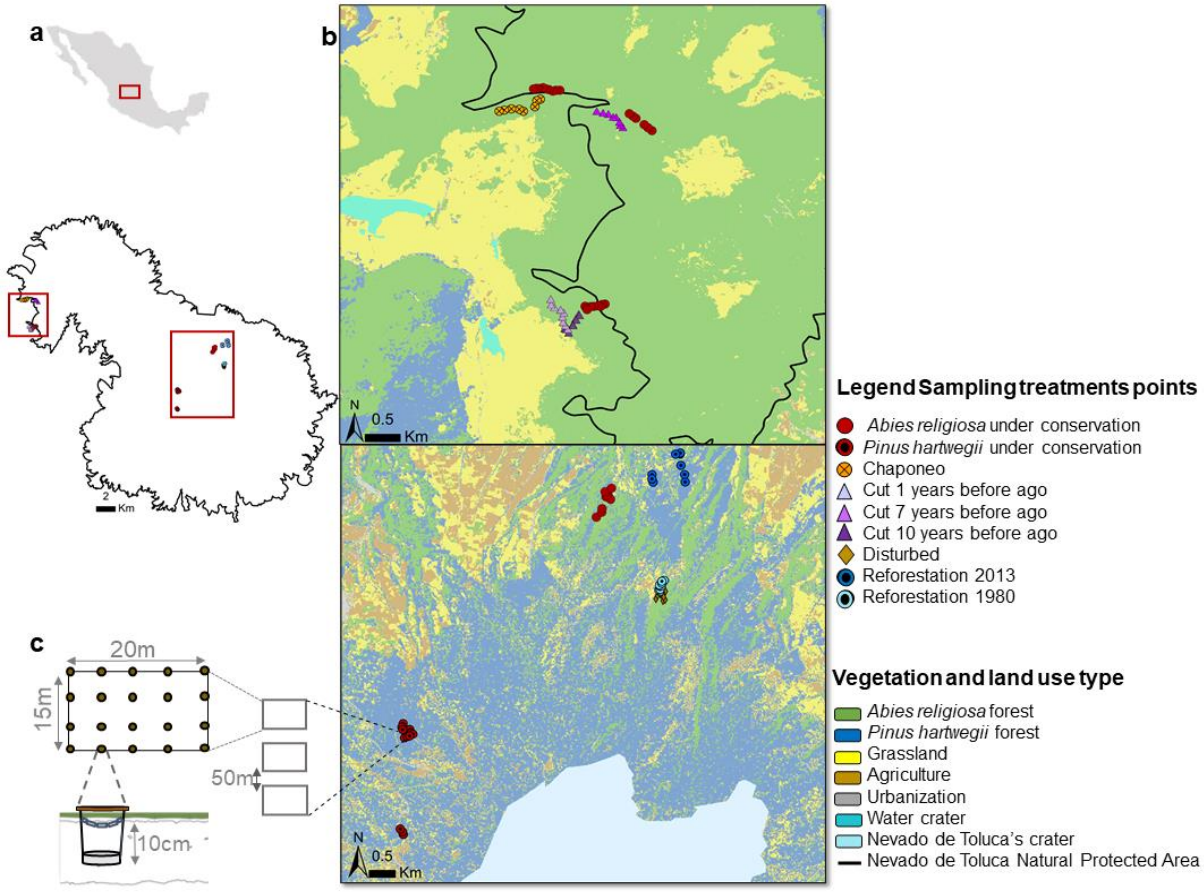

Figure 1. Study site and sampling design. **a)** Location of the study area in Mexico and Map of the protected natural area Nevado de Toluca. **b)** Sampling points area over a vegetation and land-use map. At each sampling point (symbols) we sampled three sampling blocks included under conservation (CON: red empty circle), management forest includes chaponeo activity (chaponeo: yellow circle), Cut one year ago (Y1: triangle purple light), Cut seven years ago (Y7: triangle magenta), Cut ten years ago (Y10: triangle purple) and disturbed (DIS: brown symbol) *Abies religiosa* forest. The reforestation forest includes reforestation 2013 (R13: blue light full circle) and reforestation 1980 (R80: blue full circle) within *Pinus hartwegii* forest. **c)** Sampling blocks were separated by 50 m (within sampling points) to up to 20 km (between sampling points), each block of 20 x 15 m had 20 pitfall traps equidistantly distributed. Treatments sampling points were distributed in five sites to the West (ABB: Agua Bendita and ASB: San Bartolo) and East (SJH: San Juan de las Huertas, TCL: Tlacotepec and ZIN: Zinancatep) hillsides of Nevado de Toluca. For vegetation types of details see González-Fernández et al. (2019).

Molecular biology methods

Samples (comprising 20 pooled pitfalls each) were cleaned up following the Flotation-Filtration-Stereoscope protocol (FFS) that allowed us to have 'clean' bulk specimens ready for DNA genomic extraction (see Gálvez-Reyes et al., 2021). Due to the samples having considerably different body sizes, we divided specimens by size class boundaries as suggested by Gálvez-Reyes et al. (2021): small (e.g., size *Drosophila melanogaster* <=3 mm), medium (e.g., size adult *Apis mellifera* >3mm and <=15 mm), or large specimens (e.g., adult grasshopper >15mm). After size classification, we divided each specimen according to Gálvez-Reyes et al. (2021) as follows:complete bodies for small arthropods, the thorax (head included) from medium-sized arthropods and two legs from large arthropods. This procedure generated 288 subsamples of bulk specimens for DNA extraction and library prep processed independently and then pooled (see below), but using the same library barcode identifier per subsample (thus, the final number of libraries retrieved was 115, one library per sample). For DNA genomic extractions, each bulk-sample was independently homogenized with an electric disruptor (Qiagen TissueLyser II: two times for 1 min at 30 Hz and 30 s at 30 Hz) using sterile 7mm stainless steel beads and Qiagen DNeasy Blood & Tissue Kit and BSA buffer 1x. We quantified the DNA concentration using Qubit 3.0 Fluorometer.

The preparation of metabarcoding libraries was performed as follows: we used the double dual tagging method with the XT Illumina Adapter Kit (Arribas et al., 2016). We PCR amplified a 418-base pair within the standard barcode region from the 5' end of mitochondrial COI gene for metazoa with the primers B_F 5' CCIGAYATRGCITTYCCICG 3' (Shokralla et al., 2015) and Fol-degen-rev 5' TANACYTCNGGRTGNCCRAARAAYCA 3' (Yu et al., 2012) modified to include Illumina overhang adaptors for subsequent PCR. For each subsample (n=288), we performed and pooled together equimolar amounts of three

independent PCR replicates for each arthropod-size sample. We included a negative control reaction with no DNA template in all experiments. Amplifications were performed following the PCR condition and information regarding PCR reagents are available in Arribas et al. (2016) and Gálvez-Reyes et al. (2021). Each pool of PCR amplicons was purified of remaining primers and primer dimers using magnetic bead-based Agencourt AMPure XP beads protocol (Beckman Coulter). Then, amplicons were used as template for a limited-cycle PCR amplification to add dual-indices barcodes (N7 index and S5 index) and the P5 and P7 Illumina sequencing adapters using the Nextera XT Index Kit from Illumina (Gálvez-Reyes et al., 2021). PCR products were visualized on agarose gel and red gel. The 115 metabarcoding libraries and two negative controls (DNA extraction and PCR amplification) were sequenced on two lanes of Illumina MiSeq 2x300-bp at the Cornell Institute of Biotechnology, Cornell University, USA.

Metabarcoding bioinformatics read processing

The paired-end reads processing of the 115 libraries were quality filtered following procedures described by Arribas et al. (2020). Briefly, processing included quality checking, primer removal, pair merging, quality filtering, denoising, and clustering each library independently. Raw reads were quality checked using Fastqc version 0.11.7 (Babraham Institute, 2013). Primers were trimmed at 21 and 27 bases for R1 and R2, respectively using the fastx_trimmer option of the Fastx-toolkit version 0.7 (Hannon, 2010) and reads were processed in Trimmomatic-0.36 (Bolger et al., 2014) using TRAILING:20 to remove bases at the end of a read. R1 and R2 reads were used to search paired sequences with pairfq-0.17 (Staton, 2013/2019). After this, libraries were processed with several steps including the following: reads were merged with the fastq_mergepairs command, -fastq_minovlen 50 and

-fastq_maxdiffs 15 from Usearch-9.2 (Edgar, 2013). Quality-filtered (Maxee = 1), dereplicated (-fastx_uniques) and trimmed to full length amplicons of 416-420 bp by sorted (-sortbylength) options in Usearch-10 (Edgar & Flyvbjerg, 2015). Surviving reads were denoised and chimera checked (-unoise3: -alpha 2, -minsize 4) to generate Amplicon Sequence Variants (ASVs) sensu (Callahan et al., 2016), which are suggested to be predicted biological sequences to be used for direct analysis without the need of OTU clustering (Callahan et al., 2017; Edgar, 2016).

High-level taxonomic categories (order level) were then assigned to ASVs for each of the following orders: Diptera, Collembola, Arachnida, Coleoptera, Hymenoptera, and Lepidoptera, using the lowest common ancestor (LCA) algorithm of MEGAN-6 (Huson et al., 2016). Taxonomic identification of each read was done using BLAST search against a reference library including the NCBI nt database and BOLD (accessed on January 03 2022; blastn -outfmt 5 -evalue 0.001). Remaining sequences were not further considered. The tree was exported, visualized, and edited using figtree-1.4.3 (Rambaut, 2015/2023). Each ASV dataset was aligned in Geneious-8.0.2 with MAFFT, using the FFT-NS-1 algorithm, a scoring matrix of 200/PAM/K=2, GAP open penalty of 3, and the translation align option. All sequences were inspected for insertions, deletions or stop codons disrupting the reading frame, which then were excluded.

A community table was then generated with read-counts (haplotype abundance) of each retained ASV for the six orders by matching ASVs against the complete collection of reads (i.e., reads before the dereplicating and denoising steps) using the -search_exact command with usearch-10 (Edgar & Flyvbjerg, 2015). Additional filtering according to ASV abundances in community tables (one ASV per taxa) was performed as described by Arribas et al. (2020). To summarize, we first removed from each library those ASVs with abundances

of four or fewer reads (same criteria as denoising). Second, within each library we removed ASVs that comprised less than 1% of reads within the taxonomic group that they belonged to. Although the recently released metaMATE software (Andújar et al., 2021) can provide for a more robust evaluation of filtering strategies, its application here was compromised by the limited availability of reference sequences matching the obtained ASVs (only 5 ASVs were classified as verified-authentic ASV). However, we were able to compare the efficiency of the 1% threshold used here with metaMATE, but only indirectly as metaMATE excludes ASVs only if they fall below a given threshold in all libraries (or binning groups) where they are found. Consequently, we used metaMATE to explore the effect of filtering by taxonomic group and library with thresholds of 0.8%, 1%, and 1.5%. These three thresholds removed respectively 79%, 83%, and 88% of ASVs identified as verified non-authentic by metaMATE, indicating that our filtering appropriately reduces the prevalence of NUMTs and sequencing noise for the estimation of community-wide geographical patterns within the multi-hierarchical approach. Community tables of filtered haplotypes were then transformed into incidence (presence/absence) data and used for downstream analyses.

To identify lineages within each of the orders we use a range of clustering levels. ASVs filtered data sets were used to generate an UPGMA tree under corrected distances using F84 model, and based on this tree, filtered ASVs were considered haplotypes, which were then nested into lineages following the genetic similarity at different thresholds of genetic similarity (0.5, 1.5, 3, 5 and 7.5%), plus an additional threshold derived from a species delimitation analyses conducted with the generalized mixed Yule-coalescent (GMYC) model in R (Pons et al., 2006). Analyses were performed using Vegan (Oksanen et al., 2019), Cluster (Maechler et al., 2022), PMCMR (Pohlert, 2021), Hier.part (Nally & Walsh, 2004), Ecodist (Goslee & Urban, 2007), and Betapart (Baselga & Orme, 2012).

Community diversity, composition, and ordination of the treatments

The subsets of ASVs of six arthropod orders were used to conduct analyses of community diversity and composition among treatments, considering haplotypes (raw ASVs) and all lineages (0.5, 1.5, 3, 5 and 7.5% lineages, plus that corresponding to the GMYC species delimitation). Total accumulation richness curves were first estimated for multiple levels of genetic similarity (haplotypes, lineages at 3% and 5%) using the R package Betapart (Baselga & Orme, 2012). Then, ANOVAs were used to test for significant differences in alpha diversity between the communities of different treatments of each sampling point.

The community composition variation from different treatments were assessed. Total beta diversity (Sorensen index, βsor), additive turnover (Simpson index, βsim; species replacement, without the effect of variation in richness) and nestedness (Sorensen–Simpson index, βsne; pure richness effect) components (Baselga, 2010) were then estimated based on community compositions at different hierarchical levels. We used the R package Vegan (Oksanen et al., 2019) and community composition matrices to perform non-parametric multidimensional scaling ordination (NMDS) based on Sorensen similarity. Plots were created with the ordispider option to visualize the compositional ordination of the communities according to respective treatment. An analysis of similarity (Anosim) was performed for each taxonomic group to compare arthropod community composition among sites. Anosim is a non-parametric method for analyzing variance and testing multivariate differences between groups, based on a resemblance matrix and rank dissimilarity (Clarke, 1993). Plots were visualized with the R package ggplot2 (Wickham et al., 2020).

Distance decay of similarity of communities

To examine the effect of distance on the variation of arthropods community composition but considering the multi-hierarchical approach of Baselga et al. (2013), we tested for distance decay of similarity across the different lineages. We considered the following as independent variables: (i) each local site and treatment (1 – pairwise beta diversity) and (ii) geographic distance (i.e., isolation by distance, IBD) using the Euclidean distance in kilometers. We focused on Diptera, Collembola, Arachnida, Coleoptera, Hymenoptera and Lepidoptera assemblages for IBD, because they were the six sampled orders showing the highest completeness values. For response variables, pairwise similarity of assemblages at the haplotype and higher lineages (0.5%, 1.5%, 3%, 5%, 7.5% and GMYC) among all treatments were used. The analysis was done with all six orders for the entire treatment sampling. At each lineage threshold, patterns of similarity between pairs of treatment were calculated using the means of the Simpson's similarity index (1- pairwise $\beta sim = a/[a + min(b,c)]$ diversity) using the R package betapart (Baselga & Orme, 2012), where a is the number of species present in both territories, and b and c the number of species unique to one or another, respectively.

For the multihierarchical assessment of the variation in community composition (distance decay of similarity patterns) with spatial distance (computed in meters as the Euclidean distance) were assessed at each level of genetic similarity (from haplotypes to 7.5% lineages). A negative exponential function 'decay.model' in the R package Betapart (Baselga & Orme, 2012) was used to adjust a negative exponential function to a generalized linear model (GML). We used Simpson similarity (1 – βsim; Baselga, 2010) as a response variable, pairwise effective distances of each resistance surface as predictors, log link and Gaussian error (Arribas et al., 2020; Gómez-Rodríguez & Baselga, 2018). Analyses and

graphical representations of data were performed with the R packages Vegan (Oksanen et al., 2019), Betapart (Baselga et al., 2018) and Ecodist (Goslee & Urban, 2007).

Results

Sequencing metabarcoding data and ASVs recovering

In total, we processed 115 metabarcoding libraries performed by double dual tagging of COI amplicons with Miseq Illumina sequencing that generated 30,410,338 paired reads. We obtained from 95,756 to 466,792 reads in each direction per sample. Of these, from 39,484 to 192,273 sequences remained after quality filtering (totalling 13,061,708). After read merging and sequence filtering to a length of 418 bp, each sample comprised between 31,006 and 167,120 sequences (totalling 11,229,150). Taxonomic assignments with usearch showed high similarity to a broad range of arthropod species (Figure S2). The 4,400,273 sequences comprised 2,315 ASVs (unique variants: Diptera, 685; Collembola, 470; Arachnida, 255; Coleoptera, 236; Hymenoptera, 183; Lepidoptera, 71). The GMYC threshold values obtained were: Diptera, 0.9%; Collembola, 2.9%; Arachnida, 1.3%; Coleoptera, 0.7%; Hymenoptera, 1%; and Lepidoptera 1.5% (Table S2).

The results of rarefaction curves indicated that sampling effort was sufficient to achieve from 60 to 91% completeness at the levels of individual haplotypes and lineages at 3% and 5%, for four of the six taxa (Diptera, Collembola, Araneae, Hymenoptera, Coleoptera, and Lepidoptera, Figure S3). Diptera, Collembola, Arachnida, Coleoptera, Hymenoptera and Lepidoptera presented significant differences in richness and structure among sampling treatments, but patterns were not consistent across groups (Figure 2, 3).

Managed forests

In the dataset including chaponeo vs conserved forests (CON) we found from 2 to 130 haplotypes and from 2 to 20 lineages at 3%, depending on the order of arthropods (Figure 2). Of all arthropod orders, only Araneae and Coleoptera showed statistical differences in richness between the chaponeo and conservation treatments, both at the haplotypes, and lineages at 3% and 5%. We detected higher levels of turnover between sites in chaponeo treatment than in conservation (Figure 5). Low to moderate beta diversity was found across communities in all orders, from haplotypes to higher hierarchical levels (Figure 5; Figure S3). The results of differences in species diversity and composition with ordination plots of NMDS showed that the chaponeo activity and conservation forest have distinct assemblages, with each treatment forming a separate group (Figure 5). The Anosim showed that the Arthropoda community differed significantly between the chaponeo vs conserved forests based on Bray-Curtis dissimilarity (r^2 = 0.15, P < 0.01; r^2 = 0.36, P < 0.001 at haplotypes level). Collembola and Lepidoptera groups have significant differences in community structure from haplotypes to lineages 3% and 5%, and Diptera has significant differences at the haplotype level. In the four orders (Diptera, Collembola, Arachnida, and Hymenoptera) where isolation by distance analyses were carried out, we found that decay of similarity of communities decreases with spatial distance at all levels of lineage divergence (from 0.5 to 7.5%) (Figure 9c, 10c, 11c, 12c). Distance decay was low in all orders (Figure 6, Figure S7), but it is more marked in groups with more limited dispersal abilities (e.g., Collembola and Arachnida) than in groups with higher dispersal abilities (e.g., Diptera and Hymenoptera). Interestingly, these patterns occur even at fine (<1.5 km) geographic distances (Figure 9c; 10c; 11c, 12c).

In the comparison including *A. religiosa* managed forests that were cut 1, 7 or 10 years ago vs conserved forests (CutY1, CutY7, CutY10, CON, respectively) our results showed a richness of 2 to 120 haplotypes and from 2 to 20 lineages 3%, depending on the order of arthropods (Figure 3). Most of the groups yielded high richness within communities (alpha diversity by sample), and only the Diptera, Collembola, Araneae and Hymenoptera groups presented significant differences in richness among different cut years and conservation (Figure 3). Likewise, we detected higher levels of turnover between sites in all treatments than in conservation forests. Low to high beta diversity was presented across communities in all orders, from haplotypes to high hierarchical levels (Figure 6; Figure S3), with significant differentiation among cut sites vs conserved forests. Ordination plots showed that CutY1, CutY7, CutY10 and conservation have clearly distinct assemblages, with each treatment type showing separated groups for the six arthropods orders (Figure 6). Specifically, CutY1, CutY10, and conservation forests were distinct from CutY7 forests in Collembola, Araneae, Diptera, Coleoptera and Hymenoptera. The Anosim detected that the Arthropoda community differed significantly across the cut vs conserved forests based on Bray-Curtis dissimilarity (r^2 = 0.08, P < 0.01; r^2 = 0.67, P < 0.001). We found isolation by distance (IBD), in even less than three km (Figure 9a, 10a, 11a, 12a). The IBD occurs in all groups but was more marked in Collembola and Arachnida (Figure 10a. 11a).

Disturbed forests

In the sites that include disturbed vs conservation treatments, we observed that the richness presented from 2 to 150 haplotypes and from 2 to 20 lineages at 3%. Diptera's lineages at 3% have significant differences in richness between disturbed and conserved forests. However, the richness of all other arthropod groups showed non-significant differences

between the disturbed treatment versus conservation. The estimation of beta diversity showed that the total beta diversity ranged from 0.13 for Lepidoptera to 0.8 for Collembola (Figure S1). For each taxon, community composition variation was explained mainly by turnover (βsim; species replacement between sites), rather than nestedness (βnes; species loss from one site to another) (Figure S2). Collembola, Araneae, Coleoptera and Diptera at disturbed forests have higher levels of turnover between sites than conservation forests. The former holds from haplotypes to lineages at 3% and 5%. However, patterns of species turnover varied by the disturbed treatment and geographical distance. Ordination plots showed that the disturbed and conserved forests presented different groups assemblages. This NMDS and Anosim analysis showed a high-resolution power of arthropod communities to distinguish the different sampling sites of treatment (Figure 8). Our study demonstrates a distance-decay pattern of community similarity in conserved and disturbed forests. We found isolation by distance (IBD), in even less than 2.5 km (9b, 10b, 11b, 12b). This occurs in all groups but was more marked in Collembola and Arachnida (Figure 10b. 11b).

Reforested forest

Overall, richness presented from 2 to 80 at haplotypes level and from 2 to 30 lineages at 3%. Richness α-diversity was significantly higher in the forest under conservation than in reforested forests for Collembola, Araneae and Hymenoptera, but there were no differences between forest reforested in 1980 versus 2013, and this increase did not affect overall α-diversity patterns in other groups (Figure 4). In addition, we detected higher levels of turnover between sites in conservation forests than in reforestation (both 1980 and 2013) for Diptera, Collembola, Coleoptera, Arachnida, and Hymenoptera (Figure 7). Similarly, higher levels of turnover were detected in both reforestation treatments (Figure S1). Ordination plots

showed that the forest under conservation have distinct assemblages, with each reforestation date forming a separate group (Figure 7). Interestingly, forests reforested two years before the sampling were also distinct from forests reforested 35 years ago. Pattern that holds from haplotype to lineages 3% and 5%. The Anosim showed that the Arthropoda community differed significantly across the reforested and conservation forest sites based on Bray-Curtis dissimilarity (r^2 = 0.46, P < 0.001; r^2 = 0.87, P < 0.001). Finally, across reforested sampling sites (max distance 8 km), pairwise similarity within communities at all levels of clustering decreased with geographic distance (Figure 9d, 10d, 11d,12d). The fit of the IBD was high in Collembola (from r^2 = 0.404, b = -2.07, p < 0.001 at the haplotype level, Table S3; Figure 10d). Similar IBD results were found for Diptera, Arachnida, and Hymenoptera (Figure 9d, 11d, 12d; Table S2, Table S3).

Discussion

The wcMBC recovers species richness and composition of diverse arthropod communities, making it a reliable tool for assessing biodiversity across sites. Although turnover patterns were slightly similar across studied treatments, the communities in each treatment were distinct and non-overlapping. Community composition (as represented by NMDS) exhibited clear differences among treatments and between conservation forest and other treatments, as confirmed by the Anosim test. Our results indicate that distance decay is significant even within a maximum distance of only 8 km between treatments, and that forest management treatments have an effect. These patterns are observed across multiple hierarchical levels, from haplotypes to higher lineages, suggesting that dispersal limitations and human activities drive community composition. This pattern holds even at finer geographic distances (<1.5

km), but only for groups with low dispersal ability, such as non-winged Collembola and Arachnida, in the chaponeo treatment.

Haplotypic richness differences between treatments and forest conservation

Tropical mountains are important centers for the generation and maintenance of biodiversity owing to their ecological and evolutionary variables (Fjeldså et al., 2012; Rahbek et al., 2019). In particular, the Nevado de Toluca is of great ecological and evolutionary significance because it (i) encompasses forests that exist in large, relatively continuous sections, (ii) hosts genetically diverse tree populations, and (iii) comprises forests located in an area of long-term climatic stability, a feature that promotes the accumulation of genetic diversity (Mastretta-Yanes et al., 2014).

A major challenge in forest conservation is to quantify the impact of new public policies on biodiversity, particularly in regions undergoing changes in land use (Wang et al., 2019). In an extreme case, forest management can lead to the complete replacement of dominant tree species with other more economically valuable species, resulting in significant changes to the habitat (Brunet et al., 2010). In a less severe scenario, forest management can affect forest-dwelling arthropod communities either directly, by reducing population size during harvesting activities, or indirectly, by affecting habitat availability (e.g., shelter, overwintering structures), habitat heterogeneity, or prey availability (Lange et al., 2014).

Although the forests and natural grasslands of Nevado de Toluca have been subject to deterioration due to agricultural expansion (which occurred before 1972), grazing, mining activities for construction materials, and legal and illegal logging throughout its history (Mastretta-Yanes et al., 2014), the impact of forestry activities on forest-dwelling animal communities had not yet been studied. In this study, we aimed to investigate whether forest

management practices affect the richness and composition of animal communities in different types of forests, including those undergoing management, reforestation, and conservation, using wcMCB data. We found significant differences in arthropod community richness at the haplotype level, and low or none differences at the species level (lineages 3% or GMYC) for most groups.

We observed variation in haplotype richness between different years after logging and conservation treatments, with most arthropod groups showing differences (Figure 3). In the reforestation treatment, Collembola had slightly increased haplotype richness compared to the forest under conservation, while Diptera, Araneae, and Hymenoptera showed variation between treatments. The richness of arthropod haplotypes in the reforestation treatment was only half that of other *A. religiosa* treatments, such as forest management, disturbance, and conservation (Figure 4). Notably, Araneae and Coleoptera showed high variation in haplotype richness in the conservation chaponeo treatment, while only Diptera with 3% lineages showed slightly more variation in disturbed forests than conserved forests. These results indicate that studies looking for changes at the species level (3% linages) may overlook a loss of diversity that may be already occurring at the genetic level. In other words, while species may still be present thus not affecting patterns of genetic diversity, population bottlenecks within species may be occurring, which get reflected at the haplotypic level. This is congruent with studies in many animal taxa showing that genetic diversity is decreasing within species due to human activities (Allentoft & O'Brien, 2010; Des Roches et al., 2021; Garner et al., 2005). Case studies in arthropods have also found a reduction of genetic diversity due to habitat fragmentation, environmental modifications and their consequential reduction of local population sizes, affecting more in particular specialist species (Ferreira-Neto et al., 2017; Gaublomme et al., 2013; Ortego et al., 2015).

Although the signal was weaker than at the haplotype level, we also found some differences in richness at 3% lineages and higher levels. Interestingly, we observed slightly higher arthropod richness in treatments under forest management, disturbance, and reforestation compared to conservation for some groups. In other studies, native forests and mixed plantations have been found to have the highest levels of arthropod diversity, particularly in regions where they mainly consist of small-scale monocultures planted in a checkerboard pattern (Wang et al., 2019). In our study, sites under forest management were located close to forests under conservation, which facilitated the dispersal of individuals from the populations of origin and the impact of the forest on microclimate (Arribas et al., 2020; Baker et al., 2014).

The richness of most groups of arthropods and treatments did not show significant differences. In contrast, previous studies have suggested that arthropod richness can be a reliable predictor of plant richness or vice versa in tropical forests (Basset et al., 2012; Lynggaard et al., 2020; Zhang et al., 2016). The lack of significant results in richness for some treatments and groups may indicate that species richness is highly sensitive to the wcMBC bioinformatics strategy (Liu et al., 2020). For instance, Barsum et al. (2019) cautioned about the transferability of these taxa-specific responses in different spatial and temporal contexts, as they found no significant differences in spider species richness between stand types at all sampling intervals. Therefore, the current focus of biodiversity studies is on changes in community composition rather than species richness (Lynggaard et al., 2020; Magurran, 2016; Wang et al., 2019). Consistent with this, Lynggaard et al. (2020) reported that species richness appears to recover rapidly after revegetation, and hence, changes in community composition should be the main focus. The argument is that anthropogenically disturbed communities can maintain species richness and phylogenetic diversity, even when

local or endemic species go extinct and are replaced by cosmopolitan species (Wang et al., 2019). Moreover, changes in alpha diversity may not be noticeable due to taxa substitution (Lynggaard et al., 2020). To address this issue, several filter steps have been employed to remove false OTUs (Andújar et al., 2021), and it has been proposed to use "iNextPD" to generate robust alpha diversity comparisons by estimating phylogenetic diversity instead of species richness (Wang et al., 2019).

Significant turnover in the community composition between treatments and forest conservation

Community composition and structure are key indicators of changes in biodiversity (Lynggaard et al., 2020; Wang et al., 2019). Changes in species structure and composition can be attributed to both natural and anthropogenic causes (Lynggaard et al., 2020; Magurran, 2016; Wang et al., 2019). Our results support this hypothesis, as forest management, disturbance, and reforestation had marked effects on arthropod groups' structure and composition. Specifically, forests regenerating 1, 7, and 10 years after logging were significantly different from those under conservation for all groups. Notably, there was no overlap between sites under conservation and regeneration forest in the ordination (NMDS), indicating a significant separation between arthropod communities in forests under conservation and at different times after logging. This finding is consistent with studies of beetle community composition in other forests, where the authors examined regenerating forests between 40 and 58 years ago and found that forests remained significantly different from mature forests after clearcutting 40 and 58 years ago (Liu et al., 2021). Similarly, there was a response of arthropod composition to land use change in the tropics (Beng et al., 2016). Even subtle differences in forest beetle communities after disturbance have been observed

(Liu et al., 2020). These findings can be explained by the higher diversity of vegetation and variation in height and three-dimensional structure, allowing for greater sunlight penetration into the understory, resulting in the availability of food and other resources (Wang et al., 2019). Moreover, changes in forests due to anthropogenic activities translate into changes in ecosystem functioning in forests due to changes in the arthropod community (Brandon-Mong et al., 2018).

In our study, we found that chaponeo, and disturbances had no effect on the structure and composition of some groups of arthropods. Our results indicate that the structure and composition of the arthropod community in logged forests differed significantly from that in conservation forests for Collembola, Lepidoptera and Diptera (at the haplotype level). However, there was a high degree of overlap between sites under conservation and management in the ordenation (NMDS), suggesting that the communities of Arachnida, Hymenoptera, Coleoptera, and Diptera (3% and 5% lineages) showed an incipient but non-significant separation between in logged forests and forest under conservation. Liu et al. (2020) reported subtle but significant differences in beetle community composition in the renewal of the arthropod community in regenerating very late successional forests, approximately 55 years old, and mature unlogged forests. Interestingly, one mechanism that may drive arthropod composition is the effect of light availability on understory vegetation and microclimate (Cai et al., 2021).

Mixed plantations can provide increased diversity and composition of arthropods being highly correlated with native forests (Wang et al., 2019). The effects of reforestation in Mexican forests show that the structure and composition do not overlap, and this separation was significant for all groups. This is similar with one study, where croplands support an arthropod community similar in richness and diversity to mixed plantations and just below

native forests, but where the species composition of cropland is different from native forest, and cropland cannot compensate for the loss of biodiversity that native forest depends on (Wang et al., 2019). In addition, the composition and structure of plants (or vice versa) in tropical landscapes (Basset et al., 2012; Lynggard et al., 2020; Zhang et al., 2016), especially in native forests, the high diversity of arthropods prevails (Lynggard et al., 2020; Zhang et al., 2016; Wang et al., 2019).

The composition of arthropod communities exhibited considerable variation across treatments and arthropod groups. Additionally, a correlation was observed between geographical distance and changes in species composition, indicating that forest management, disturbances, reforestation, and distance work synergistically to shape biodiversity and community composition. This is consistent with previous findings which have established strong correlations between environmental gradients and changes in species composition across forest types, suggesting that differences in β diversity could be driven by interactions between forest type, environmental heterogeneity, and the life history of species (Beng et al., 2016). Moreover, dispersal constraints have been shown to play a critical role in shaping community structure across multiple hierarchical levels at fine-scale geographical features, within a habitat-based framework (Arribas et al., 2020; Noguerales et al., 2021; Gálvez-Reyes et al., 2021). Therefore, it is crucial to examine subtle differences in community composition to gain a better understanding of the underlying mechanisms driving biodiversity turnover in communities and to inform adaptive management strategies (Barsoum et al., 2019). For instance, bark beetle-induced forest dieback does not result in detectable differences in species diversity but does result in changes in community composition (Cai et al., 2021). Hence, heterogeneous forests are more resistant to pests and diseases and harbor higher diversity (Cai et al., 2021).

Forest management treatments and dispersal limitations are driving community structure

To investigate whether dispersal constraints shape community structure, we evaluated the dispersal distance relationship (DDR) from haplotype levels to communities in different forest management and reforestation activities. We calculated isolation by distance (IBD) for four orders with contrasting dispersal capacities: Diptera, Collembola, Hymenoptera, and Arachnida, and considered all our sampling locations in Nevado de Toluca, with a maximum distance of 8 km between treatment sampling points and a minimum distance of 50 m within plots. We found that dispersal limitation plays an important role within each treatment because the similarity of communities decreases with spatial distance from the haplotype level to all levels of lineage divergence (from 0.5% to 7.5%), and the same pattern was observed with molecular putative species GMYC (Figures 9, 10, 11, and 12). The decay of similarity of arthropod communities across distance was high in all orders and treatments (Figures 9, 10, 11, and 12), but it was more marked for all groups within the reforestation treatment (Figures 9d, 10d, 11d, and 12d). Similarly, it was more pronounced in groups with more limited dispersal abilities, such as Collembola and Arachnida, than in groups with higher dispersal abilities, such as Diptera and Hymenoptera. Interestingly, for the Collembola and Arachnida group, a marked decay of similarity was found within forest management with chaponeo at less than 1.5 km. This is consistent with other studies showing that dispersal limitations drive community structure across multiple hierarchical levels at fine geographical scales within Collembola and Arachnida, resulting in community differentiation at very short geographic distances (<2 km) within forests (Arribas et al., 2020; Gálvez-Reyes et al., 2021).

Baselga et al. (2012) suggest that dispersal capacity is a key factor driving community similarity at multihierarchical levels. Recent studies have demonstrated that dispersal restrictions play a critical role even at relatively small spatial scales of a few kilometers in

soil microarthropods (Arribas et al., 2020). Consistent with these findings, our results show that the decrease in similarity with increasing distance is more pronounced in wingless Collembola (r^2 = 0.404 at the haplotype level; Table S3; Figure 10) than in winged Diptera (r^2 = 0.15 at the haplotype level), within forests at finer geographic scales <1.5 km. Similar patterns have been reported in arthropods inhabiting conservation forests, where the distance decrease is greater in wingless Collembola (r^2 = 0.704 and r^2 = 0.599) at the haplotype and GMYC levels than in winged Diptera (r^2 = 0.293 and r^2 = 0.195) at the haplotype and GMYC levels (Gálvez-Reyes et al., 2021). However, as dispersal capabilities become limited, landscape features that restrict dispersal are expected to play a more prominent role in shaping diversity (Vellend, 2010), including human modifications to the ecosystems. Thus, different land uses are likely to have varying impacts on community dissimilarity (Barsoum et al., 2019; Wang et al., 2019). Our results support that the decrease in similarity is given by treatment for example Chaponeo. But also, within in the dissimilarity is greater in wingless Collembola (r^2 = 0.24 at the haplotype level; Table S3; Figure 10) than in winged Diptera (r^2 = 0.16 at the haplotype level; Table S3; Figure 9) at finer geographic scales <1.5 km. This suggests that there is a synergistic effect between distance and treatment in shaping the composition of most groups. Finally, we observed that the proportion of Diptera was significantly affected by the different treatments. Our observations are consistent with those of Brandon-Mong et al. (2018) and Zhang et al. (2016), who also reported similar proportion of Diptera in anthropogenic forests. Furthermore, Brandon-Mong et al. (2018) emphasized the positive feedback mechanism between land use and temperature, which poses threats to arthropod communities and has implications for ecosystem functioning and human well-being.

Opportunities and limitations of metabarcoding in measures of biodiversity monitory

The wcMBC of arthropods is a valuable tool for evaluating biodiversity from bulk samples, which can be used to inform management practices and policy development (Ji et al., 2013; Yu et al., 2012). High-quality wcMBC data may contribute to global efforts to generate sequence data for many species on Earth (Creedy et al., 2019; Ji et al., 2013), particularly in poorly understood, diverse ecosystems such as tropical mountain forests (Gálvez-Reyes et al., 2021). Our results revealed six taxonomic groups that are sensitive to different forest management and reforestation activities, from haplotype levels to multiple hierarchical lineages. Similar results were obtained when evaluating alpha, beta, and gamma diversity of soil mesofauna in forests and grasslands (Arribas et al., 2020), within forest types (Noguerales et al., 2021). Other studies have evaluated the richness and composition of arthropods in native forests, forest plantations, monocultures, disturbed and restored forests (Barsoum et al., 2019; Beng et al., 2016; Liu et al., 2020; Lynggaard et al., 2020; Wang et al., 2019).

However, our focus on wcMB of bulk samples has certain limitations when it comes to obtaining accurate ASV-based data. The success of metabarcoding studies for assessing invertebrate diversity in bulk samples depends on various methodological factors, such as sampling design, sampling method, bulk-sample cleaning, DNA extraction, primer selection, library tagging by PCR (to avoind tag-jumps), sequencing coverage, bioinformatic methods, and database reference (Bohmann et al., 2021; Creedy et al., 2019; Liu, Clarke, et al., 2020; Elbrecht et al., 2017, 2019; Elbrecht & Leese, 2015; Krehenwinkel et al., 2017, 2018; Piñol et al., 2015; Schnell et al., 2015; van der Loos & Nijland, 2021). Key sources of error at the wet-lab stage can strongly influence the absolute and relative amounts of invertebrates detected (Elbrecht et al., 2017). For instance, according to Creedy et al. (2019), the size of

the arthropods can have implications for canopy arthropod studies, but this can be overcome by ensuring sufficient sequencing depth. It is necessary to overcome challenges related to size and biomass control, and to consider the sampling period, which can indicate compositional changes over time (Barsoum et al., 2019; Creedy et al., 2019; Elbrecht et al., 2017). The lack of a comprehensive reference database makes taxonomic assignment to operational taxonomic units (OTUs) difficult (Elbrecht & Leese, 2017; Liu et al., 2019; Liu et al., 2020). To avoid generating faulty data, researchers should consider reviewing methodological options and applications to improve the accuracy and reliability of results. The wcMB approach captures fine-scale temporal variations in the composition of arthropod communities, emphasizing the importance of controlling for temporal effects in sampling. In our study, the sampling periods were two weeks apart.

Implications for forest management in the Nevado de Toluca

Our data evaluate the richness and composition of arthropods in relation to forest management and reforestation activities carried out within the *A. religiosa* forest and reforested *P. hartwegii* forests. We suggest the establishment of a monitoring program that integrates forest conservation and the monitoring of arthropod diversity and composition within the forests of Nevado de Toluca. Such a program would allow for the assessment of the effectiveness of forest management practices and the identification of any potential threats to the arthropod communities within the protected area (Barsoum et al., 2019). Similarly, environmental policies should promote appropriate forest management practices that minimize negative impacts, limit forest plantations to small areas, and maintain a dense and old canopy to ensure suitable habitats for shade-requiring and old forests arthropod specialists. Conservation of forest connectivity is also crucial. It is important to consider that

the response to high temperatures may vary among arthropod orders and forest types (Barsoum et al., 2019; Creedy et al., 2019; Wang et al., 2019) when formulating policies. Conservation strategies for managing biodiversity in tropical high mountains must take into account specific taxonomic differences. In our study, we identified approximately 2,277 haplotypes in the entire arthropod community. It is also important to consider functional diversity to assess the health status of ecosystems (Lynggaard et al., 2020). For instance, springtails (Entomobryidae) are known to play a crucial role in soil rehabilitation, and their large presence in the initial stages is a key factor in the recovery of soil functions (Langmaack et al., 2001). To better detect differences in biodiversity levels, forest monitoring could include temporal assessments of arthropod diversity over time (Barsoum et al., 2019; Creedy et al., 2019).

Finally, it is essential to report on the landscape-scale impacts and habitat connectivity of forest management practices, disturbance, and reforestation to improve monitoring programs. Weather conditions should be considered to refine management practices. The effects of mature forest on harvested areas (1, 7, and 10 years after logging) are likely the result of a combination of interacting features, including the influence of forest, landscape context, logging-related features, landscape setting, and distance. Biodiversity data should be collected to identify forests of conservation concern, detect threats such as climate change and new pests and pathogens, and measure the effectiveness of forest policy measures designed to enhance forest biodiversity (Barsoum et al., 2019; Lynggaard et al., 2020; Wang et al., 2019). To mitigate the deterioration caused by the opening of agricultural fields, grazing within forests, mining for construction material extraction, and legal and illegal logging throughout the history of Nevado de Toluca (Mastretta-Yanes et al., 2014), it is crucial to inform the indices of biodiversity and community composition. Haplotype data

should also be incorporated into biomonitoring as a complement to species-level diversity to promote conservation and forest management strategies.

Conclusion

Biodiversity assessments are a fundamental requirement for ecosystem management. To achieve this, a holistic integration of high-throughput technology and efficient sampling, combined with the COI marker, can be used to interpret the biodiversity of the entire arthropod community. Metabarcoding provides an efficient method for assessing biodiversity samples from forests under conservation, stands under forest management and reforestation, and different geographic distances. These factors drive changes in the richness and composition of arthropod communities. However, the utility of DNA metabarcoding for ecological management and biodiversity monitoring could be enhanced with more comprehensive reference databases (Liu et al., 2020).

To better inform policy discussions and monitoring efforts, it is important to consider the link between the richness and composition of arthropod communities and plant diversity in different forest ecosystems, as well as the impact of climate change on forest dynamics. Promoting interdisciplinary research on indicators of biodiversity composition and forest management can lead to the development of effective biodiversity management guidelines for forest and conservation managers based on recent scientific results. By doing so, management of biodiversity in forest ecosystems can be more assertive.

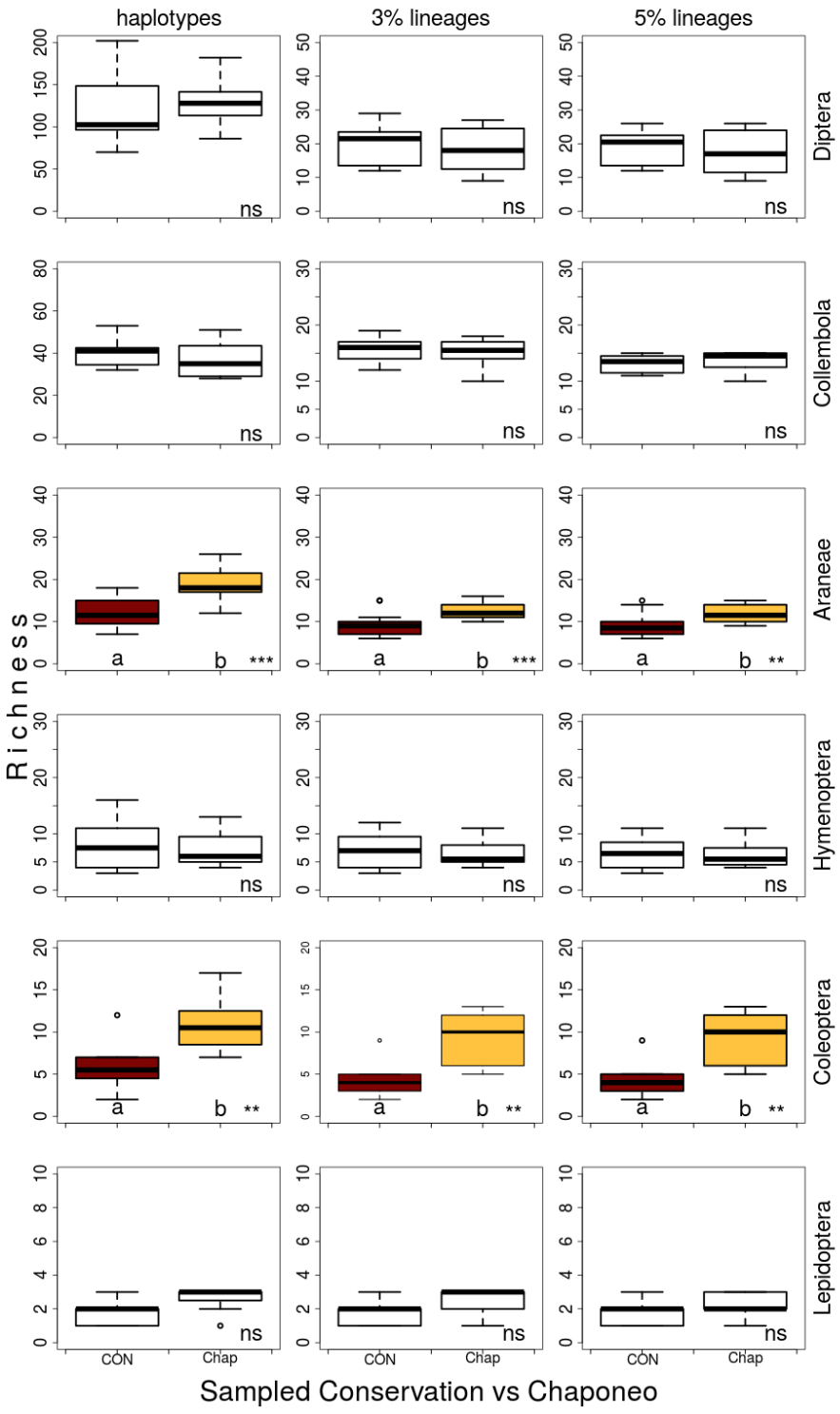

Figure 2. Arthropods richness comparing managed (chaponeo) and conservation treatments. CON = Under conservation (red color), Chaponeo activity= chaponeo (pink color). Same letters indicate that there are no significant differences among those treatments $p<0.05$. Significance codes *:$p<0.05$, **:$p<0.01$, ns: non-significant.

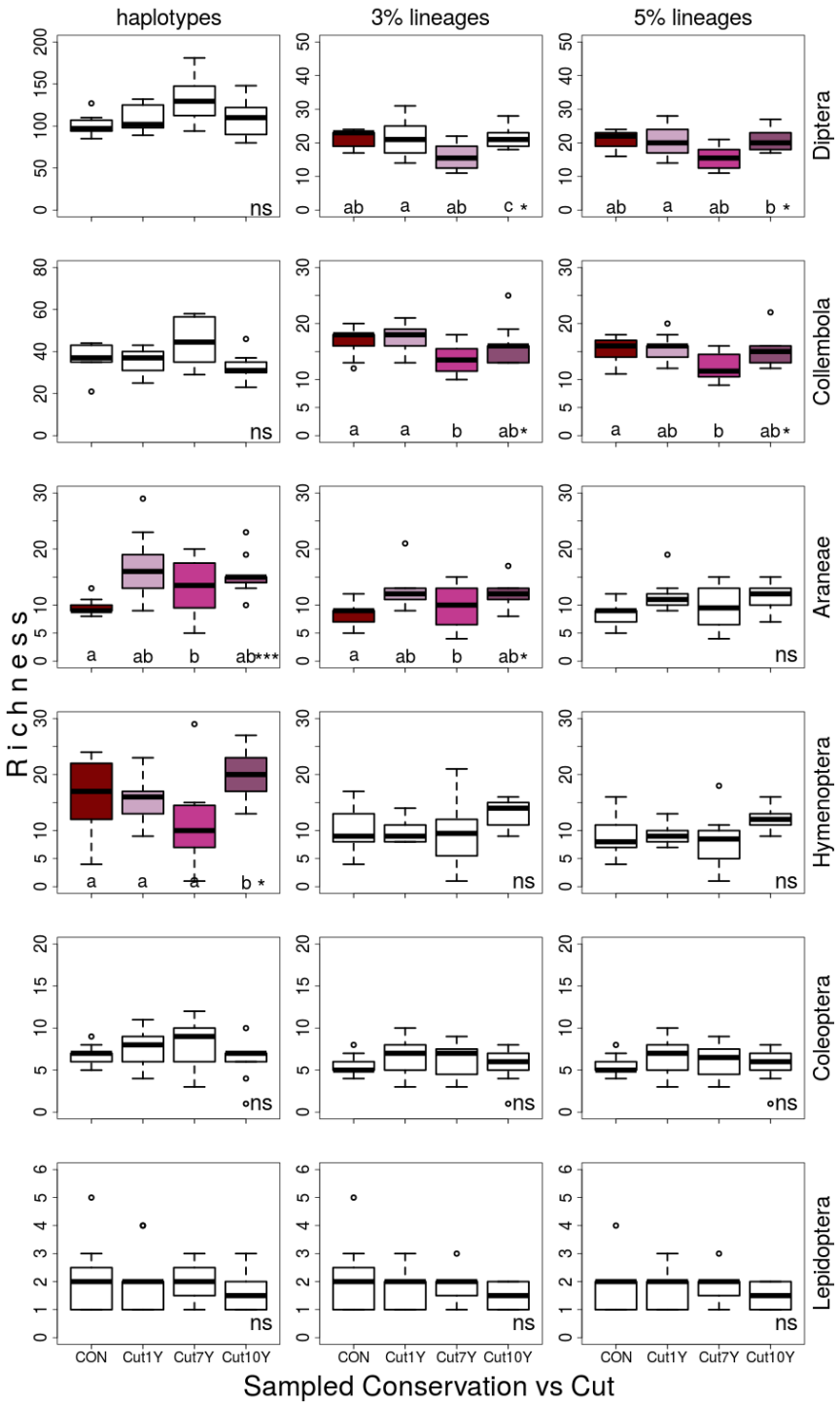

Figure 3. Arthropod richness by sampling cut years before treatment. CON = Under conservation, CutY1= Cut one year before, CutY7=Cut seven years before, CutY10=Cut ten years before. Same letters indicate that there are no significative differences among those treatments $p<0.05$. Significance codes *:$p<0.05$, ***:$p<0.001$, ns: non-significant.

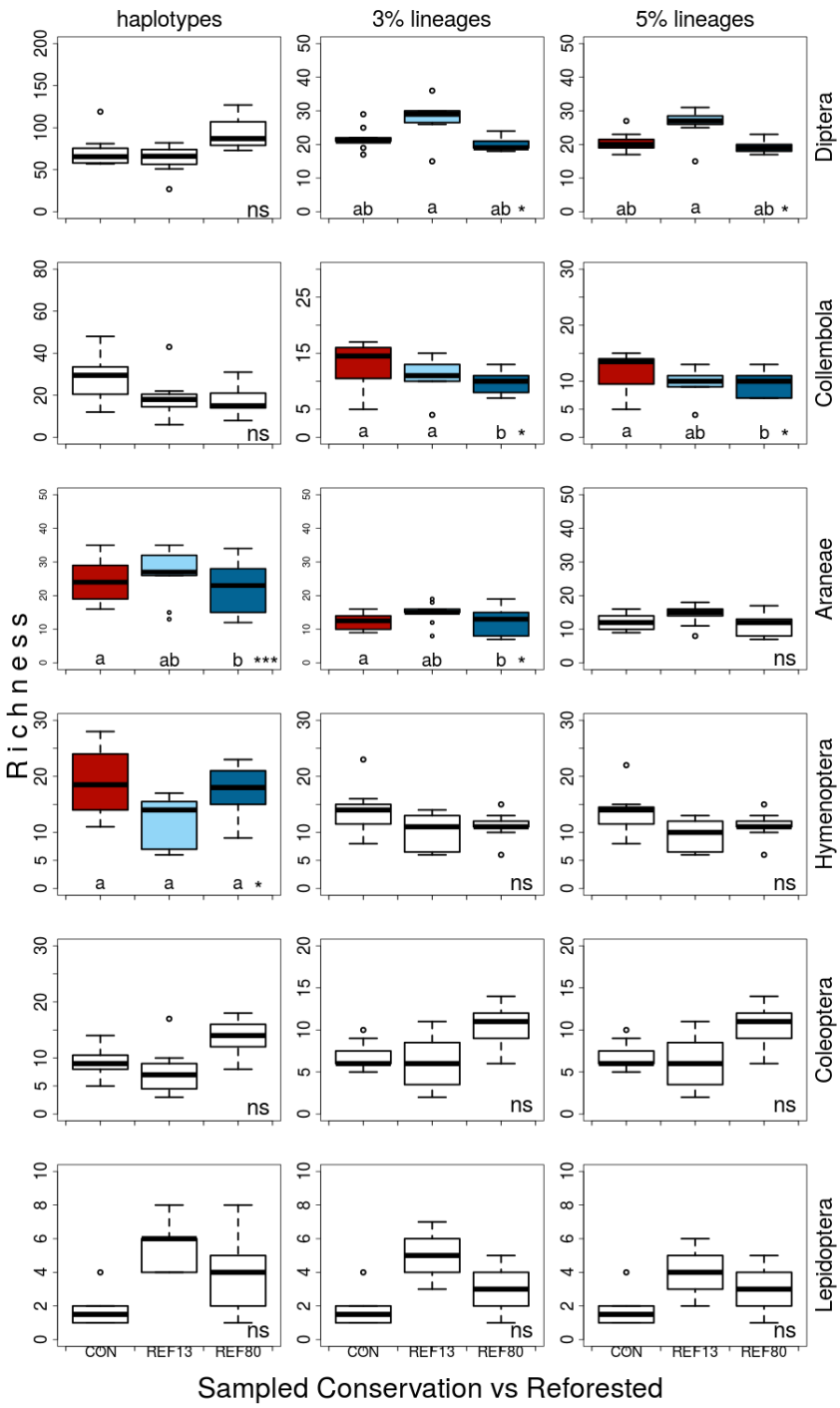

Figure 4. Arthropods richness by sampling reforested treatment. The reforestation forest includes under conservation (CON), reforestation two year ago before the sampling (REF13) and reforestation 35 year ago before the sampling (REF80). Same letters indicate that there are no significative differences among those treatments $p<0.05$. Significance codes *:$p<0.05$, ***:$p<0.001$, ns: non-significant.

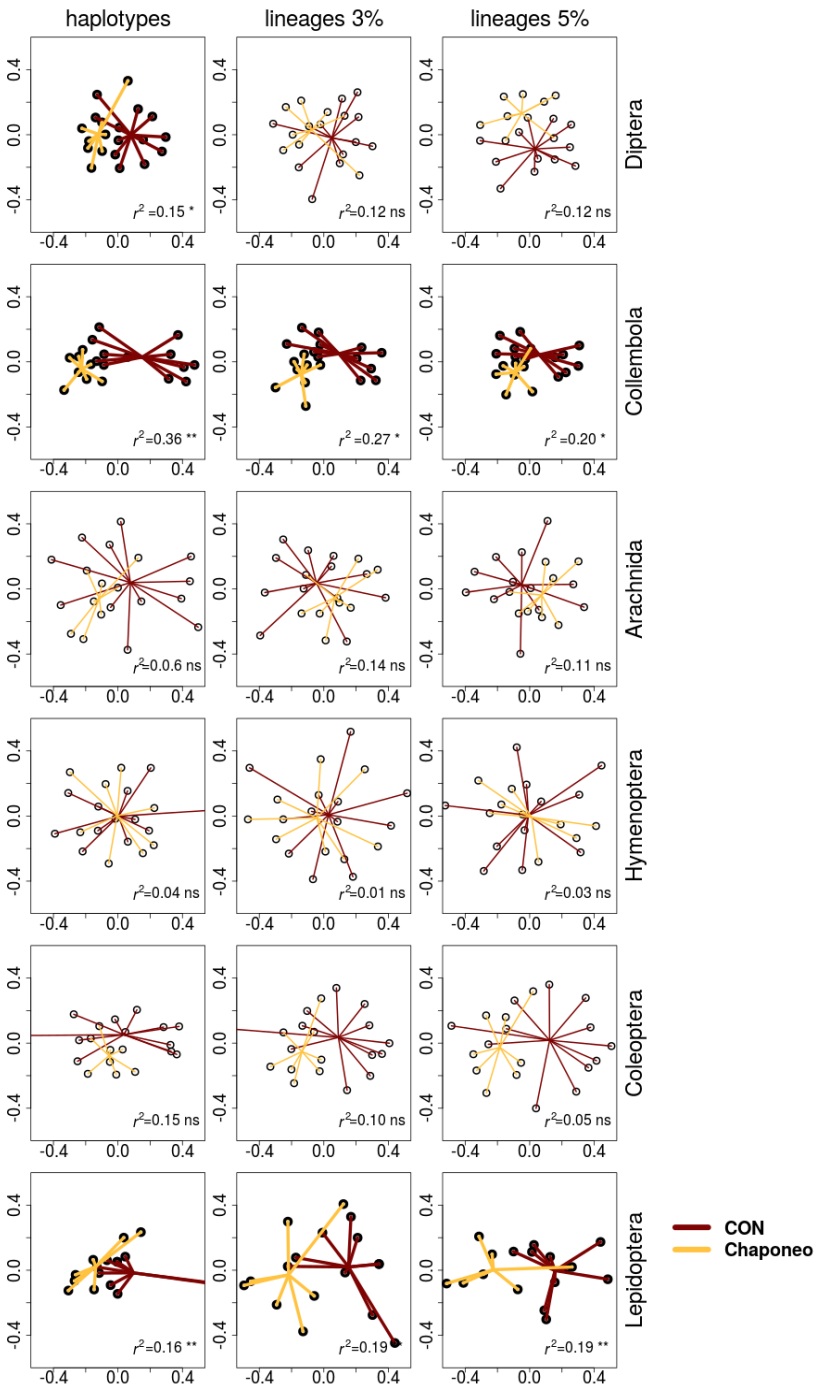

Figure 5. Non-Metric Multidimensional scaling (NMDS) ordinations of community similarity (Simpson index, βsim) in forest under conservation (CON: red color) and chaponeo activity (chaponeo: yellow color) at the haplotype, 3% and 5% lineages of the six taxonomic orders studied. The asterisk represents significant differences of community structure among sites estimated with Anosim at *:p<0.05, **:p<0.01 ***:p<0.001 and ns: non-significant values.

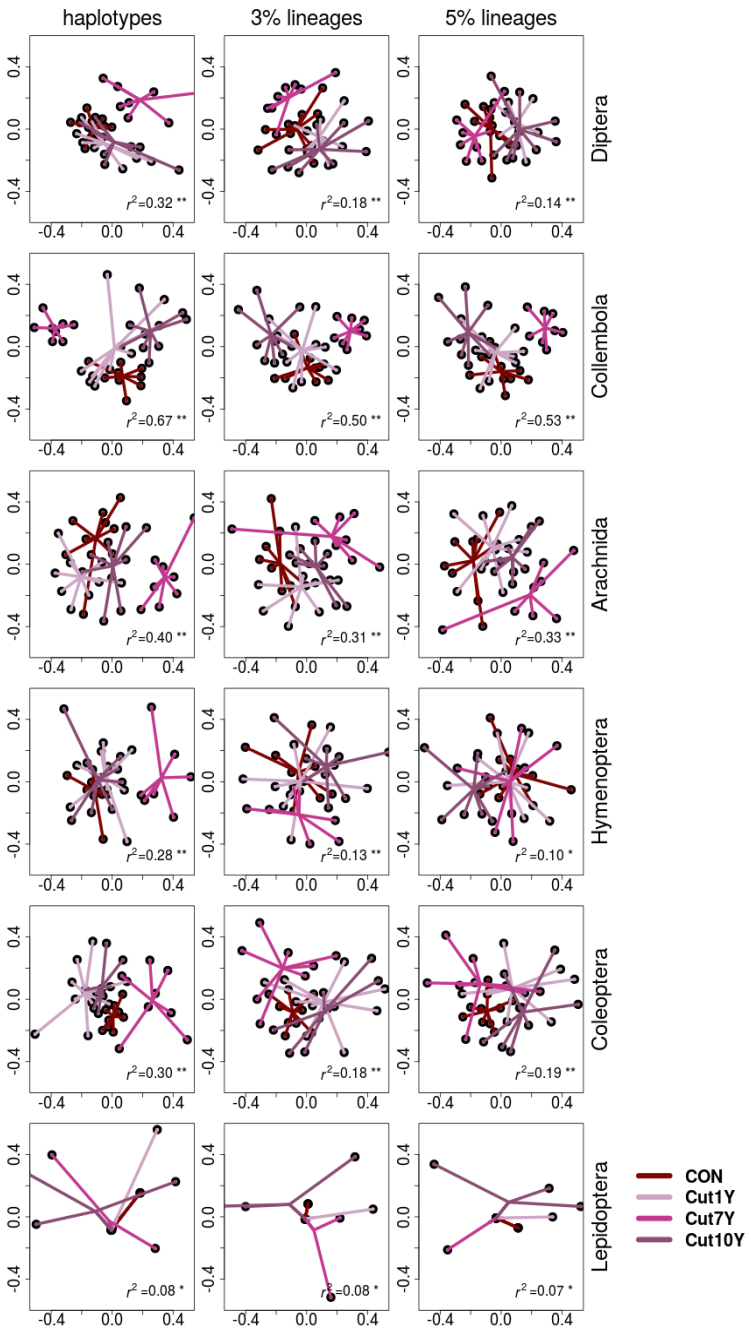

Figure 6. Non-Metric Multidimensional scaling (NMDS) ordinations of community similarity (Simpson index, βsim) in forest under conservation (CON: red color), cut one year before (Cut1Y: pink color), cut seven years before (Cut7Y: purple color), and cut ten years before (Cut10Y: violet color) at the haplotype, 3% and 5% lineages of the six taxonomic orders studied. The asterisk represents significant differences of community structure among sites estimated with Anosim at *:p<0.05, **:p<0.01 ***:p<0.001 and ns: non-significant values.

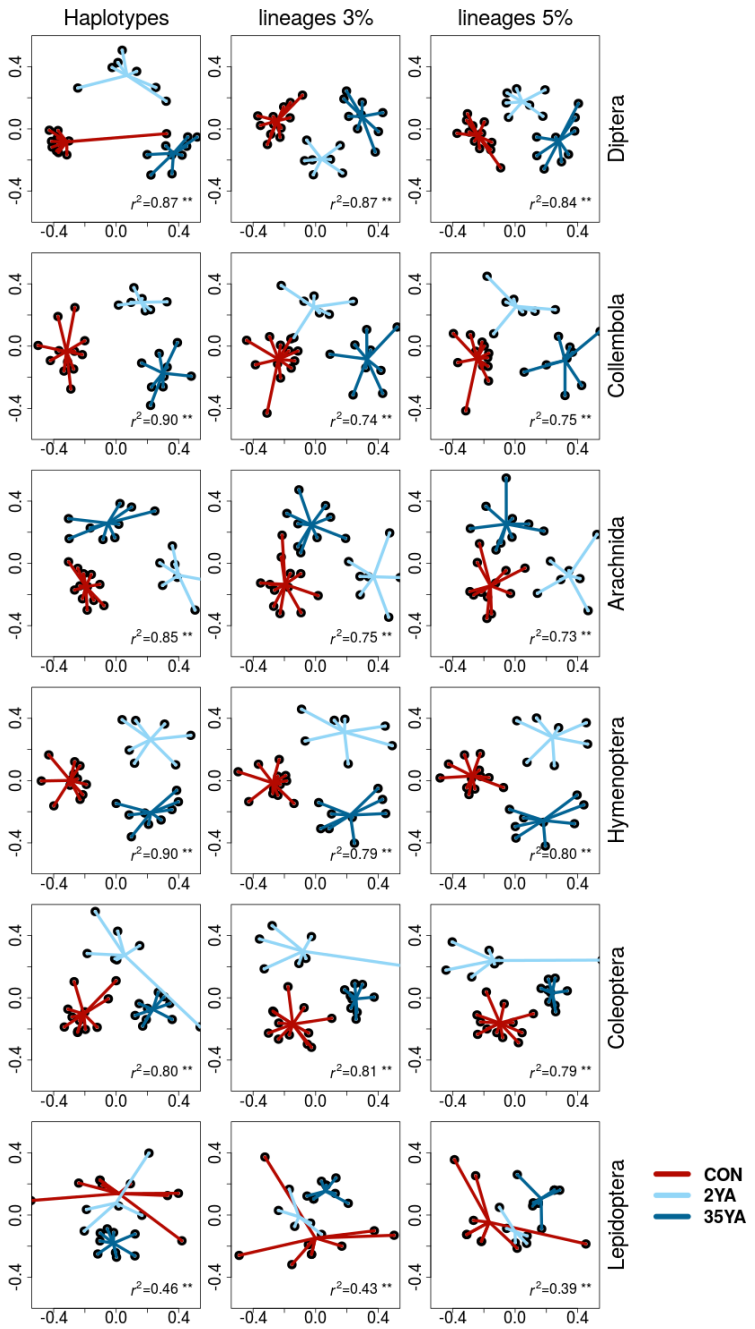

Figure 7. Non-Metric Multidimensional scaling (NMDS) ordinations of community similarity (Simpson index, βsim) in forest under conservation (CON: red color), reforestation two year ago before the sampling (REF13: light blue color) and reforestation 35 year ago before the sampling (REF80: blue color) at the haplotype, 3% and 5% lineages of the six taxonomic orders studied. The asterisk represents significant differences of community structure among sites estimated with Anosim at *:p<0.05, **:p<0.01 ***:p<0.001 and ns: non-significant values.

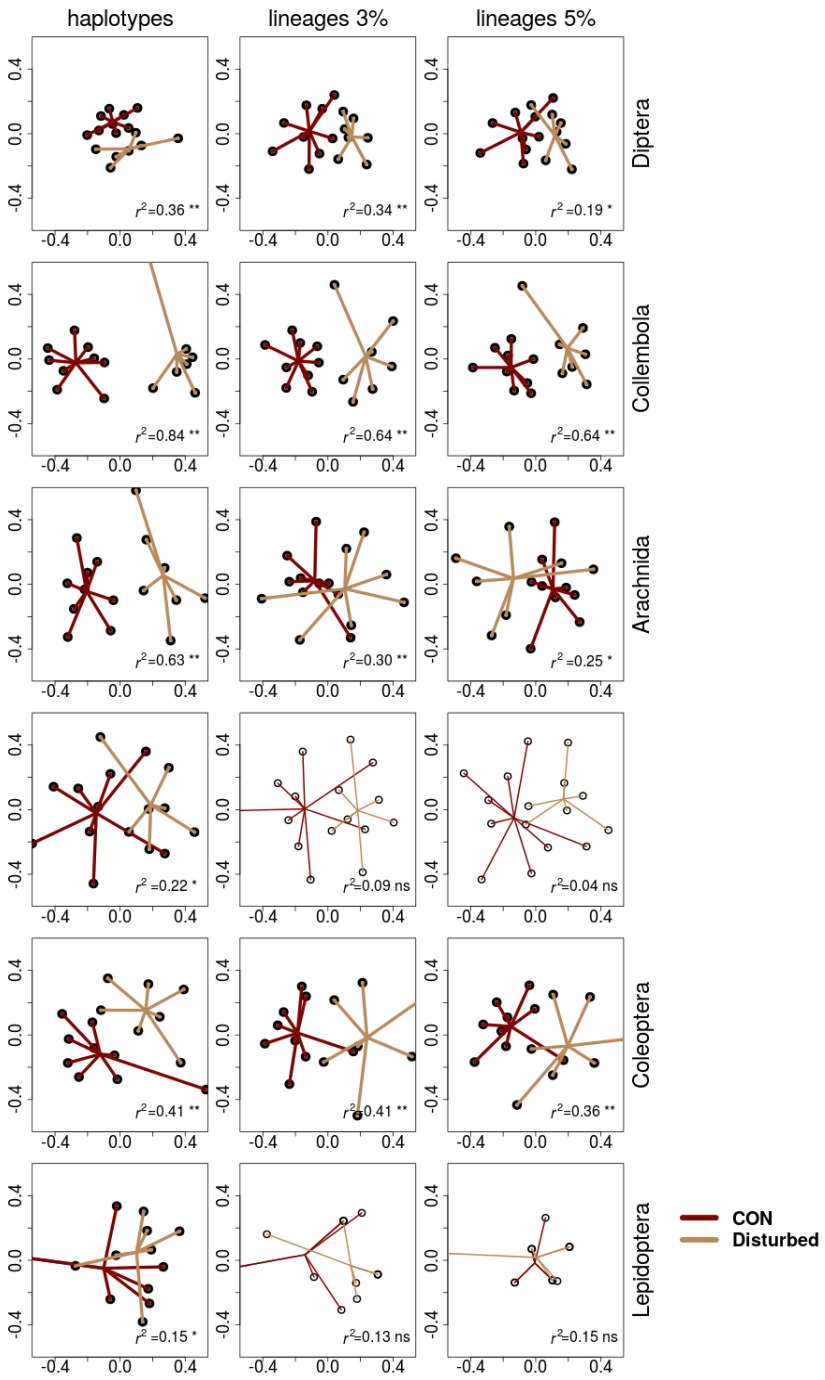

Figure 8. Non-Metric Multidimensional scaling (NMDS) ordinations of community similarity (Simpson index, βsim) in forest under conservation (CON: red color) and forest under disturbed (Disturbed: yellow color) at the haplotype, 3% and 5% lineages of the six taxonomic orders studied. The asterisk represents significant differences of community structure among sites estimated with Anosim at *:p<0.05, **:p<0.01 ***:p<0.001 and ns: non-significant values.

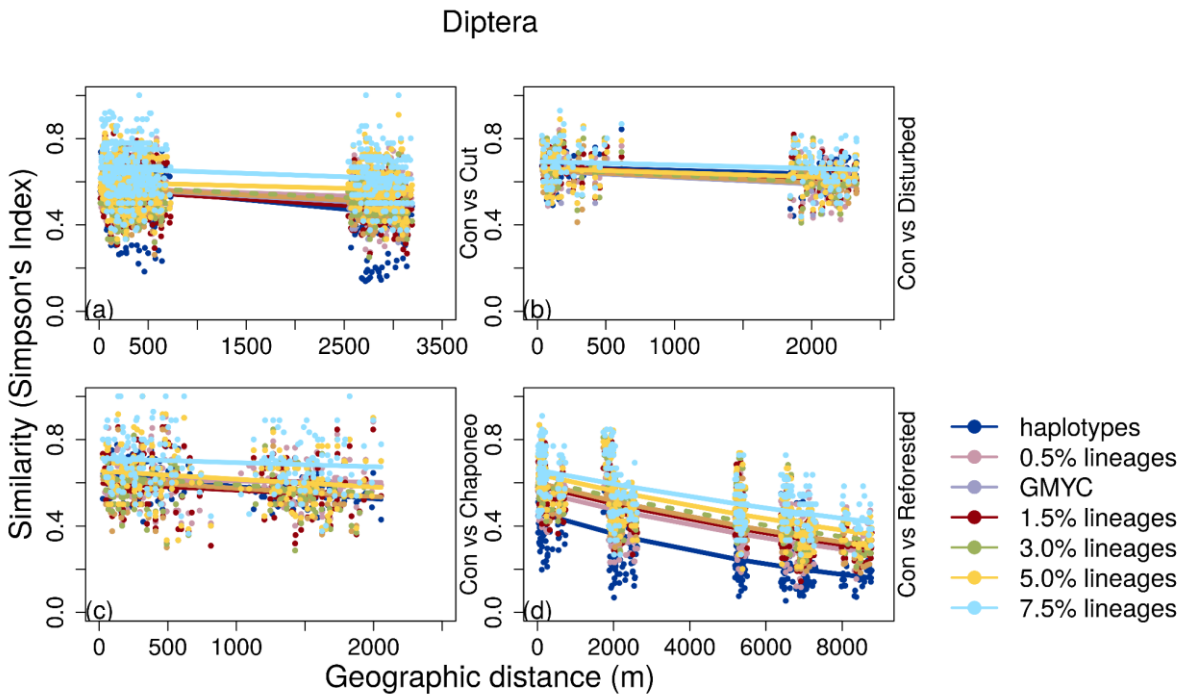

Figure 9. Distance decay of Diptera community similarity at multiple levels of genetic similarity in Nevado de Toluca conserved and managed forests. (a) Decay of similarity is shown against geographic distance for Diptera in *A. religiosa* forest under conservation and forest cut one year before, cut seven years before, and cut ten years before. (b) Decay of similarity versus geographic distance within *A. religiosa* forest under conservation and forest under disturbed. (c) Decay of similarity versus geographic distance in *A. religiosa* forest under conservation and chaponeo treatment. (d) Decay of similarity versus geographic distance in *P. hartwegii* forest under conservation versus reforestation two years ago before the sampling and reforestation 35 years ago before the sampling. Dots represent similarity between pairs of sampling sites and lines are the fitted model for each lineage. The significance levels and r^2 are in Table S3.

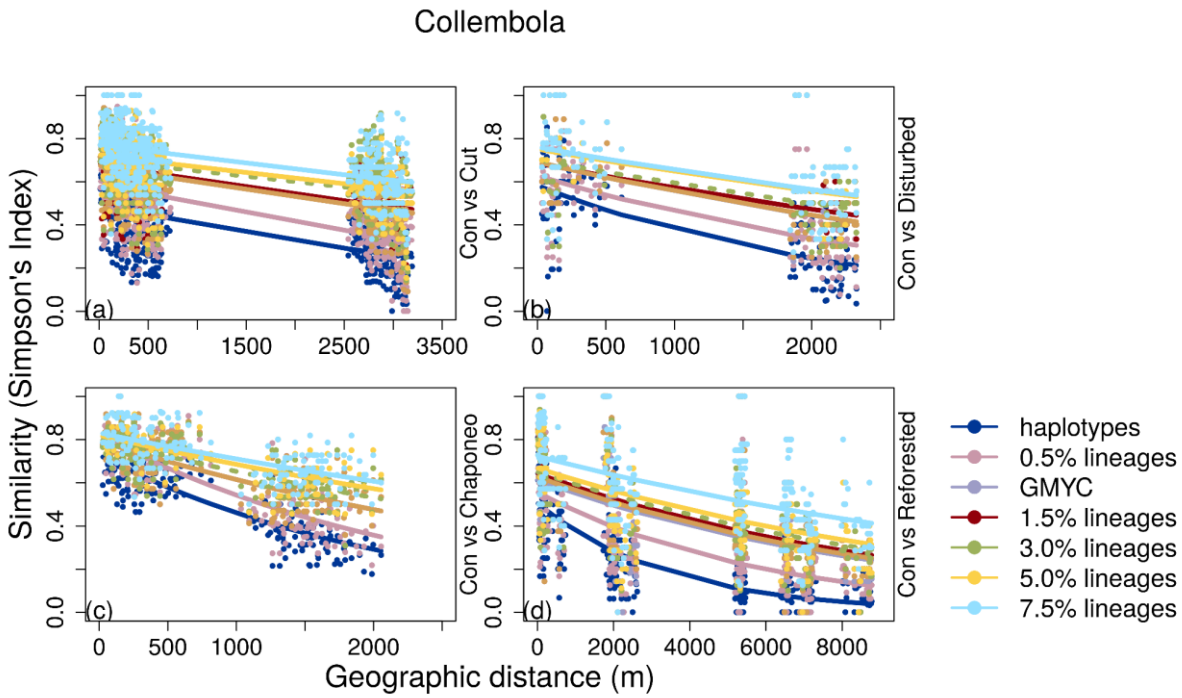

Figure 10. Distance decay of Collembola community similarity at multiple levels of genetic similarity in Nevado de Toluca. (a) Decay of similarity is shown against geographic distance for Collembola in *A. religiosa* forest under conservation and forest cut one year before, cut seven years before, and cut ten years before. (b) Decay of similarity versus geographic distance within *A. religiosa* forest under conservation and forest under disturbed. (c) Decay of similarity versus geographic distance in *A. religiosa* forest under conservation and chaponeo treatment. (d) Decay of similarity versus geographic distance in *P. hartwegii* forest under conservation versus reforestation two years ago before the sampling and reforestation 35 years ago before the sampling. Dots represent similarity between pairs of sampling sites and lines are the fitted model for each lineage. The significance levels and r^2 are in Table S3.

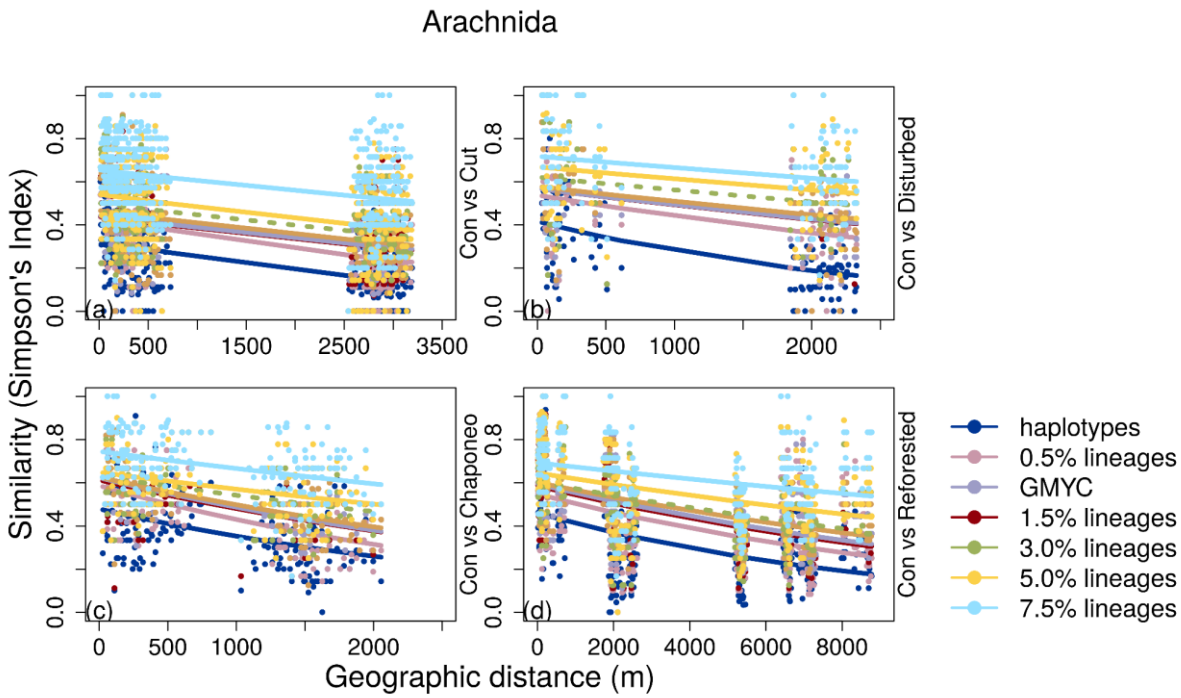

Figure 11. Distance decay of Arachnida community similarity at multiple levels of genetic similarity in Nevado de Toluca. (a) Decay of similarity is shown against geographic distance for Arachnida in *A. religiosa* forest under conservation and forest cut one year before, cut seven years before, and cut ten years before. (b) Decay of similarity versus geographic distance within *A. religiosa* forest under conservation and forest under disturbed. (c) Decay of similarity versus geographic distance in *A. religiosa* forest under conservation and chaponeo treatment. (d) Decay of similarity versus geographic distance in *P. hartwegii* forest under conservation versus reforestation two years ago before the sampling and reforestation 35 years ago before the sampling. Dots represent similarity between pairs of sampling sites and lines are the fitted model for each lineage. The significance levels and r^2 are in Table S3.

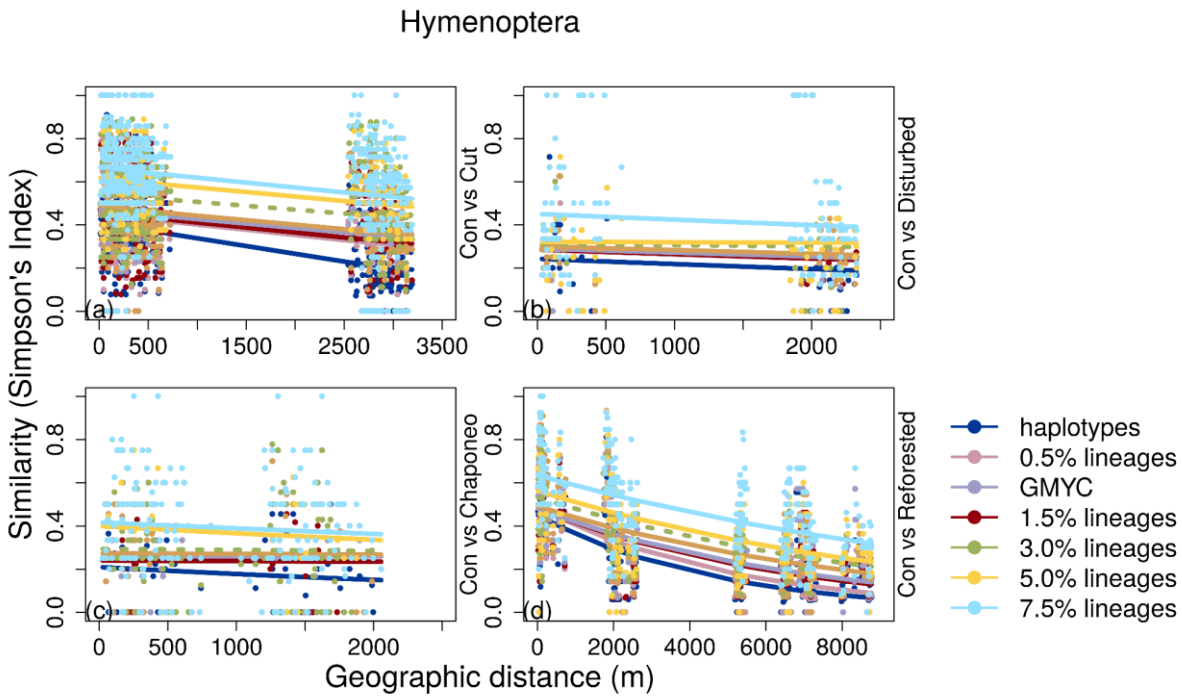

Figure 12. Distance decay of Hymenoptera community similarity at multiple levels of genetic similarity in Nevado de Toluca. (a) Decay of similarity is shown against geographic distance for Hymenoptera in *A. religiosa* forest under conservation and forest cut one year before, cut seven years before, and cut ten years before. (b) Decay of similarity versus geographic distance within *A. religiosa* forest under conservation and forest under disturbed. (c) Decay of similarity versus geographic distance in *A. religiosa* forest under conservation and chaponeo treatment. (d) Decay of similarity versus geographic distance in *P. hartwegii* forest under conservation versus reforestation two years ago before the sampling and reforestation 35 years ago before the sampling. Dots represent similarity between pairs of sampling sites and lines are the fitted model for each lineage. The significance levels and r^2 are in Table S3.

References

Andújar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology*, *27*(20), 3968-3975. https://doi.org/10.1111/mec.14844

Andújar, C., Creedy, T. J., Arribas, P., López, H., Salces-Castellano, A., Pérez-Delgado, A. J., Vogler, A. P., & Emerson, B. C. (2021). Validated removal of nuclear

pseudogenes and sequencing artefacts from mitochondrial metabarcode data. *Molecular Ecology Resources, 21*(6), 1772-1787. https://doi.org/10.1111/1755-0998.13337

Allentoft, M. E., & O'Brien, J. (2010). Global amphibian declines, loss of genetic diversity and fitness: a review. *Diversity*, *2*(1), 47-71.

Arribas, P., Andújar, C., Hopkins, K., Shepherd, M., & Vogler, A. P. (2016). Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. *Methods in Ecology and Evolution*, *7*(9), 1071-1081. https://doi.org/10.1111/2041-210X.12557

Arribas, P., Andújar, C., Salces-Castellano, A., Emerson, B. C., & Vogler, A. P. (2020). The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding. *Molecular Ecology*, *30*(1), 48-61. https://doi.org/10.1111/mec.15591

Babraham Institute. (2013). *Babraham Bioinformatics—FastQC a quality control tool for highq throughput sequence data*. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Back, P., Suominen, A., Malo, P., Tahvonen, O., Blank, J., & Deb, K. (2020). *Towards sustainable forest management strategies with MOEAs*. 1046-1054. https://doi.org/10.1145/3377930.3389837

Baker, T. P., Jordan, G. J., Steel, E. A., Fountain-Jones, N. M., Wardlaw, T. J., & Baker, S. C. (2014). Microclimate through space and time: Microclimatic variation at the edge of regeneration forests over daily, yearly and decadal time scales. *Forest Ecology and Management*, *334*, 174-184. https://doi.org/10.1016/j.foreco.2014.09.008

Barsoum, N., Bruce, C., Forster, J., Ji, Y.-Q., & Yu, D. W. (2019). The devil is in the detail: Metabarcoding of arthropods provides a sensitive measure of biodiversity response to forest stand composition compared with surrogate measures of biodiversity. *Ecological Indicators*, *101*, 313-323. https://doi.org/10.1016/j.ecolind.2019.01.023

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, *19*(1), 134-143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Baselga, A., Fujisawa, T., Crampton-Platt, A., Bergsten, J., Foster, P. G., Monaghan, M. T., & Vogler, A. P. (2013). Whole-community DNA barcoding reveals a spatio-temporal continuum of biodiversity at species and genetic levels. *Nature Communications*, *4*(1), Article 1. https://doi.org/10.1038/ncomms2881

Baselga, A., Gómez-Rodríguez, C., & Vogler, A. P. (2015). Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes. *Global Ecology and Biogeography*, *24*(8), 873-882. https://doi.org/10.1111/geb.12322

Baselga, A., & Orme, C. D. L. (2012). betapart: An R package for the study of beta diversity. *Methods in Ecology and Evolution*, *3*(5), 808-812. https://doi.org/10.1111/j.2041-210X.2012.00224.x

Baselga, A., Orme, D., Villeger, S., Bortoli, J. D., Leprieur, F., Logez, M., & Henriques-Silva, R. (2018). *betapart: Partitioning Beta Diversity into Turnover and Nestedness Components* (1.5.1). https://CRAN.R-project.org/package=betapart

Basset, Y., Cizek, L., Cuénoud, P., Didham, R. K., Guilhaumon, F., Missa, O., Novotny, V., Ødegaard, F., Roslin, T., Schmidl, J., Tishechkin, A. K., Winchester, N. N., Roubik, D. W., Aberlenc, H. P., Bail, J., Barrios, H., Bridle, J. R., Castaño-Meneses, G., Corbara, B., … Leponce, M. (2012). Arthropod diversity in a tropical forest. *Science*, *338*(6113), 1481-1484. https://doi.org/10.1126/science.1226727

Beng, K. C., Tomlinson, K. W., Shen, X. H., Surget-Groba, Y., Hughes, A. C., Corlett, R. T., & Slik, J. W. F. (2016). The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Scientific reports*, *6*(October 2015), 24965. https://doi.org/10.1038/srep24965

Betts, M. G., Phalan, B. T., Wolf, C., Baker, S. C., Messier, C., Puettmann, K. J., Green, R., Harris, S. H., Edwards, D. P., Lindenmayer, D. B., & Balmford, A. (2021). Producing wood at least cost to biodiversity: Integrating Triad and sharing–sparing approaches to inform forest landscape management. *Biological Reviews*, *96*(4), 1301-1317. https://doi.org/10.1111/brv.12703

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A. J., & Creer, S. (2021). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, *22*(4), 1231-1246. https://doi.org/10.1111/1755-0998.13512

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Brandon-Mong, G. J., Littlefair, J. E., Sing, K. W., Lee, Y. P., Gan, H. M., Clare, E. L., & Wilson, J. J. (2018). Temporal changes in arthropod activity in tropical anthropogenic forests. *Bulletin of Entomological Research*, *108*(6), 792-799. https://doi.org/10.1017/s000748531800010x

Brunet, J., Fritz, Ö., & Richnau, G. (2010). Biodiversity in European beech forests—A review with recommendations for sustainable forest management. *Ecological Bulletins*, *53*, 77-94. https://www.jstor.org/stable/41442021

Buss, D. F., Carlisle, D. M., Chon, T., Culp, J., Harding, J. S., Keizer-vlek, H. E., Robinson, W. A., Strachan, S., Thirion, C., & Hughes, R. M. (2015). *Stream biomonitoring using macroinvertebrates around the globe: A comparison of large-scale programs*. https://doi.org/10.1007/s10661-014-4132-8

Cai, W., Yang, C., Wang, X., Wu, C., Larrieu, L., Lopez-Vaamonde, C., Wen, Q., & Yu, D. W. (2021). The ecological impact of pest-induced tree dieback on insect biodiversity in Yunnan pine plantations, China. *Forest Ecology and Management*, *491*, 119173. https://doi.org/10.1016/j.foreco.2021.119173

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, *11*(12), Article 12. https://doi.org/10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), Article 7. https://doi.org/10.1038/nmeth.3869

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, *18*(1), 117-143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

Colwell, R. K., Brehm, G., Cardelús, C. L., Gilman, A. C., & Longino, J. T. (2008). Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics. science, 322(5899), 258-261. https://doi.org/10.1126/science.1162547

Creedy, T. J., Ng, W. S., & Vogler, A. P. (2019). Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy. *Ecology and Evolution*, *9*(6), 3105-3116. https://doi.org/10.1002/ece3.4839

Depraz, S., Sanial, E., Catalán, A. K. R., & Rojas, A. S. (2017). Less protection for better conservation? A politicised relationship between a city and its protected area in the vicinity of Nevado de Toluca (Mexico). *Articulo - Journal of Urban Research*, *16*, Article 16. https://doi.org/10.4000/articulo.3261

Des Roches, S., Pendleton, L. H., Shapiro, B., & Palkovacs, E. P. (2021). Conserving intraspecific variation for nature's contributions to people. *Nature Ecology & Evolution*, 5(5), 574-582.

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*(10), Article 10. https://doi.org/10.1038/nmeth.2604

Edgar, R. C. (2016). UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv*, 81257. https://doi.org/10.1101/081257

Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, *31*(21), 3476-3482. https://doi.org/10.1093/bioinformatics/btv401

Edwards, D. P., Magrach, A., Woodcock, P., Ji, Y., Lim, N. T.-L., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Khen, C. V., Chung, A. Y. C., Reynolds, G., Fisher, B., Laurance, W. F., Wilcove, D. S., Hamer, K. C., & Yu, D. W. (2014). Selective-logging and oil palm: Multitaxon impacts, biodiversity indicators, and trade-offs for conservation planning. *Ecological Applications*, *24*(8), 2029-2049. https://doi.org/10.1890/14-0010.1

Elbrecht, V., Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Hajibabaei, M., Wright, M., Zakharov, E. V., Hebert, P. D. N., & Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, *7*, e7745. https://doi.org/10.7717/peerj.7745

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—Sequence relationships with an innovative metabarcoding protocol. *Peer J*.

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, *7*(17), 6918-6926. https://doi.org/10.1002/ece3.3192

Fernandes, K., van der Heyde, M., Coghlan, M., Wardell-Johnson, G., Bunce, M., Harris, R., & Nevill, P. (2019). Invertebrate DNA metabarcoding reveals changes in communities across mine site restoration chronosequences. *Restoration Ecology*, *27*(5), 1177-1186. https://doi.org/10.1111/rec.12976

Ferreira-Neto, C. A., dos Santos Cruz, G. A., de Amorim, I. C., Balbino, V. Q., & de Cássia de Moura, R. (2017). Effects of fragmentation and anthropic pressure on the genetic structure of Canthon (Peltecanthon) staigi (Coleoptera: Scarabaeidae) populations in the Atlantic Forest domain. *Journal of insect conservation*, *21*, 267-276.

Fjeldså, J., Bowie, R. C. K., & Rahbek, C. (2012). The Role of Mountain Ranges in the Diversification of Birds. *Annual Review of Ecology, Evolution, and Systematics*, *43*(1), 249-265. https://doi.org/10.1146/annurev-ecolsys-102710-145113

Fuller, R. J., Oliver, T. H., & Leather, S. R. (2008). Forest management effects on carabid beetle communities in coniferous and broadleaved forests: Implications for conservation. *Insect Conservation and Diversity*, *1*, 242-252. https://doi.org/10.1111/j.1752-4598.2008.00032.x

Gálvez-Reyes, N., Arribas, P., Andújar, C., Emerson, B. C., Piñero, D., & Mastretta-Yanes, A. (2021). Dispersal limitations and long-term persistence drive differentiation from haplotypes to communities within a tropical sky-island: Evidence from community metabarcoding. *Molecular Ecology*, *30*(24), 6611-6626. https://doi.org/10.1111/mec.16195

Garner, A., Rachlow, J. L., & Hicks, J. F. (2005). Patterns of genetic diversity and its loss in mammalian populations. *Conservation Biology*, *19*(4), 1215-1221.

Gaublomme, E., Maebe, K., Van Doninck, K., Dhuyvetter, H., Li, X., Desender, K., & Hendrickx, F. (2013). Loss of genetic diversity and increased genetic structuring in response to forest area reduction in a ground dwelling insect: a case study of the flightless carabid beetle C arabus problematicus (C oleoptera, C arabidae). *Insect Conservation and Diversity*, *6*(4), 473-482.

Gómez-Rodríguez, C., & Baselga, A. (2018). Variation among European beetle taxa in patterns of distance decay of similarity suggests a major role of dispersal processes. *Ecography*, *41*(11), 1825-1834. https://doi.org/10.1111/ecog.03693

Goslee, S. C., & Urban, D. L. (2007). The ecodist Package for Dissimilarity-based Analysis of Ecological Data. *Journal of Statistical Software*, *22*, 1-19. https://doi.org/10.18637/jss.v022.i07

Graham, C. H., & Fine, P. V. A. (2008). Phylogenetic beta diversity: Linking ecological and evolutionary processes across space in time. *Ecology letters*, *11*(12), 1265-1277. https://doi.org/10.1111/j.1461-0248.2008.01256.x

Hajibabaei, M., Baird, D. J., Fahner, N. A., Beiko, R., & Golding, G. B. (2016). A new way to contemplate Darwin's tangled bank: How DNA barcodes are reconnecting biodiversity science and biomonitoring. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *371*(1702). https://doi.org/10.1098/rstb.2015.0330

Hannon, G. J. (2010). *FASTX-Toolkit.* http://hannonlab.cshl.edu/fastx_toolkit/

Hubbell, S. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press; JSTOR. https://doi.org/10.2307/j.ctt7rj8w

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN Community Edition—Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLOS Computational Biology*, *12*(6), e1004957. https://doi.org/10.1371/journal.pcbi.1004957

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P. M., Woodcock, P., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Hamer, K. C., Wilcove, D. S., Bruce, C., Wang, X., Levi, T., Lott, M., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245-1257. https://doi.org/10.1111/ele.12162

Kozak, K. H., & Wiens, J. J. (2006). Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution*, *60*(12), 2604. https://doi.org/10.1554/06-334.1

Krehenwinkel, H., Fong, M., Kennedy, S., Huang, E. G., Noriyuki, S., Cayetano, L., & Gillespie, R. (2018). The effect of DNA degradation bias in passive sampling devices on metabarcoding studies of arthropod communities and their associated microbiota. *PLoS ONE*, *13*(1), 1-14. https://doi.org/10.1371/journal.pone.0189188

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, *7*(1), 1-12. https://doi.org/10.1038/s41598-017-17333-x

Lange, M., Türke, M., Pašalić, E., Boch, S., Hessenmöller, D., Müller, J., Prati, D., Socher, S. A., Fischer, M., Weisser, W. W., & Gossner, M. M. (2014). Effects of forest management on ground-dwelling beetles (Coleoptera; Carabidae, Staphylinidae) in Central Europe are mainly mediated by changes in forest structure. *Forest Ecology and Management*, *329*, 166-176. https://doi.org/10.1016/j.foreco.2014.06.012

Langmaack, M., Schrader, S., & Helming, K. (2001). Effect of mesofaunal activity on the rehabilitation of sealed soil surfaces. *Applied Soil Ecology*, *16*(2), 121-130. https://doi.org/10.1016/S0929-1393(00)00108-6

Li, L., Zheng, B., & Liu, L. (2010). Biomonitoring and bioindicators used for river ecosystems: Definitions, approaches and trends. *Procedia Environmental Sciences*, *2*, 1510-1524. https://doi.org/10.1016/j.proenv.2010.10.164

Liu, M., Baker, S. C., Burridge, C. P., Jordan, G. J., & Clarke, L. J. (2020). DNA metabarcoding captures subtle differences in forest beetle communities following disturbance. *Restoration Ecology*, *28*(6), 1475-1484. https://doi.org/10.1111/rec.13236

Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., & Burridge, C. P. (2020). A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology*, *45*(3), 373-385. https://doi.org/10.1111/een.12831

Liu, M., Jordan, G. J., Burridge, C. P., Clarke, L. J., & Baker, S. C. (2021). Metabarcoding reveals landscape drivers of beetle community composition approximately 50 years after timber harvesting. *Forest Ecology and Management*, *488*, 119020. https://doi.org/10.1016/j.foreco.2021.119020

Lynggaard, C., Yu, D. W., Oliveira, G., Caldeira, C. F., Ramos, S. J., Ellegaard, M. R., Gilbert, M. T. P., Gastauer, M., & Bohmann, K. (2020). DNA-Based Arthropod Diversity Assessment in Amazonian Iron Mine Lands Show Ecological Succession Towards Undisturbed Reference Sites. *Frontiers in Ecology and Evolution*, *8*. https://www.frontiersin.org/articles/10.3389/fevo.2020.590976

Maechler, M., original), P. R. (Fortran, original), A. S. (S, original), M. H. (S, Hornik [trl, K., maintenance(1999-2000)), ctb] (port to R., Studer, M., Roudier, P., Gonzalez, J., Kozlowski, K., pam()), E. S. (fastpam options for, & Murphy (volume.ellipsoid({d >= 3})), K. (2022). *cluster: Finding Groups in Data: Cluster Analysis Extended Rousseeuw et al.* (2.1.3). https://CRAN.R-project.org/package=cluster

Magurran, A. E. (2016). How ecosystems change. *Science*, *351*(6272), 448-449. https://doi.org/10.1126/science.aad6758

Mastretta-Yanes, A., Cao, R., Arzata, S., Quadri, P., Espinosa, T., Arredondo, L., & Piñero, D. (2014). *¿Será exitosa la estrategia de cambio de categoría para mantener la biodiversidad del Nevado de Toluca? 12*, 7-17.

Mastretta-Yanes, A., Moreno-Letelier, A., Piñero, D., Jorgensen, T. H., & Emerson, B. C. (2015). Biodiversity in the Mexican highlands and the interaction of geology,

geography and climate within the Trans-Mexican Volcanic Belt. *Journal of Biogeography*, *42*(9), 1586-1600. https://doi.org/10.1111/jbi.12546

McGill, B. J. (2010). Towards a unification of unified theories of biodiversity. *Ecology Letters*, *13*(5), 627-642. https://doi.org/10.1111/j.1461-0248.2010.01449.x

Morrone, J. J., & Márquez, J. (2001). Halffter's Mexican Transition Zone, beetle generalized tracks, and geographical homology: Halffter's Mexican Transition Zone. *Journal of Biogeography*, *28*(5), 635-650. https://doi.org/10.1046/j.1365-2699.2001.00571.x

Nally, R. M., & Walsh, C. J. (2004). Hierarchical Partitioning Public-domain Software. *Biodiversity & Conservation*, *13*(3), 659-660. https://doi.org/10.1023/B:BIOC.0000009515.11717.0b

Nghiem, N., & Tran, H. (2016). The Biodiversity Benefits and Opportunity Costs of Plantation Forest Management: A Modelling Case Study of Pinus radiata in New Zealand. *Forests*, *7*(12), Article 12. https://doi.org/10.3390/f7120297

Noguerales, V., Meramveliotakis, E., Castro-Insua, A., Andújar, C., Arribas, P., Creedy, T. J., Overcast, I., Morlon, H., Emerson, B. C., Vogler, A. P., & Papadopoulou, A. (2021). Community metabarcoding reveals the relative role of environmental filtering and spatial processes in metacommunity dynamics of soil microarthropods across a mosaic of montane forests. *Molecular Ecology*, *n/a*(n/a), 1-19. https://doi.org/10.1111/mec.16275

Obrist, M. K., & Duelli, P. (2010). Rapid biodiversity assessment of arthropods for monitoring average local species richness and related ecosystem services. *Biodiversity and Conservation*, *19*(8), 2201-2220. https://doi.org/10.1007/s10531-010-9832-y

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package* (2.5-6). https://CRAN.R-project.org/package=vegan

Ortego, J., Aguirre, M. P., Noguerales, V., & Cordero, P. J. (2015). Consequences of extensive habitat fragmentation in landscape-level patterns of genetic diversity and structure in the M editerranean esparto grasshopper. *Evolutionary Applications*, *8*(6), 621-632.

Paillet, Y., Bergès, L., Hjältén, J., Odor, P., Avon, C., Bernhardt-Römermann, M., Bijlsma, R.-J., De Bruyn, L., Fuhr, M., Grandin, U., Kanka, R., Lundin, L., Luque, S., Magura, T., Matesanz, S., Mészáros, I., Sebastià, M.-T., Schmidt, W., Standovár, T., … Virtanen, R. (2010). Biodiversity differences between managed and unmanaged forests: Meta-analysis of species richness in Europe. *Conservation biology : the journal of the Society for Conservation Biology*, *24*(1), 101-112. https://doi.org/10.1111/j.1523-1739.2009.01399.x

Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, *15*(4), 819-830. https://doi.org/10.1111/1755-0998.12355

Pohlert, T. (2021). *PMCMR: Calculate Pairwise Multiple Comparisons of Mean Rank Sums* (4.4). https://CRAN.R-project.org/package=PMCMR

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-based species

delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595-609. https://doi.org/10.1080/10635150600852011
Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-Bravo, D., Rasmussen, C. M. Ø., Richardson, K., Rosing, M. T., Whittaker, R. J., & Fjeldså, J. (2019). Building mountain biodiversity: Geological and evolutionary processes. *Science*, *365*(6458), 1114-1119. https://doi.org/10.1126/science.aax0151
Rambaut, A. (2023). *Rambaut/figtree* [Java]. https://github.com/rambaut/figtree (Original work published 2015)
Rambo, T., Schowalter, T., & North, M. (2014). Canopy arthropod responses to thinning and burning treatments in old-growth mixed-conifer forest in the Sierra Nevada, California. *Forest Ecology and Management*, *326*, 91-100. https://doi.org/10.1016/j.foreco.2014.04.014
Rosindell, J., Hubbell, S. P., & Etienne, R. S. (2011). The Unified Neutral Theory of Biodiversity and Biogeography at Age Ten. *Trends in Ecology and Evolution*, *26*(7), 340-348. https://doi.org/10.1016/j.tree.2011.03.024
Rzendowski, J. (2006). Bosque de coníferas. *Vegetación de México*, 295-327.
Salces-Castellano, A., Patiño, J., Alvarez, N., Andújar, C., Arribas, P., Braojos-Ruiz, J. J., del Arco-Aguilar, M., García-Olivares, V., Karger, D. N., López, H., Manolopoulou, I., Oromí, P., Pérez-Delgado, A. J., Peterman, W. E., Rijsdijk, K. F., & Emerson, B. C. (2020). Climate drives community-wide divergence within species over a limited spatial scale: Evidence from an oceanic island. *Ecology Letters*, *23*(2), 305-315. https://doi.org/10.1111/ele.13433
Schnell, I. B., Bohmann, K., & Gilbert, M. T. P. (2015). Tag jumps illuminated—Reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular ecology resources*, *15*(6), 1289-1303. https://doi.org/10.1111/1755-0998.12402
Shokralla, S., Porter, T. M., Gibson, J. F., Dobosz, R., Janzen, D. H., Hallwachs, W., Golding, G. B., & Hajibabaei, M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Scientific Reports*, *5*(1), Article 1. https://doi.org/10.1038/srep09687
Staton, E. (2019). *Sestaton/Pairfq* [Perl]. https://github.com/sestaton/Pairfq (Original work published 2013)
Torres-Rojo, J. M., Moreno-Sánchez, R., & Mendoza-Briseño, M. A. (2016). Sustainable Forest Management in Mexico. *Current Forestry Reports*, *2*(2), 93-105. https://doi.org/10.1007/s40725-016-0033-0
Uscanga, A., López, H., Piñero, D., Emerson, B. C., & Mastretta-Yanes, A. (2021). Evaluating species origins within tropical sky-islands arthropod communities. *Journal of Biogeography*, *48*(9), 2199-2210. https://doi.org/10.1111/jbi.14144
van der Loos, L. M., & Nijland, R. (2021). Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. *Molecular Ecology*, *30*(13), 3270-3288. https://doi.org/10.1111/mec.15592
Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*. https://doi.org/10.1086/652373
Wang, X., Hua, F., Wang, L., Wilcove, D. S., & Yu, D. W. (2019). The biodiversity benefit of native forests and mixed-species plantations over monoculture plantations. *Diversity and Distributions*, *25*(11), 1721-1735. https://doi.org/10.1111/ddi.12972
Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & RStudio. (2020). *ggplot2: Create Elegant Data*

Visualisations Using the Grammar of Graphics (3.3.2). https://CRAN.R-project.org/package=ggplot2

Wiens, J. J. (2004). Speciation and ecology revisited: Phylogenetic niche conservatism and the origin of species. *Evolution*, *58*(1), 193-197. https://doi.org/10.1111/j.0014-3820.2004.tb01586.x

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, *3*(4), 613-623. https://doi.org/10.1111/j.2041-210X.2012.00198.x

Zhang, K., Lin, S., Ji, Y., Yang, C., Wang, X., CY, Y., H, W., Jiang, H., Harrison, R., & Yu, D. (2016). Plant diversity accurately predicts insect diversity in two tropical landscapes. *Molecular ecology*, *68125438*. https://doi.org/10.1111/MEC.13770

Zou, Y., Sang, W., Wang, S., Warren-Thomas, E., Liu, Y., Yu, Z., Wang, C., & Axmacher, J. C. (2015). Diversity patterns of ground beetles and understory vegetation in mature, secondary, and plantation forest regions of temperate northern China. *Ecology and evolution*, *5*(3), 531-542. https://doi.org/10.1002/ece3.1367

DISCUSIÓN GENERAL

Las comunidades de artrópodos recuperadas por ADN wcMBC, bloques de muestreo y trampas pitfall demostraron ser exitosas, ya que logramos obtener datos a nivel de haplotipos con enfoques basados en ASV. Esto nos permite evaluar las comunidades de artrópodos muestreadas en un área montañosa bajo conservación y bajo la influencia de actividades antropogénicas.

En los bosques bajo conservación, se encontraron más de 1,200 especies de artrópodos distintas en una sola montaña, y la composición de estas comunidades varió significativamente en 42 puntos de muestreo diferentes. El estudio reveló que la distancia entre los sitios de muestreo fue un factor clave en la estructura de estas comunidades, lo que indica que las limitaciones de dispersión desempeñan un papel importante en su configuración. Este patrón se observó tanto en distancias geográficas grandes como pequeñas, pero solo en especies de artrópodos con capacidades de dispersión limitadas.

En los bosques sujetos a actividades antropogénicas y bajo conservación, la evaluación de la biodiversidad en 115 sitios permitió recuperar 2,315 especies de artrópodos diferentes y analizar tanto la riqueza de especies como la composición de las comunidades de artrópodos. Se evidenciaron diferencias significativas en la composición de las comunidades entre los diferentes tratamientos y en comparación con el bosque de conservación. Los resultados señalan que la distancia entre los sitios de muestreo ejerce un efecto notable, incluso en distancias cortas de hasta 8 km entre tratamientos, y que los tratamientos de manejo forestal influyen en la composición de la comunidad. Este patrón

se mantuvo incluso en distancias geográficas más reducidas (<1,5 km), pero solo fue evidente para grupos de artrópodos con capacidad de dispersión limitada, como Collembola y Arachnida sin alas, en el tratamiento de chaponeo.

Avanzando hacia enfoques de haplotipos basados en ADN wcMBC para el monitoreo eficiente de comunidades de artrópodos terrestres

La integración de conceptos holísticos y la implementación de recomendaciones prácticas son fundamentales para mejorar la precisión de los datos de biodiversidad y los estudios de monitoreo que emplean enfoques basados en ADN wcMBC y ASV. En este contexto, el éxito del enfoque wcMBC depende de varios factores, tales como la limpieza de las muestras a granel y la homogeneización del tamaño de los artrópodos, la eficiencia de la extracción de ADN, la cuantificación en la construcción de bibliotecas, el diseño de las etiquetas, la afinidad de los primer y la amplificación para la identificación taxonómica, la etiquetación de las bibliotecas mediante PCR, la prevención de la contaminación cruzada, la cobertura de secuenciación, los métodos bioinformáticos, el filtrado de secuencias, los datos de referencia y el periodo de muestreo (Barsoum et al., 2019; Creedy et al., 2019; Elbrecht et al., 2017, 2019; Elbrecht & Leese, 2015, 2017; Gálvez-Reyes et al., 2021; Krehenwinkel et al., 2017, 2018; Liu et al., 2019; Liu et al., 2020; Piñol et al., 2015; Schnell et al., 2015; van der Loos y Nijland, 2021).

Aunque existen desafíos y sesgos técnicos, los datos obtenidos mediante wcMBC proporcionan aproximaciones de índices de riqueza y composición que permite evaluar el estado de los ecosistemas. Los primeros estudios basados en haplotipos (Elbrecht et al.,

2018; Arribas et al., 2020), que utilizan variantes de secuencia de amplicones (ASV; sensu Callahan et al., 2017), han abordado la biodiversidad de las comunidades. En este sentido, la incorporación de datos a nivel de haplotipos con enfoques basados en ASV puede contribuir a responder preguntas ecológicas y evolutivas, así como evaluar el impacto de los cambios ambientales a nivel de población, mediante el estudio de la diversidad intraespecífica de especies, brindando nuevas oportunidades para el monitoreo de la biodiversidad a través de la ampliación de ADN wcMBC (Elbrecht et al., 2018; Tsuji et al., 2020; Turon et al., 2020; Zizka et al., 2020; Arribas et al., 2020; Gálvez-Reyes et al., 2021; Noguerales et al., 2021).

Riqueza, estructura de comunidades de artrópodos: influencia de limitaciones de dispersión y su implicación en la evolución de sistemas de montañas tropicales

Las comunidades de artrópodos recuperadas mediante ADN wcMBC permitieron conocer la riqueza y la estructura de artrópodos en regiones de alta diversidad biológica y complejidad ecológica. Estudios previos también emplearon métodos similares para evaluar la comunidad de especies en bosques y determinar el esfuerzo de campo necesario para obtener muestras representativas (Andújar et al., 2017; Arribas et al., 2016; Creedy et al., 2019). Además, nuestro muestreo y clasificación por tamaño nos permitieron identificar ocho órdenes principales de artrópodos, lo cual concuerda con otros análisis de ADN wcMBC (Elbrecht, Peinert y Leese, 2017; Elbrecht et al., 2018; Creedy et al., 2019). El flujo deanálisis de datos de ADN wcMBC nos facilitó analizar patrones de diversidad y composición a nivel de cada orden de artrópodos en un marco multijerárquico, evitando el

sesgo introducido por la taxonomía tradicional (Creedy et al., 2019). Sin embargo, es importante tener en cuenta las posibles limitaciones de nuestro enfoque, como el uso de cebadores universales para PCR, diferencias en biomasa y salto de etiquetas (Piñol et al., 2015; Elbrecht, Peinert y Leese, 2017; Schnell, Bohmann y Gilbert, 2015), que mitigamos mediante la realización de PCR por triplicado, la clasificación de muestras por tamaño y el uso de primers desalinizados con precipitación de cromatografía de líquidos de alta resolución HPLC por sus siglas en inglés (High Performance Liquid Chromatography).

El recambio de comunidades de artrópodos a niveles multijerárquicos fue notable mediante la identificación de 1,277 especies de artrópodos pertenecientes a ocho órdenes. Este recambio en las comunidades de artrópodos fue significativo entre diferentes áreas de muestreo y tipos de vegetación. El recambio de especies demostró ser más dominante que el anidamiento en la estructura de las comunidades, y esta diferenciación persistió incluso al considerar especies de manera conservadora a nivel molecular. La distancia se reveló como un factor clave que explicó la estructura de la comunidad, desde haplotipos hasta linajes, sugiriendo que las limitaciones de dispersión impulsan la estructura de la comunidad en distintos niveles multijerárquicos dentro de una sola isla del cielo tropical. Además, se identificó una división Este-Oeste en la estructura de la comunidad dentro del área de estudio, correspondiente a laderas opuestas de la montaña, lo que podría influir en el ensamblaje de la comunidad. Esto se debe a que las laderas orientadas al Este con exposición matutina al sol, pueden presentar condiciones ambientales diferentes a las laderas orientadas al Oeste, que son más frías y con mayor presencia de niebla (Rahbek et al., 2019). De hecho, se espera que los bosques de *A. religiosa*, crezcan en condiciones

similares dentro de una misma montaña, con un nicho ambiental restringido asociado a sitios húmedos y fríos (Rzendowski, 2006).

Las limitaciones de dispersión impulsan la estructura de la comunidad en niveles multijerárquicos en escalas geográficas finas dentro de una montaña, a nivel de haplotipos, utilizando análisis de IBD e IBR. Se observó que la disminución de la similitud entre las comunidades disminuye con la distancia espacial, especialmente en grupos con capacidades de dispersión más limitadas. Las comunidades con buenos dispersores no solo son homogéneas debido a su capacidad para dispersarse a distancias mayores, sino también porque pueden superar barreras geográficas que separan hábitats adecuados. Este patrón fue particularmente evidente para Collembola a distancias geográficas más finas (<2 km), consistente con estudios genéticos dentro de Collembola que muestran diferenciación genética en distancias geográficas muy cortas (Cicconardi et al., 2013; Arribas et al., 2020). Este fenómeno también se ha observado en otros escarabajos de hojas y acuáticos (Baselga et al., 2013; 2015; Gómez-Rodríguez y Baselga, 2018) a escalas geográficas mucho más amplias (cientos de km), y sólo recientemente a escala de pocos kilómetros, en Collembola, Acari y Coleoptera (Arribas et al., 2020). Además, las diferencias altitudinales tienen un mayor impacto en la similitud que la distancia, la pendiente o el tipo de vegetación por sí solos. Las características del paisaje también influyen en la estructura de la comunidad de Diptera, aunque no para Collembola debido su capacidad limitada de dispersión. La distancia efectiva entre los sitios de muestreo depende del modelo de elevación utilizado para establecer los valores de conductancia. Esta observación es congruente con la predicción de Janzen de que las barreras de montañas son más pronunciadas en regiones

tropicales (Janzen, 1967), respaldadas por estudios empíricos recientes (Polato et al., 2018). Aunque la conectividad del paisaje contribuye a la limitación de la dispersión, la distancia geográfica parece desempeñar un papel predominante en ambos órdenes. Se sugiere que la limitación de la dispersión actúa a lo largo del tiempo evolutivo y puede explicar la diversificación a pequeña escala espacial de ciertos escarabajos *Scarelus* en las montañas tropicales (Bray y Bocak, 2016). Por último, se destaca la utilidad de los enfoques multijerárquicos para evaluar la variación en los ensamblajes biológicos impulsados por la dispersión que sigue una geometría fractal en la que la distribución de haplotipos y linajes superiores es similar consigo misma (Baselga et al., 2013, 2015; Gómez-Rodríguez et al., 2019).

Las implicaciones para la evolución en sistemas insulares y montañas tropicales revelaron que las limitaciones de dispersión, derivadas de la geografía y topografía de estas áreas, afectan la estructura de la comunidad de manera similar a lo observado en las islas oceánicas (Salces-Castellano et al., 2019). Es probable que las restricciones de dispersión sean más significativas en los trópicos debido a las características propias de las especies tropicales, como tolerancias térmicas estrechas y menor capacidad de dispersión, lo que conduce a un mayor aislamiento por distancia y aislamiento por elevación (Polato et al., 2018). En consecuencia, la combinación de factores geográficos, topográficos y las características específicas de dispersión, adaptaciones particulares o estrategias de colonización de los linajes taxonómicos puede influir en la dinámica de dispersión y, por ende, en la estructura de la comunidad en estos entornos (Salces-Castellano et al., 2019; Polato et al., 2018). El archipiélago de islas del cielo, que incluye el Nevado de Toluca, podría

explicar por qué las montañas tropicales son puntos críticos de biodiversidad, ya que se ha propuesto que la diferenciación de poblaciones pequeñas y aisladas espacialmente, combinada con su mantenimiento a largo plazo, puede conducir a la especiación (Rahbek et al., 2019). Este estudio evidencia que una sola isla del cielo puede actuar como cuna de diferenciación poblacional, y que este proceso puede persistir a lo largo de escalas de tiempo evolutivas. A diferencia de estudios anteriores centrados en taxones específicos de baja dispersión que llegaron a conclusiones similares (Bray y Bocak, 2016), nuestros datos a nivel de comunidad sugieren que este fenómeno podría ser generalizado entre los artrópodos tropicales de montaña. Este descubrimiento se ve respaldado por diversos estudios que exploran las limitaciones de la dispersión espacial en comunidades de artrópodos del suelo, moluscos terrestres, escarabajos de las hojas y microartrópodos del suelo mediante el *metabarcoding* a nivel de haplotipo (Arribas et al., 2020; Gómez-Rodríguez et al., 2019; Andújar et al., 2022; Noguerales et al., 2021). A medida que se profundiza en el análisis de la biodiversidad del suelo en islas oceánicas, "Islas en el Cielo", regiones montañosas y bosques nubosos, se destaca la relevancia del filtrado ambiental, la topografía y los procesos espaciales en el ensamblaje comunitario (Andújar et al., 2022; Gálvez-Reyes et al., 2021; Noguerales et al., 2021). Las comunidades completas de artrópodos en gradientes de elevación dentro de una isla (Lim et al., 2022) revelaron que el conservadurismo del nicho climático es un factor importante que da forma al ensamblaje ecológico a lo largo de la elevación, sugiriendo que la complejidad topográfica es un impulsor significativo de la diversificación. De manera similar, los enfoques de *metabarcoding* de toda la comunidad de artrópodos del suelo revelaron fuertes

limitaciones en el filtrado y la dispersión del hábitat como impulsores del ensamblaje de comunidades y la metafilogeografía de la biodiversidad dentro de una isla oceánica (Andújar et al., 2022; Noguerales et al., 2021). La exploración de gradientes altitudinales en bosques subtropicales y montañas tropicales revela patrones consistentes en la diversidad microbiana del suelo, resaltando la influencia de los microclimas, la variabilidad climática (Ma et al., 2022; Zhang et al., 2023) y las diferentes elevaciones, respaldando la hipótesis de Janzen relacionada con la variabilidad climática (Feng et al., 2023). Aunque los organismos difieren, la sensibilidad de las comunidades microbianas sugiere aplicaciones extrapolativas a otros organismos y ambientes. En las montañas neotropicales, se observa un interés en el recambio comunitario a lo largo de gradientes altitudinales, poniendo a prueba la hipótesis de que 'los pasos de montaña son más altos en el trópico' para las comunidades de insectos (Shimabukuro et al., 2023). A pesar de las diferencias en organismos y ecosistemas, se destaca que las metodologías y enfoques utilizados ofrecen una perspectiva valiosa para la investigación de la biodiversidad y el ensamblaje comunitario en diversos entornos. Los procesos generales que se pueden caracterizar del ensamblaje comunitario incluyen la importancia relativa de la estocasticidad, el aislamiento por distancia, la diferenciación asociada al hábitat o al huésped, filtrado ambiental y limitación de dispersión, que sugieren un potencial de extrapolación, aunque se destaca la necesidad de replicar estudios en diferentes regiones y ecosistemas para validar la generalidad de los patrones observados.

Impacto del manejo forestal y las limitaciones de dispersión en la estructura y riqueza de comunidades: Implicaciones para el monitoreo de biodiversidad

La conservación de los bosques se enfrenta desafíos derivados de cambios en el uso del suelo impulsados por políticas públicas (Wang et al., 2019). Las prácticas de manejo forestal pueden tener impactos significativos en la biodiversidad, como la sustitución de especies de árboles, efectos de la heterogeneidad del hábitat y la disponibilidad de presas para artrópodos (Brunet et al., 2010; Lange et al., 2014). El estudio reveló diferencias significativas en la riqueza de artrópodos entre distintos tratamientos de manejo forestal, observándose un ligero aumento en la diversidad genética en áreas reforestadas en comparación con la conservación, aunque menor que otros tratamientos. Estos resultados concuerdan con estudios previos que han identificado ensamblajes únicos de artrópodos en diferentes tipos de bosques y hábitats, y destacando la importancia de la conservación y manejo adecuado de los bosques para preservar la biodiversidad (Beng et al., 2016; Arribas et al., 2020; Gálvez-Reyes et al., 2021; Noguerales et al., 2021; Wang et al., 2019; Baker et al., 2014). Curiosamente, la riqueza de algunos grupos de artrópodos y tratamientos no presentó diferencias significativas. Aunque algunos estudios han sugerido que la riqueza de artrópodos puede ser un indicador confiable de la riqueza de plantas en bosques tropicales (Basset et al., 2012; Lynggaard et al., 2020; Zhang et al., 2016), y entre diferentes tipos de rodales en intervalos de tiempo (Barsoum et al., 2019), los resultados de este estudio indican que la estrategia bioinformática utilizada puede influir en los resultados de la riqueza de especies (Liu et al., 2020). Similarmente, en nuestros propios resultados, algunos grupos de artrópodos tampoco mostraron diferencias significativas en la riqueza de

especies. En este contexto, se destaca que la composición de la comunidad se presenta como un enfoque más relevante en los estudios de biodiversidad (Lynggaard et al., 2020; Magurran, 2016; Wang et al., 2019), ya que la riqueza de especies puede recuperarse rápidamente después de la revegetación (Lynggaard et al., 2020). Además, los cambios en la diversidad alfa pueden no ser detectables debido a la sustitución de especies (Lynggaard et al., 2020). Para abordar este problema, se han empleado varios pasos de filtro para eliminar OTU falsos (Andújar et al., 2021), y se propone el uso de "iNextPD" para estimar la diversidad filogenética en lugar de la riqueza de especies (Wang et al., 2019).

Se observó un cambio significativo en la composición de la comunidad entre los tratamientos y la conservación del bosque. Nuestro estudio examinó el impacto de la gestión forestal, las perturbaciones y la reforestación en la composición y estructura de la comunidad de artrópodos, que pueden ser indicadores de la biodiversidad (Lynggaard et al., 2020; Wang et al., 2019). Se encontraron cambios notables en la composición de la comunidad entre los bosques regenerados después de la tala y los bosques de conservación, especialmente en ciertos grupos de artrópodos. Estos hallazgos son consistentes con estudios previos sobre la composición de la comunidad de escarabajos en bosques en regeneración (Liu et al., 2021) y la respuesta de la composición de artrópodos al cambio de uso de la tierra y la regeneración en los trópicos (Beng et al., 2016; Lynggaard et al., 2020; Magurran, 2016; Wang et al., 2019). Sin embargo, algunos grupos de artrópodos mostraron una separación incipiente pero no significativa entre los bosques explotados y los bosques de conservación. Se observó que las plantaciones mixtas aumentaban la diversidad y composición de artrópodos, pero no compensaban la pérdida de biodiversidad en

comparación con los bosques nativos (Wang et al., 2019). Además, factores como la distancia geográfica y otros factores ambientales interactúan de manera sinérgica en la configuración de la biodiversidad y la composición de la comunidad (Beng et al., 2016; Arribas et al., 2020; Noguerales et al., 2021; Gálvez-Reyes et al., 2021). Por lo tanto, es importante comprender las sutiles diferencias en la composición de la comunidad y los mecanismos subyacentes para la gestión adaptativa (Barsoum et al., 2019). Nuestro estudio destaca la complejidad y dinamismo de las respuestas de la comunidad de artrópodos a los esfuerzos de manejo y conservación forestal, y subraya la necesidad de más investigación en este campo.

Los tratamientos de manejo forestal y las limitaciones de dispersión están impulsando la estructura de la comunidad. El papel de las limitaciones de dispersión en la configuración de la estructura de la comunidad en los ecosistemas forestales fue investigado por Arribas et al. (2020). El presente estudio evaluó la influencia de las limitaciones de dispersión en la estructura de la comunidad en ecosistemas forestales, centrándose específicamente en el Nevado de Toluca, México. Se encontró que las limitaciones de dispersión desempeñan un papel importante en la configuración de la estructura de la comunidad, con una disminución de la similitud de las comunidades a medida que aumenta la distancia espacial. Este patrón es más pronunciado en grupos con capacidades de dispersión más limitadas, como Collembola y Arachnida, en comparación con grupos con capacidades de dispersión más amplias, como Díptera e Himenóptera. Asimismo, se observó que el manejo forestal, como el chaponeo, interactúa con la distancia a menos de 1,5 km, influenciando la configuración de la composición de la comunidad de

Collembola y Arachnida. Esto indica que las limitaciones de dispersión impulsan la estructura de la comunidad a escalas geográficas finas dentro de los bosques bajo conservación, como se informó en otros estudios (Arribas et al., 2020; Gálvez-Reyes et al., 2021). Además, se evidenció que el manejo forestal específico tiene un impacto en la estructura de la comunidad, respaldando hallazgos de estudios previos de Brandon-Mong et al. (2018) y Zhang et al. (2016), quienes también encontraron patrones similares en bosques afectados por actividades antropogénicas.

El wcMBC es una herramienta que evalúa la biodiversidad en muestras a granel y proporciona información sobre prácticas de gestión y políticas de desarrollo (Ji et al., 2013; Yu et al., 2012). Algunos estudios han demostrado que el wcMBC puede revelar grupos taxonómicos sensibles a diferentes actividades de manejo forestal y reforestación en niveles multijerárquicos (Arribas et al., 2020; Noguerales et al., 2021; Gálvez-Reyes et al., 2021). Aunque existen varios factores metodológicos que pueden afectar la precisión y confiabilidad de los resultados, como el diseño de muestreo, la extracción de ADN, la selección de primers, la cobertura de secuenciación, los datos de referencias, y los métodos bioinformáticos (Bohmann et al., 2021; Creedy et al., 2019; Liu et al., 2020; Elbrecht et al., 2017, 2019; Elbrecht & Leese, 2015; Krehenwinkel et al., 2017, 2018; Piñol et al., 2015; Schnell et al., 2015; van der Loos & Nijland, 2021); se ha explorado la riqueza y composición de artrópodos en diferentes tipos de bosques y uso de la tierra (Barsoum et al., 2019; Beng et al., 2016; Liu et al., 2020; Lynggaard et al., 2020; Wang et al., 2019). Por lo tanto, se sugiere revisar cuidadosamente las opciones y aplicaciones metodológicas para mejorar la precisión y confiabilidad de los resultados, así como controlar los efectos temporales en el

muestreo para capturar variaciones temporales a escala fina en la composición de las comunidades de artrópodos (Creedy et al., 2019; Elbrecht et al., 2017).

Las implicaciones para el manejo forestal en el Nevado de Toluca sugieren la implementación de un programa de monitoreo que integre el manejo forestal, la conservación forestal y la detección de amenazas potenciales para la diversidad de artrópodos. Nuestros hallazgos aportan evidencias sobre los efectos del manejo forestal y la reforestación sobre la riqueza y composición de artrópodos. Para lograr esto, se recomienda promover prácticas de manejo forestal que minimicen los impactos negativos y considerar la diversidad funcional de la comunidad de artrópodos (Lynggaard et al., 2020). Para detectar las diferencias en los niveles de biodiversidad, el monitoreo forestal podría incluir evaluaciones temporales de la diversidad de artrópodos (Barsoum et al., 2019; Creedy et al., 2019). Se deben recopilar datos de biodiversidad para identificar los bosques de interés para la conservación, detectar amenazas como el cambio climático y nuevas plagas y patógenos, y medir la eficacia de las medidas de política forestal destinadas a mejorar la biodiversidad forestal (Barsoum et al., 2019; Lynggaard et al., 2020; Wang et al., 2019). Se sugiere la necesidad de evaluar la diversidad de artrópodos en diferentes momentos y considerar las condiciones climáticas y los impactos a escala del paisaje en la implementación de prácticas de manejo forestal. Para mitigar los impactos negativos que han ocurrido en la historia del Nevado de Toluca (Mastretta-Yanes et al., 2014), se debe informar sobre los índices de biodiversidad y la composición de la comunidad de artrópodos y considerar los datos de haplotipos en los esfuerzos de biomonitoreo.

Perspectivas y estudios futuros en bosques de altas montañas

En los bosques de las montañas tropicales, es esencial contar con la capacidad técnica (laboratorio y bioinformática) y financiera para el monitoreo de artrópodos. Hasta ahora, se ha demostrado que el ADN wcMBC es una herramienta eficaz para el seguimiento de cambios en ecosistemas en proceso de restauración (Lynggaard et al., 2021; Edwards et al., 2014; Fernandes et al., 2019) y los bosques puede experimentar variaciones significativas en períodos cortos de tiempo (Barsoum et al., 2019). Además, se ha evidenciado que la riqueza de especies de plantas puede predecir con precisión la riqueza de especies de artrópodos (Zhang et al., 2016), y la biodiversidad es mayor en los bosques nativos y las plantaciones mixtas que en los monocultivos (Wang et al., 2019). Aunque, se está acumulando evidencia sobre la disminución de las poblaciones de insectos (Butchart et al., 2010; Del-Val et al., 2021), persisten incertidumbres con respecto a la magnitud y distribución de estas disminuciones (Van Klink et al., 2020). Si bien las poblaciones de insectos terrestres han disminuido, las poblaciones de agua dulce han mostrado un aumento general, posiblemente debido a iniciativas de agua limpia y cambios climáticos (Van Klink et al., 2020). Las variaciones en las tendencias sugieren factores localizados que impulsan los cambios poblacionales (Van Klink et al., 2020), destacando oportunidades para esfuerzos de conservación específicos.

Las islas en el cielo tropicales, como el Nevado de Toluca, se identifican como prioridades críticas para el monitoreo debido a que albergan una alta diversidad de flora y fauna (Mastretta-Yanes et al., 2018; Uscanga et al., 2019). Sin embargo, estudiar esta diversidad en áreas poco investigadas presenta un desafío, ya que puede resultar en la falta

de una base de datos de referencia completa, lo que dificulta la asignación taxonómica de las OTUs (Lynggaard et al., 2021; Liu et al., 2019; Cuff et al., 2021). La caracterización de las comunidades de artrópodos basada en el ADN es una herramienta crucial para el monitoreo a largo plazo (Gálvez-Reyes et al., 2021) y evaluación de impacto, pero depende de la confiabilidad taxonómica de las bases de datos genéticas para las asignaciones taxonómicas (Leray et al., 2019). Un monitoreo temporal y un inventario coordinado pueden proporcionar datos de referencia para la planificación de la conservación (Arribas et al., 2020; Emerson et al., 2023). Además, implementaciones más sofisticadas de muestreo y HTS (por sus siglas en inglés; High-Throughput Sequencing) con datos de abundancia tienen el potencial de contribuir a los esfuerzos de conservación (Emerson et al., 2023). Por lo tanto, es necesario crear una línea base de datos de referencia de artrópodos de las montañas de la Faja Volcánica Transmexicana y los bosques sujetos a manejo forestal, plantaciones y reforestaciones.

Las muestras a granel en estudios de ADN pueden presentar sesgos debido a la variabilidad en la biomasa de los especímenes, donde los especímenes grandes pueden dominar los resultados, mientras que los pequeños pueden no ser detectados (Elbrecht y Leese, 2015; Elbrecht, Peinert y Leese, 2017). La estandarización del tamaño del tejido es necesario para evitar sesgos en la amplificación del ADN, pero se sugiere aumentar la profundidad de secuenciación suficiente para evaluar la biodiversidad (Creedy et al., 2019).

Para el muestreo, se deben emplear varios tipos de recolección con el fin de abarcar la diversidad de artrópodos, o según los artrópodos de interés para los estudios, como, por ejemplo, polinizadores. Para futuros estudios, se sugiere utilizar la trampa Malaise, que

captura insectos voladores como Himenópteros, Dípteros, Lepidópteros, Neuroptera, Hemípteros y Coleópteros (Brandon-Mong et al., 2015; Gibson et al., 2014; Ji et al., 2013; Shokralla et al., 2015; Yang et al., 2014; Yu et al., 2012), ya que parece apropiada para el ADN wcMBC a partir de muestras almacenadas durante períodos no óptimos, como experimentos de degradación de dos y cuatro semanas (Krehenwinkel et al., 2018). Además, las trampas pitfall, que recolectan artrópodos como hormigas, coleópteros y micro-himenópteros (Ji et al., 2013; Gálvez-Reyes et al., 2021) que son una opción económica para la captura de artrópodos terrestres y pueden dejarse instaladas durante varias semanas, lo que prolonga el tiempo de muestreo y reduce la variación en las colecciones específicas. Ambas trampas son útiles en el ADN wcMB ya que captura variaciones temporales a escala fina en la composición de las comunidades de artrópodos (Barsoum et al., 2019; Creedy et al., 2019).

Basándonos en nuestros hallazgos, se han identificado puntos importantes para futuras investigaciones dentro de una sola montaña o diferentes montañas de la FVTM: (i) Se espera que los indicadores de biodiversidad (grupos de artrópodos) estén correlacionados con el manejo forestal y la estructura de la comunidad de plantas en el bosque dentro de la FVTM. (ii) Se espera que las condiciones climáticas influyan en la adaptación de los indicadores de biodiversidad (grupos de artrópodos). (iii) Se espera que la conectividad del hábitat a nivel de paisaje, considerando los tratamientos de manejo forestal, reforestación, plantación y disturbios, así como factores limitantes como la topografía y elevación de los sitios, moldeen la composición de las comunidades de artrópodos a nivel local (una sola montaña) y regional (entre montañas). A través de la

obtención de datos utilizando trampas pitfall y Malaise para recopilar información de toda la comunidad de artrópodos mediante ADN wcMBC, con cuatro réplicas por tratamiento y la plataforma NOVAseq. Los resultados mostrarán que los bosques nativos dentro de las montañas de la FVTM presentan los niveles más altos de riqueza de especies de artrópodos, diversidad de Shannon y Simpson, y diversidad filogenética y de Faith. Además, la mayoría de estas especies son exclusivas de los bosques nativos, lo cual es consistente con los patrones de diversidad observados en otras montañas, así como con la correlación positiva con la riqueza y composición de plantas. La similitud de las comunidades será influenciada por la distancia entre sitios y las actividades antropogénicas a lo largo del tiempo.

CONCLUSIONES

En conclusión, la realización de enfoques basados en ADN metabarcoding de toda la comunidad (wcMBC) y ASV para la evaluación de la biodiversidad requiere la integración de conceptos holísticos de monitoreo y la adhesión a recomendaciones prácticas para garantizar la relevancia de los estudios. El éxito en el ADN wcMBC depende de varios factores, que incluyen el marco conceptual, las hipótesis y las preguntas de investigación, la eficiencia de la técnica de extracción de ADN, los marcadores genéticos, los cebadores utilizados para la identificación taxonómica, los datos de referencia y los métodos bioinformáticos. A pesar de los desafíos y los sesgos técnicos, los datos moleculares son adecuados para aproximar índices bióticos basados en datos de secuencias y evaluar el estado de los ecosistemas.

Los resultados de este estudio proporcionan evidencia empírica de que las montañas tropicales hiperdiversas pueden exhibir patrones globales reforzados por la diferenciación de pequeñas poblaciones aisladas y su persistencia a largo plazo. Podría ser necesario reconsiderar el enfoque sobre el aislamiento de la población entre diferentes montañas, ya que las comunidades de artrópodos pueden mostrar una fuerte rotación a nivel intraespecífico dentro de una isla del cielo, incluso a escalas geográficas limitadas. La distancia y la elevación se identifican como impulsores significativos de la estructura de la biodiversidad, desde haplotipos individuales hasta especies y niveles superiores, incluso a escalas espaciales finas. Las limitaciones de dispersión pueden actuar como una fuente de diferenciación genética a escala local dentro de las montañas tropicales, lo que podría conducir a la diversificación a nivel de especies con el tiempo.

Las evaluaciones de la biodiversidad son fundamentales para una gestión efectiva de los ecosistemas, y la tecnología de alto rendimiento combinada con el marcador COI puede proporcionar un método eficaz para interpretar la biodiversidad de las comunidades de artrópodos en diferentes ecosistemas forestales. Sin embargo, la utilidad del ADN wcMBC para la gestión ecológica y el control de la biodiversidad podría mejorarse aún más con bases de datos de referencia más completas. Es crucial considerar el vínculo entre la riqueza y composición de la comunidad de artrópodos y la diversidad de plantas en diferentes ecosistemas forestales, así como el impacto del cambio climático en la dinámica forestal, para informar las discusiones de políticas y los esfuerzos de monitoreo. La investigación interdisciplinaria sobre indicadores de biodiversidad y gestión forestal puede conducir al desarrollo de directrices efectivas de gestión de la biodiversidad para

administradores de bosques y conservación basadas en hallazgos científicos recientes, resultando en una gestión de la biodiversidad más asertiva en los ecosistemas forestales.

REFERENCIAS

Andújar, C., Arribas, P., Gray, C., Bruce, C., Woodward, G., Yu, D. W., & Vogler, A. P. (2017). Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Molecular Ecology*, *27*(1), 146-166.https://doi.org/10.1111/mec.14410.

Andújar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology*, *27*(20), 3968-3975. https://doi.org/10.1111/mec.14844

Andújar, C., Creedy, T. J., Arribas, P., López, H., Salces-Castellano, A., Pérez-Delgado, A. J., Vogler, A. P., & Emerson, B. C. (2021). Validated removal of nuclear pseudogenes and sequencing artefacts from mitochondrial metabarcode data. *Molecular Ecology Resources*, *21*(6), 1772-1787. https://doi.org/10.1111/1755-0998.13337

Andujar, C., Arribas, P., López, H., Arjona, Y., Pérez-Delgado, A., Oromí, P., ... & Emerson, B. C. (2022). Community assembly and metaphylogeography of soil biodiversity: Insights from haplotype-level community DNA metabarcoding within an oceanic island. *Molecular Ecology*, *31*(15), 4078. https://doi.org/10.1111/mec.16560.

Arribas, P., Andújar, C., Hopkins, K., Shepherd, M., & Vogler, A. P. (2016). Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. *Methods in Ecology and Evolution*, *7*(9), 1071-1081. https://doi.org/10.1111/2041-210X.12557

Arribas, P., Andújar, C., Salces-Castellano, A., Emerson, B. C., & Vogler, A. P. (2020). The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding. *Molecular Ecology*, *30*(1), 48-61. https://doi.org/10.1111/mec.15591

Back, P., Suominen, A., Malo, P., Tahvonen, O., Blank, J., & Deb, K. (2020). *Towards sustainable forest management strategies with MOEAs*. 1046-1054. https://doi.org/10.1145/3377930.3389837

Baker, T. P., Jordan, G. J., Steel, E. A., Fountain-Jones, N. M., Wardlaw, T. J., & Baker, S. C. (2014). Microclimate through space and time: Microclimatic variation at the edge of regeneration forests over daily, yearly and decadal time scales. *Forest Ecology and Management*, *334*, 174-184. https://doi.org/10.1016/j.foreco.2014.09.008

Barsoum, N., Bruce, C., Forster, J., Ji, Y.-Q., & Yu, D. W. (2019). The devil is in the detail: Metabarcoding of arthropods provides a sensitive measure of biodiversity response to forest stand composition compared with surrogate measures of biodiversity. *Ecological Indicators*, *101*, 313-323. https://doi.org/10.1016/j.ecolind.2019.01.023

Basset, Y., Cizek, L., Cuénoud, P., Didham, R. K., Guilhaumon, F., Missa, O., Novotny, V., Ødegaard, F., Roslin, T., Schmidl, J., Tishechkin, A. K., Winchester, N. N., Roubik, D. W., Aberlenc, H. P., Bail, J., Barrios, H., Bridle, J. R., Castaño-Meneses, G., Corbara, B., … Leponce, M. (2012). Arthropod diversity in a tropical forest. *Science*, *338*(6113), 1481-1484. https://doi.org/10.1126/science.1226727

Baselga, A., Fujisawa, T., Crampton-Platt, A., Bergsten, J., Foster, P. G., Monaghan, M. T., & Vogler, A. P. (2013). Whole-community DNA barcoding reveals a spatio-temporal continuum of biodiversity at species and genetic levels. *Nature Communications*, *4*(1), Article 1. https://doi.org/10.1038/ncomms2881

Baselga, A., Gómez-Rodríguez, C., & Vogler, A. P. (2015). Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes. *Global Ecology and Biogeography*, *24*(8), 873-882. https://doi.org/10.1111/geb.12322

Beng, K. C., Tomlinson, K. W., Shen, X. H., Surget-Groba, Y., Hughes, A. C., Corlett, R. T., & Slik, J. W. F. (2016). The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Scientific reports*, *6*(October 2015), 24965. https://doi.org/10.1038/srep24965

Bell, G. (2001). Neutral macroecology. *Science*, *293*(5539), 2413-2418.

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A. J., & Creer, S. (2021). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, *22*(4), 1231-1246. https://doi.org/10.1111/1755-0998.13512

Brandon-Mong, G. J., Littlefair, J. E., Sing, K. W., Lee, Y. P., Gan, H. M., Clare, E. L., & Wilson, J. J. (2018). Temporal changes in arthropod activity in tropical anthropogenic forests. *Bulletin of Entomological Research*, *108*(6), 792-799. https://doi.org/10.1017/s000748531800010x

Bray, T. C., & Bocak, L. (2016). Slowly dispersing neotenic beetles can speciate on a penny coin and generate space-limited diversity in the tropical mountains. *Scientific Reports*, *6*(1), 33579. https://doi.org/10.1038/srep33579

Bruce, C. (2013). *From metacommunity dynamics to rapid biodiversity assessment: DNA-based approaches expand horizons in both fundamental and applied ecology Catharine Bruce Submitted for the qualification of PhD* (Número November).

Brunet, J., Fritz, Ö., & Richnau, G. (2010). Biodiversity in European beech forests—A review with recommendations for sustainable forest management. *Ecological Bulletins*, *53*, 77-94. https://www.jstor.org/stable/41442021

Buss, D. F., Carlisle, D. M., Chon, T., Culp, J., Harding, J. S., Keizer-vlek, H. E., Robinson, W. A., Strachan, S., Thirion, C., & Hughes, R. M. (2015). *Stream biomonitoring using macroinvertebrates around the globe: A comparison of large-scale programs*. https://doi.org/10.1007/s10661-014-4132-8

Butchart, S. H., Walpole, M., Collen, B., Van Strien, A., Scharlemann, J. P., Almond, R. E., ... & Watson, R. (2010). Global biodiversity: indicators of recent declines. *Science*, *328*(5982), 1164-1168. https://doi.org/10.1126/science.1187512.

Callahan, B.J., McMurdie, P.J. and Holmes, S.P. (2017) 'Exact sequence variants should replace operational taxonomic units in marker-gene data analysis', *The ISME Journal*, 11(12), pp. 2639–2643. https://doi.org/10.1038/ismej.2017.119.

Ceballos, G., Díaz-Pardo, E., Espinosa, H., Flores-Villela, O., García, A., & Martínez, L. (2009). *Zonas críticas y de alto riesgo para la conservación de la biodiversidad de México, en Capital natural de México, vol. II: Estado de conservación y tendencias de cambio. II*, 575-600.

Challenger, A., & Dirzo, R. (2009). *Factores de cambio y estado de la biodiversidad, en Capital natural de México, vol. II: Estado de conservación y tendencias de cambio. II*, 37-73.

Chapela, F. (2012). Estado de los bosques de México. *México, DF: Consejo civil mexicano para la silvicultura sostenible AC*, *52*, 217.

Cicconardi, Francesco, Pietro P. Fanciulli, and Brent C. Emerson. "Collembola, the biological species concept and the underestimation of global species richness." *Molecular ecology* 22.21 (2013): 5382-5396. https://doi.org/10.1111/mec.12472

Coote, L., Dietzsch, A. C., Wilson, M. W., Graham, C. T., Fuller, L., Walsh, A. T., Irwin, S., Kelly, D. L., Mitchell, F. J. G., Kelly, T. C., & O'Halloran, J. (2013). Testing indicators of biodiversity for plantation forests. *Ecological Indicators*, *32*, 107-115. https://doi.org/10.1016/j.ecolind.2013.03.020

Creedy, T.J., Ng, W.S. and Vogler, A.P. (2019) 'Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy', *Ecology and Evolution*, 9(6), pp. 3105–3116. https://doi.org/10.1002/ece3.4839.

Cuff, J. P., Windsor, F. M., Tercel, M. P., Kitson, J. J., & Evans, D. M. (2021). Overcoming the pitfalls of merging dietary metabarcoding into ecological networks. *Methods in Ecology and Evolution*, *13*(3), 545-559. https://doi.org/10.1111/2041-210X.13796.

Del-Val, E., Ramírez, E., & Astier, M. (2021). Comparison of arthropod communities between high and low input maize farms in Mexico. *CABI Agriculture and Bioscience*, *2*, 1-10. https://doi.org/10.1186/s43170-021-00060-9

Depraz, S., Sanial, E., Catalán, A. K. R., & Rojas, A. S. (2017). Less protection for better conservation? A politicised relationship between a city and its protected area in the vicinity of Nevado de Toluca (Mexico). *Articulo - Journal of Urban Research*, *16*, Article 16. https://doi.org/10.4000/articulo.3261

Dong, Y. W., Huang, X. W., Wang, W., Li, Y., & Wang, J. (2016). The marine 'great wall'of China: local-and broad-scale ecological impacts of coastal infrastructure on intertidal macrobenthic communities. *Diversity and Distributions*, *22*(7), 731-744. https://doi.org/10.1111/ddi.12443

Edwards, D. P., Magrach, A., Woodcock, P., Ji, Y., Lim, N. T.-L., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Khen, C. V., Chung, A. Y. C., Reynolds, G., Fisher, B., Laurance, W. F., Wilcove, D. S., Hamer, K. C., & Yu, D. W. (2014). Selective-logging and oil palm: Multitaxon impacts, biodiversity indicators, and trade-offs for conservation planning. *Ecological Applications*, *24*(8), 2029-2049. https://doi.org/10.1890/14-0010.1

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, *7*(17), 6918-6926. https://doi.org/10.1002/ece3.3192

Elbrecht, V., Vamos, E. E., Steinke, D., & Leese, F. (2018). Estimating intraspecific genetic diversity from community DNA metabarcoding data. *PeerJ*, *6*, e4644. https://doi.org/10.7717/peerj.4644.

Elbrecht, V., Braukmann, T. W., Ivanova, N. V., Prosser, S. W., Hajibabaei, M., Wright, M., ... & Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, *7*, e7745. https://doi.org/10.7717/peerj.7745.

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—sequence relationships with an innovative metabarcoding protocol. *PloS one*, *10*(7), e0130324. https://doi.org/10.1371/journal.pone.0130324

Elbrecht, V., & Leese, F. (2017). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science*, *5*, 11. https://doi.org/10.3389/fenvs.2017.00011

Emerson, B. C., Borges, P. A., Cardoso, P., Convey, P., deWaard, J. R., Economo, E. P., ... & Andújar, C. (2023). Collective and harmonized high throughput barcoding of insular arthropod biodiversity: Toward a Genomic Observatories Network for islands. *Molecular Ecology*, *32*(23), 6161-6176. https://doi.org/10.1111/mec.16683

Feng, Y., Wang, J., Zhang, J., Qi, X., Long, W., Ding, Y., & Liu, L. (2023). Soil microbes support Janzen's mountain passes hypothesis: The role of local-scale climate variability along a tropical

montane gradient. *Frontiers in Microbiology*, *14*, 1135116. https://doi.org/10.3389/fmicb.2023.1135116

Fernandes, K., van der Heyde, M., Coghlan, M., Wardell-Johnson, G., Bunce, M., Harris, R., & Nevill, P. (2019). Invertebrate DNA metabarcoding reveals changes in communities across mine site restoration chronosequences. *Restoration Ecology*, *27*(5), 1177-1186. https://doi.org/10.1111/rec.12976

Ferrari, L., Conticelli, S., Vaggelli, G., Petrone, C. M., & Manetti, P. (2000). Late Miocene volcanism and intra-arc tectonics during the early development of the Trans-Mexican Volcanic Belt. *Tectonophysics*, *318*(1-4), 161-185. https://doi.org/10.1016/S0040-1951(99)00310-8

Ferrari, L., Orozco-Esquivel, T., Manea, V., & Manea, M. (2012). The dynamic history of the Trans-Mexican Volcanic Belt and the Mexico subduction zone. *Tectonophysics*, *522*, 122-149. https://doi.org/10.1016/j.tecto.2011.09.018

Forestry Commission (2011). UK Forestry Standard Guidelines: forests and biodiversity, Forestry Commission, Edinburgh, 108 pp

Fuller, R. J., Oliver, T. H., & Leather, S. R. (2008). Forest management effects on carabid beetle communities in coniferous and broadleaved forests: Implications for conservation. *Insect Conservation and Diversity*, *1*, 242-252. https://doi.org/10.1111/j.1752-4598.2008.00032.x

Gálvez-Reyes, N., Arribas, P., Andújar, C., Emerson, B. C., Piñero, D., & Mastretta-Yanes, A. (2021). Dispersal limitations and long-term persistence drive differentiation from haplotypes to communities within a tropical sky-island: Evidence from community metabarcoding. *Molecular Ecology*, *30*(24), 6611-6626. https://doi.org/10.1111/mec.16195

Gibson, J., Shokralla, S., Porter, T. M., King, I., van Konynenburg, S., Janzen, D. H., ... & Hajibabaei, M. (2014). Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics. *Proceedings of the National Academy of Sciences*, *111*(22), 8007-8012. https://doi.org/10.1073/pnas.1406468111.

Gómez-Rodríguez, C., & Baselga, A. (2018). Variation among European beetle taxa in patterns of distance decay of similarity suggests a major role of dispersal processes. *Ecography*, 1825–1834. https://doi.org/10.1111/ecog.03693

Gómez-Rodríguez, C., Miller, K. E., Castillejo, J., Iglesias-Piñeiro, J., & Baselga, A. (2019). Understanding dispersal limitation through the assessment of diversity patterns across phylogenetic scales below the species level. *Global Ecology and Biogeography*, *28*(3), 353-364. https://doi.org/10.1111/geb.12857

Gómez-Tuena, A., Orozco-Esquivel, M. T., & Ferrari, L. (2007). Igneous petrogenesis of the Trans-Mexican volcanic belt. https://doi.org/10.1130/2007.2422(05)

Gollan, J. R., de Bruyn, L. L., Reid, N., & Wilkie, L. (2013). Monitoring the ecosystem service provided by dung beetles offers benefits over commonly used biodiversity metrics and a traditional trapping method. *Journal for Nature Conservation*, *21*(3), 183-188. https://doi.org/10.1016/j.jnc.2012.12.004

Graham, C. H., & Fine, P. V. A. (2008). Phylogenetic beta diversity: Linking ecological and evolutionary processes across space in time. *Ecology letters*, *11*(12), 1265-1277. https://doi.org/10.1111/j.1461-0248.2008.01256.x

Hajibabaei, M., Baird, D. J., Fahner, N. A., Beiko, R., & Golding, G. B. (2016). A new way to contemplate Darwin's tangled bank: How DNA barcodes are reconnecting biodiversity science and biomonitoring. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *371*, 20150330. https://doi.org/10.1098/rstb.2015.0330

Hajibabaei, M., Spall, J. L., Shokralla, S., & van Konynenburg, S. (2012). Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental

barcoding of DNA from preservative ethanol. *BMC ecology*, *12*(1), 28. https://doi.org/10.1186/1472-6785-12-28

Hubbell, S. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press; JSTOR. https://doi.org/10.2307/j.ctt7rj8w

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P. M., Woodcock, P., Edwards, F. a, Larsen, T. H., Hsu, W. W., Benedick, S., Hamer, K. C., Wilcove, D. S., Bruce, C., Wang, X., Levi, T., Lott, M., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology letters*. https://doi.org/10.1111/ele.12162

Janzen, D. H. (1967). Why Mountain Passes are Higher in the Tropics Author (s): Daniel H. Janzen Source: The American Naturalist , Vol. 101 , No. 919 (May—Jun ., 1967), pp. 233-249 Published by: The University of Chicago Press for The American Society of Naturalist. *The American Naturalist*, *101*(919), 233–249.

Kozak, K. H., & Wiens, J. J. (2006). Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution*, *60*(12), 2604. https://doi.org/10.1554/06-334.1

Krehenwinkel, H., Fong, M., Kennedy, S., Huang, E. G., Noriyuki, S., Cayetano, L., & Gillespie, R. (2018). The effect of DNA degradation bias in passive sampling devices on metabarcoding studies of arthropod communities and their associated microbiota. *PLoS one*, *13*(1), e0189188. https://doi.org/10.1371/journal.pone.0189188

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific reports*, *7*(1), 17668. https://doi.org/10.1038/s41598-017-17333-x.

Lambeck, R. J. (1997). Focal species: A multi-species umbrella for nature conservation. *Conservation Biology*, *11*(4), 849-856. https://doi.org/10.1046/j.1523-1739.1997.96319.x

Lange, M., Türke, M., Pašalić, E., Boch, S., Hessenmöller, D., Müller, J., Prati, D., Socher, S. A., Fischer, M., Weisser, W. W., & Gossner, M. M. (2014). Effects of forest management on ground-dwelling beetles (Coleoptera; Carabidae, Staphylinidae) in Central Europe are mainly mediated by changes in forest structure. *Forest Ecology and Management*, *329*, 166-176. https://doi.org/10.1016/j.foreco.2014.06.012

Launer, A. E., & Murphy, D. D. (1994). Umbrella species and the conservation of habitat fragments: A case of a threatened butterfly and a vanishing grassland ecosystem. *Biological Conservation*, *69*(2), 145-153. https://doi.org/10.1016/0006-3207(94)90054-X

Leray, M., Knowlton, N., Ho, S. L., Nguyen, B. N., & Machida, R. J. (2019). GenBank is a reliable resource for 21st century biodiversity research. *Proceedings of the National Academy of Sciences*, *116*(45), 22651-22656. https://doi.org/10.1073/pnas.1911714116.

Li, L., Zheng, B., & Liu, L. (2010). Biomonitoring and bioindicators used for river ecosystems: Definitions, approaches and trends. *Procedia Environmental Sciences*, *2*, 1510-1524. https://doi.org/10.1016/j.proenv.2010.10.164

Liu, M., Baker, S. C., Burridge, C. P., Jordan, G. J., & Clarke, L. J. (2020). DNA metabarcoding captures subtle differences in forest beetle communities following disturbance. *Restoration Ecology*, *28*(6), 1475-1484. https://doi.org/10.1111/rec.13236

Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., & Burridge, C. P. (2019). A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology*, *45*(3), 373-385. https://doi.org/10.1111/een.12831

Lim, J. Y., Patiño, J., Noriyuki, S., Cayetano, L., Gillespie, R. G., & Krehenwinkel, H. (2022). Semi-quantitative metabarcoding reveals how climate shapes arthropod community assembly

along elevation gradients on Hawaii Island. *Molecular Ecology*, *31*(5), 1416-1429. https://doi.org/10.1111/mec.16323

Lynggaard, C., Yu, D. W., Oliveira, G., Caldeira, C. F., Ramos, S. J., Ellegaard, M. R., Gilbert, M. T. P., Gastauer, M., & Bohmann, K. (2020). DNA-Based Arthropod Diversity Assessment in Amazonian Iron Mine Lands Show Ecological Succession Towards Undisturbed Reference Sites. *Frontiers in Ecology and Evolution*, *8*. https://doi.org/10.3389/fevo.2020.590976

Ma, L., Liu, L., Lu, Y., Chen, L., Zhang, Z., Zhang, H., ... & Zhang, J. (2022). When microclimates meet soil microbes: Temperature controls soil microbial diversity along an elevational gradient in subtropical forests. *Soil Biology and Biochemistry*, *166*, 108566. https://doi.org/10.1016/j.soilbio.2022.108566

Magurran, A. E. (2016). How ecosystems change. *Science*, *351*(6272), 448-449. https://doi.org/10.1126/science.aad6758

Malcolm, D. C., Mason, W. L., & Clarke, G. C. (2001). The transformation of conifer forests in Britain Ð regeneration, gap size and silvicultural systems. *Forest Ecology and Management*, 17. https://doi.org/10.1016/S0378-1127(00)00692-7

Mastretta-Yanes, A., Cao, R., Arzata, S., Quadri, P., Espinosa, T., Arredondo, L., & Piñero, D. (2014). *¿Será exitosa la estrategia de cambio de categoría para mantener la biodiversidad del Nevado de Toluca? 12*, 7-17.

Mastretta-Yanes, A., Moreno-Letelier, A., Piñero, D., Jorgensen, T. H., & Emerson, B. C. (2015). Biodiversity in the Mexican highlands and the interaction of geology, geography and climate within the Trans-Mexican Volcanic Belt. *Journal of Biogeography*, *42*(9), 1586-1600. https://doi.org/10.1111/jbi.12546

Meyer, A., Boyer, F., Valentini, A., Bonin, A., Ficetola, G. F., Beisel, J. N., ... & Usseglio-Polatera, P. (2020). Morphological vs. DNA metabarcoding approaches for the evaluation of stream ecological status with benthic invertebrates: Testing different combinations of markers and strategies of data filtering. *Molecular Ecology*, *30*(13), 3203-3220. https://doi.org/10.1111/mec.15723

McCormack, J. E., Huang, H., Knowles, L. L., Gillespie, R., Clague, D., & Gillespie, R. G. (2009). Encyclopedia of islands.

McGill, B. J. (2010). Towards a unification of unified theories of biodiversity. *Ecology Letters*, *13*(5), 627-642. https://doi.org/10.1111/j.1461-0248.2010.01449.x

Morrone, J. J., & Márquez, J. (2001). Halffter's Mexican Transition Zone, beetle generalized tracks, and geographical homology: Halffter's Mexican Transition Zone. *Journal of Biogeography*, *28*(5), 635-650. https://doi.org/10.1046/j.1365-2699.2001.00571.x

Múrria, C., Bonada, N., Vellend, M., Zamora-Muñoz, C., Alba-Tercedor, J., Sainz-Cantero, C. E., ... & Derka, T. (2017). Local environment rather than past climate determines community composition of mountain stream macroinvertebrates across Europe. *Molecular ecology*, *26*(21), 6085-6099. https://doi.org/10.1111/mec.14346

Noguerales, V., Meramveliotakis, E., Castro-Insua, A., Andújar, C., Arribas, P., Creedy, T. J., Overcast, I., Morlon, H., Emerson, B. C., Vogler, A. P., & Papadopoulou, A. (2021). Community metabarcoding reveals the relative role of environmental filtering and spatial processes in metacommunity dynamics of soil microarthropods across a mosaic of montane forests. *Molecular Ecology*, *n/a*(n/a), 1-19. https://doi.org/10.1111/mec.16275

Noss, R. F. (1990). Indicators for Monitoring Biodiversity: A Hierarchical Approach. *Conservation Biology*, *4*, 355-364. https://doi.org/10.1111/j.1523-1739.1990.tb00309.x

Obrist, M. K., & Duelli, P. (2010). Rapid biodiversity assessment of arthropods for monitoring average local species richness and related ecosystem services. *Biodiversity and Conservation*, *19*(8), 2201-2220. https://doi.org/10.1007/s10531-010-9832-y

Paillet, Y., Bergès, L., Hjältén, J., Odor, P., Avon, C., Bernhardt-Römermann, M., Bijlsma, R.-J., De Bruyn, L., Fuhr, M., Grandin, U., Kanka, R., Lundin, L., Luque, S., Magura, T., Matesanz, S., Mészáros, I., Sebastià, M.-T., Schmidt, W., Standovár, T., … Virtanen, R. (2010). Biodiversity differences between managed and unmanaged forests: Meta-analysis of species richness in Europe. *Conservation biology : the journal of the Society for Conservation Biology*, *24*(1), 101-112. https://doi.org/10.1111/j.1523-1739.2009.01399.x

Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, *15*(4), 819-830. https://doi.org/10.1111/1755-0998.12355

Polato, N. R., Gill, B. A., Shah, A. A., Gray, M. M., Casner, K. L., Barthelet, A., Messer, P. W., Simmons, M. P., Guayasamin, J. M., Encalada, A. C., Kondratieff, B. C., Flecker, A. S., Thomas, S. A., Ghalambor, C. K., Poff, N. L., Funk, W. C., & Zamudio, K. R. (2018). Narrow thermal tolerance and low dispersal drive higher speciation in tropical mountains. *Proceedings of the National Academy of Sciences*, *115*(49), 12471–12476. https://doi.org/10.1073/pnas.1809326115

Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-Bravo, D., Rasmussen, C. M. Ø., Richardson, K., Rosing, M. T., Whittaker, R. J., & Fjeldså, J. (2019). Building mountain biodiversity: Geological and evolutionary processes. *Science*, *365*(6458), 1114–1119. https://doi.org/10.1126/science.aax0151

Rambo, T., Schowalter, T., & North, M. (2014). Forest Ecology and Management Canopy arthropod responses to thinning and burning treatments in old-growth mixed-conifer forest in the Sierra Nevada , California. *Forest Ecology and Management*, *326*, 91-100. https://doi.org/10.1016/j.foreco.2014.04.014

Rosindell, J., Hubbell, S. P., & Etienne, R. S. (2011). The Unified Neutral Theory of Biodiversity and Biogeography at Age Ten. *Trends in Ecology and Evolution*, *26*(7), 340-348. https://doi.org/10.1016/j.tree.2011.03.024

Rzedowski, J. (2006). Vegetación de México. 1ra. *Edición digital, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México*, *504*. https://www.biodiversidad.gob.mx/publicaciones/librosDig/pdf/VegetacionMxPort.pdf

Salces-Castellano, A., Patiño, J., Alvarez, N., Andújar, C., Arribas, P., Braojos-Ruiz, J. J., ... & Emerson, B. C. (2019). Climate drives community-wide divergence within species over a limited spatial scale: evidence from an oceanic island. *Ecology Letters*, *23*(2), 305-315. https://doi.org/10.1111/ele.13433

Sarremejane, R., Mykrä, H., Bonada, N., Aroviita, J., & Muotka, T. (2017). Habitat connectivity and dispersal ability drive the assembly mechanisms of macroinvertebrate communities in river networks. *Freshwater Biology*, *62*(6), 1073-1082. https://doi.org/10.1111/fwb.12926

Semarnat (Secretaría de Medio Ambiente y Recursos Naturales). (2013). Programa Sectorial de Medio Ambiente y Recursos Naturales 2013-2018.

Smith, G. F., Gittings, T., Wilson, M., French, L., Oxbrough, A., O'Donoghue, S., O'Halloran, J., Kelly, D. L., Mitchell, F. J. G., Kelly, T., Iremonger, S., McKee, A.-M., & Giller, P. (2008). Identifying practical indicators of biodiversity for stand-level management of plantation forests. *Biodiversity and Conservation*, *17*(5), 991-1015. https://doi.org/10.1007/s10531-007-9274-3

Sheldon, K. S., Huey, R. B., Kaspari, M., & Sanders, N. J. (2018). Fifty Years of Mountain Passes: A Perspective on Dan Janzen's Classic Article. *The American Naturalist*, *191*(5), 553–565. https://doi.org/10.1086/697046

Shimabukuro, E. M., Gómez-Rodríguez, C., Lamas, C. J. E., & Baselga, A. (2023). Mountain passes are higher at low latitudes for madicolous insect communities of the Neotropical region. *Diversity and Distributions*, *29*(9), 1118-1128. https://doi.org/10.1111/ddi.13747

Schnell, I.B., Bohmann, K. and Gilbert, M.T.P. (2015) 'Tag jumps illuminated--reducing sequence-to-sample misidentifications in metabarcoding studies.', *Molecular ecology resources*, 15(6), pp. 1289–303. https://doi.org/10.1111/1755-0998.12402.

Shokralla, S., Porter, T. M., Gibson, J. F., Dobosz, R., Janzen, D. H., Hallwachs, W., ... & Hajibabaei, M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Scientific reports*, *5*(1), 9687. https://doi.org/10.1038/srep09687.

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8), 1789-1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x

Torres-Rojo, J. M., Moreno-Sánchez, R., & Mendoza-Briseño, M. A. (2016). Sustainable Forest Management in Mexico. *Current Forestry Reports*, *2*(2), 93-105. https://doi.org/10.1007/s40725-016-0033-0

Tsuji, S., Shibata, N., Sawada, H., & Ushio, M. (2020). Quantitative evaluation of intraspecific genetic diversity in a natural fish population using environmental DNA analysis. *Molecular Ecology Resources*, *20*(5), 1323-1332. https://doi.org/10.1111/1755-0998.13200

Turon, X., Antich, A., Palacín, C., Præbel, K., & Wangensteen, O. S. (2020). From metabarcoding to metaphylogeography: separating the wheat from the chaff. *Ecological Applications*, *30*(2), e02036. https://doi.org/10.1002/eap.2036.

Uscanga, A., López, H., Piñero, D., Emerson, B. C., & Mastretta-Yanes, A. (2021). Evaluating species origins within tropical sky-islands arthropod communities. *Journal of Biogeography*, *48*(9), 2199-2210. https://doi.org/10.1111/jbi.14144

Van Klink, R., Bowler, D. E., Gongalsky, K. B., Swengel, A. B., Gentile, A., & Chase, J. M. (2020). Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science*, *368*(6489), 417-420. https://doi.org/10.1126/science.aax9931

van der Loos, L.M. and Nijland, R. (2021) 'Biases in bulk: DNA metabarcoding of marine communities and the methodology involved', *Molecular Ecology*, 30(13), pp. 3270–3288. https://doi.org/10.1111/mec.15592.

Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*. https://doi.org/10.1086/652373

Wang, X., Hua, F., Wang, L., Wilcove, D. S., & Yu, D. W. (2019). The biodiversity benefit of native forests and mixed-species plantations over monoculture plantations. *Diversity and Distributions*, *25*(11), 1721-1735. https://doi.org/10.1111/ddi.12972

Wiens, J. J. (2004). Speciation and ecology revisited: Phylogenetic niche conservatism and the origin of species. *Evolution*, *58*(1), 193-197. https://doi.org/10.1111/j.0014-3820.2004.tb01586.x

Wiens, J. J., Camacho, A., Goldberg, A., Jezkova, T., Kaplan, M. E., Lambert, S. M., ... & Walls, R. L. (2019). Climate change, extinction, and Sky Island biogeography in a montane lizard. *Molecular ecology*, *28*(10), 2610-2624. https://doi.org/10.1111/mec.15073

Yang, C., Wang, X., Miller, J. A., de Blécourt, M., Ji, Y., Yang, C., ... & Douglas, W. Y. (2014). Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. *Ecological Indicators*, *46*, 379-389. https://doi.org/10.1016/j.ecolind.2014.06.028.

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring.

Methods in Ecology and Evolution, *3*(4), 613-623. https://doi.org/10.1111/j.2041-210X.2012.00198.x

Zimmermann, J., Glöckner, G., Jahn, R., Enke, N., & Gemeinholzer, B. (2015). Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Molecular ecology resources*, *15*(3), 526-542. https://doi.org/10.1111/1755-0998.12336

Zhang, K., Lin, S., Ji, Y., Yang, C., Wang, X., CY, Y., H, W., Jiang, H., Harrison, R., & Yu, D. (2016). Plant diversity accurately predicts insect diversity in two tropical landscapes. *Molecular ecology*, *68125438*. https://doi.org/10.1111/MEC.13770

Zhang, Y., Aaron Hogan, J., Crowther, T. W., Xu, S., Zhao, R., Song, P., ... & Yang, J. (2023). Drivers and mechanisms that contribute to microbial β-diversity patterns and range sizes in mountains across a climatic variability gradient. *Ecography*, e07049. https://doi.org/10.1111/ecog.07049

Zou, Y., Sang, W., Wang, S., Warren-Thomas, E., Liu, Y., Yu, Z., Wang, C., & Axmacher, J. C. (2015). Diversity patterns of ground beetles and understory vegetation in mature, secondary, and plantation forest regions of temperate northern China. *Ecology and evolution*, *5*(3), 531-542. https://doi.org/10.1002/ece3.1367

Zizka, V.M.A., Weiss, M. and Leese, F. (2020) 'Can metabarcoding resolve intraspecific genetic diversity changes to environmental stressors? A test case using river macrozoobenthos', *Metabarcoding and Metagenomics*, 4, p. e51925. https://doi.org/10.3897/mbmg.4.51925.

ANEXO 1: Material suplementario artículo 1

Insect Conservation and Diversity

Gálvez-Reyes *et al.* Supporting Information

Suplementary Information for: Towards DNA metabarcoding-based haplotype for monitoring terrestrial arthropod communities

Nancy Gálvez-Reyes*[1,2], Daniel Piñero[1] & Alicia Mastretta-Yanes*[3,4].

Supplemental material table S1. Authors, collection traps, arthropods, HTS sequencing platform and list of primers normally used in metabarcoding of arthropod communities.

AUTHORS	TRAP/ BULK-SAMPLE/ HABITAT	ARTHROPOD	MARKER/ HTS/	PRIMER /SIZE
Yu et al. (2012)	Malaise / 7 bulk-samples / prefectures	Lepidoptera, Diptera, Hymenoptera, Coleoptera, Hemiptera, Psocoptera, Arachnida, Blattaria, Plecoptera, Trichoptera, Ephemeroptera, Odonata.	COI/ Roche 454	Fol-degen-for/Fol-degen-rev (658 bp)
Ji et al. (2013)	Light-trap, Pitfall / //Subtropical Forest and Temperate woodland	Lepidoptera, hymenoptera, arachnids, coleoptera,	COI/ Roche 454	Fol-degen-for/Fol-degen-rev (658 bp).

Gibson et al. (2014)	Malaise trap // Conservation Area.	Coleoptera, Diptera, Hymenoptera, Lepidoptera Araneae, Blattodea, Collembola, Decapoda, Hemiptera, Megaloptera, Neuroptera, Orthoptera, Psocoptera, Thysanoptera, Trombidiformes, Mesostigmata.	COI, 16S, 18S/ Roche 454 and Illumina MiSeq	ArF1xArR2/ArF10xArR3, ArF1xArR3/ArF1xArR6, ArF1xArR2/ArF4xArR5, ArF10xArR3/ArF10xArR5, ArF1xArR3/ArF1xArR6, ArF1xArR6/ArF10xArR3, 16Sv4F/16Sv6R, 16S v3F/16Sv3R, 18SEukF/18SEukR, 18SNemF/18SNemR.
Yang et al. (2014).	Soil, leaf-litter (Winkler), Malaise-trap, and canopy-fogging samples // Protected Forest and plantation.	Arachnida, insecta, collembola, Myriapoda.	COI, 18S/ Roche 454	Fol-degen-for/Fol-degen-rev (658 bp). SSU FO4/SSU R22 (450 bp).
Brandon-Mong et al. (2015)	Malaise trap//	Araneae, Blattodea, Coleoptera, Diptera, Hemiptera, Hymenoptera, Lepidoptera, Mantodea, Odonata, Orthoptera, Collembola.	COI / Illumina MiSeq	mlColintF/HCO2198 (313 bp) LepF1/MLepF1-Rev (218 bp)
Bittleston et al. (2016)	/3 bulk-samples/ Nepenthes pitcher plants	Orethrella, Culicoides, Drosophilidae, Bradysia, Aedes, Culex.	18S/ Illumina MiSeq	Euk1391f/ EukBr
Arribas et al. (2016)	Deep soil sampling and mesofauna extraction / Two soil samples / Grassland habitats.	Acari, Collembola, Oribatida, Astigmata, Trombidiformes, Mesostigmata, Entomobryomorpha and Poduromorpha.	cox1/ Illumina MiSeq	FoldF-FoldR / LCO1490_short-HCO2198_short (650 bp).
Beng et al. (2016)	Winkler// Forests and plantations	Arachnida, Blatodea, coleoptera, diptera, hemiptera, hymenoptera, orthoptera, chilopod	COI/ Illumina MiSeq	MhemF /dgHCO2198 (400 bp).
Sinniger et al. 2016	39 deep-sea sediment	meiobenthic taxa (OTUS)	subunit (18S) / 454 Roche GSFLX	R22mod/F04
Elbrecht et al. (2017)	18 sample water	Benthic macroinvertebrates	COI/ Illumina HiSeq 2500	BF1/BR1_A; BF1/BR1_B; BF2/BR2_A; BF2/BR2_B; BF2/BR1_A; BF2/BR1_B; BF1/BR2_A; BF1/BR2_B

Elbrecht and Leese 2017	freshwater	Freshwater Macroinvertebrate	COI/ HiSeq 2500	BF1/ BR1, BF2/ BR2
Wood et al. (2017)	/24 soils samples/ natural and planted forest, unimproved grassland, and vineyards.	Arachnids, Collembola, Insects.	COI, 18S/ 454 GS-FLX	LCO1490/HCO2198 (526 bp) 18S11b/18S2.
Bowser et al. (2017)	Sweep net samples/ /Grasslands	Hemiptera, Diptera, Coleoptera, Hymenoptera, Lepidoptera.	COI/ Illumina MiSeq	mlCOIintF /HCO2198 (313 bp). Leray et al. (2013), and Brandon-Mong et al. (2015)
Pedro et al. (2017)	Trampa pitfall / 12 bulk arthropods / coffee (Coffea arabica L.) growing	Arachnida, Collembola, Blattodea, Coleoptera, Diptera, Hymenoptera, Lepidoptera, Mecoptera, Orthoptera.	COI /Roche 454	mlCOIintF_adF / jgHCO2198 454A-MID-adF/ jgHCO2198_454R.
Fonseca et al. (2017)	Sediment sample in depth/ with corer methodology / meiofauna antartic	Arthropods	18 S rDNA/ Roche 454 GSFLX	SSU_ F04/ SSU_R22mod (450 bp).
Krehenwinkel, Kennedy, Pekár, & Gillespie, (2017)	funnel-web spiders / gut content of 27 adult Hololena spp. / UC Berkeley campus	Isopoda, Coleoptera, Diptera, Hemiptera, Hymenoptera, Lepidoptera, Orthoptera, Psocoptera, Araneae, Collembola, Dermaptera, Opiliones, Psocoptera, Thysanoptera.	COI/ Illumina MiSeq	mlCOIintF/Fol-degen-rev (363bp) Uni-MinibarF1/Uni-MinibarR1 (176bp) ZBJ-ArtF1c/ZBJ-ArtR2c (210bp) ArF1/mlCOIintR (127 bp) mlCOIintF/ARR (252bp).
Krehenwinkel et al. (2017)	beat sheets / 23 mock communities / Native rainforests	Araneae, Blattodea, Coleoptera, Collembola, Dermaptera, Diptera, Ephemenoptera, Hemiptera, Hymenoptera, Isopoda, Lepidoptera, Myriapoda, Neuroptera, Odonato, Orthoptera, Psocopotera, Trichoptera.	COI_A, COI_B, CytB, 12SrDNA, 18SrDNA-V1-2, 18SrDNA-V6-7, 28SrDNAD6, Histone-H3	ArF1/Fol-degen-rev (418 bp) mlCOIintF/Fol-degen-rev27 (313 bp) CB3 / CB4 (350 bp) 12sai / 12sbi (348 bp) SSU_FO4 / SSU_R22 (380 bp) 18s_2F / 18s_4R (304 bp) 28s_3F / 28s_4R (318 bp) H3aF58 / H3aR58 (328bp).

			/Illumina MiSeq	
Wood et al. 2017	natural forest, planted forest, unimproved grassland, improved grassland, and vineyards	bacteria, fungi, plants, and metazoans: arachnids, Collembola, insects, and nematodes **(OTUS)**	16S/ ITS/ 18S /COI 454 GS-FLX	
Andújar et al. (2017)	Surber sample /freshwater / River	Coleoptera, Ephemeroptera, Trichoptera, Arachnida, Diptera	SSU, bc50 and bc30 (cox1) / Illumina MiSeq	Fol-degen-for/Ill_C_R, Ill_B_F/Fol-degen-rev (420 bp) SSU-FO4/SSU-R22 (300 to 400).
Theissinger et al. (2018)	Emergence traps// Wetlands	Chironomid	COI / MiSeq Illumina	BF2 / BR1.
Krehenwinkel et al. (2018)	Malaise traps / Two bulk-samples (15 mock communities) /Native rainforests	Arthropods	COI / Illumina MiSeq	ARF1/Fol-degen-rev (418 bp) mlCOIintF/Fol-degen-rev (313 bp).
Oliverio et al. (2018)	80 soil samples/grass, forest, crops	Arthropods: phylotypes (unique sequence variants)	COI/158 /Illumina MiSeq	Primers as per Madden et al. 2016. Madden et al. (2016).
Elbrecht et al. 2018		Plecoptera, Trichoptera, Coleoptera and Isopoda (haplotypes)	COI / Illumina MiSeq and HiSeq systems	(250 bp).
Galan et al. (2018)	336 bat faecal samples	Arthropods	COI / Illumina MiSeq	LepF1 / EPT-long-univR (133 bp).
Barsoum et al. (2019)	forest plantation stands / Malaise and Pitfall traps		COI / Roche GS FLX	Fol-degen-for / Fol-degen-rev.
Thomsen and Sigsgaard (2019)	56 individual flowers	Arthropods	COI, 16S / IlluminaNextSeq500	COI~211bp and16S ribosomal RNA ~160bp
Wang et al. (2019)	71 samples / forest	Araneae, Formicidae, Tachinidae, Phoridae, Braconidae, Thysanoptera, Syrphidae, Isoptera, Lepidoptera,	COI / Illumina MiSeq	LCO1490 / mlCOIintR (319 bp).

		Hemiptera, Diptera, Orthoptera, Formicidae, Thysanoptera.		
Porter et al. 2019	216 soil samples/ boreal forest	Terrestrial arthropod fauna / exact sequence variants (ESVs)	COI / Illumina MiSeq	F230R_modN / 230R_modN.
Creedy et al. 2019	canopy fogging Liquidambar styracaflua	Coleoptera, Formicidae, Acari, Araneae, Hymenoptera, Hemiptera, Diptera and Collembola.	COI / 418 bp Illumina MiSeq /OTUs	Ill_B_F (Shokralla et al., 2015) and Fol_degen_rev (Yu et al., 2012).
Bush et al. 2020	126 and 138 samples / aquatic invertebrates	Arachnidae, Insecta, Gastropoda, Collembola.	CO1 amplicon using two complementary primers, BE/BR5 and F230R	
Arribas et al. 2020	144 soil samples (a) a sample containing the superficial soil layer (SUP), by extracting one square metre of leaf litter and humus up to 5 cm deep and (b) a sample of the corresponding deep soil layer (DEEP), by digging the substrate of a 30 cm diameter core to 30 cm depth, comprising ca. 20 litres of soil. / grasslands and forests	arthropod mesofauna Acari, Collembola and Coleoptera Haplotypes	COI / Illumina MiSeq (418 bp)	B_F 5' CCIGAYATRGCITTYCCICG 3' (Shokralla et al., 2015) and Fol-degen-R 5' TANACYTCNGGRTGNCCRAARAAYCA 3' (Yu et al., 2012).
Tsuji et al. 2020	Water samples	haplotypes	D-loop region of mitochondrial DNA Miseq 7 290 bp	
Turon et al. 2020	51 samples	benthic marine communities MOTUS	COI / 313 bp	
Kirse et al. (2021)	180 soil samples and 54 Malaise traps/forest	Arthropods	COI/ Illumina MiSeq	mlCOIintF / dgHCO2198 (313 bp).
Cai et al. (2021)	60 samples/ Malaise traps / forest	Diptera, Lepidoptera, Hymenoptera, Coleoptera y hemiptera.	COI/Illumina MiSeq platform	mlCOIintF–Fol-degen-rev primers (313 bp).

Kirse et al. (2021)	162 soil samples/ forest	Annelida, Arthropoda, Chordata, Cnidaria, Gastrotricha, Mollusca, Nematoda, Platyhelminthes, Rotifera, Tardigrada/ Annelida, Arthropoda, Chordata, Mollusca, Nematoda, Tardigrada	18S and COI/ Illumina Miseq	TAReuk454FWD1/TAReukREV3r (380 bp) and mlCOIintF/dgHCO2198 (313 bp).
Steinke et al. 2021	70 bulk-samples/ Malaise traps/ grassland and forest	Diptera, Hymenoptera, Lepidoptera, Hemiptera and Coleoptera.	COI/ Illumina MiSeq (421bp)	BF2 / BR2.
Shum and Palumbi 2021		haplotypes per MOTU	COI 313 bp	COI, mICOIintF: GGWACWGGWTGAACWGTWTAYCCYCC; Leray et al., 2013, matched to jgHCO2198: TAIACYTCIGGRTGICCRAARAAYCA; Geller et al., 2013).
Gálvez-Reyes et al. (2021)	42 bulk-samples / pitfalls /tropical sky island	Diptera, Collembola, Arachnida, Coleoptera, Hemiptera, Hymenoptera, Lepidoptera, and Myriapoda.	COI/ Illumina MiSeq (418 bp)	B_F 5' CCIGAYATRGCITTYCCICG 3' (Shokralla et al., 2015) and Fol-degen-R 5' TANACYTCNGGRTGNCCRAARAAYCA 3' (Yu et al., 2012).
Noguerales et al., 2021	88 soil samples / forest habitat types of the Troodos mountain	Acari, Collembola and Coleoptera.	COI/ Illumina MiSeq (418 bp)	B_F 5' CCIGAYATRGCITTYCCICG 3' (Shokralla et al., 2015) and Fol-degen-R 5' TANACYTCNGGRTGNCCRAARAAYCA 3' (Yu et al., 2012).
Lim et al. 2021		Araneae, Coleoptera, Hemiptera Lepidoptera, Orthroptera, Psocoptera,		

Supplementary material 1. **Testing trap types for collecting bulk samples.**

Sampling. Arthropods were collected in a previous work by Uscanga (2016) and Uscanga *et al.* (2021) using four different trap types: Soapy-water (S), pitfall (P) traps, Winkler (W) and beating tray (B). Sampling was done in the Transmexican Volcanic Belt, at >3,000 m.a.s.l. in alpine grasslands and *Abies religiosa* forests. The traps were set during April–March 2014. The trap types were placed according to the following details:

1) *Soapy-water (S)* for sampling pollinators. Samples were left in the field for six hours.
2) *Pitfall (P)* traps for collecting aboveground arthropods. Two sets of 20 traps distributed in 15x20 m plots were set for two weeks. Pitfall traps consisted of plastic cups of 13 cm height and 10 cm diameter, with five 2 x 1.5 cm perforations, 2 cm apart from each other, at a height of 10 cm above the base. Cups were buried to the height of the window and lids were added to prevent the entry of rainwater. Cups were filled with a mixture of 185 ml of ethanol 70% and 15 ml of glycerin, which we previously tested was enough to prevent evaporation and thus DNA degradation.
3) *Winkler (W)* traps for collecting arthropods living below ground with three soil and litter samples, separating arthropods with Berlesse funnels.
4) *Beating (B) tray net* for collecting flying arthropods. Sampling was done in a 20-m transect, with activity of 1 hour (resting every 5 min).

Wet-Lab. In total there were 12 bulk-samples of arthropod communities. Specimens were divided in three sizes according to Gálvez-Reyes *et al.,* (2021): small arthropods (*Drosophila melanogaster* <=3 mm), the head and thorax for medium-sized-arthropods (adult *Apis*

mellifera >3 mm and <=15 mm), and for large-arthropods (adult grasshopper >15 mm); only a leg was kept (Figure 2). Subsets of a few samples were weighed to determine that the head + thorax of medium sized individuals were comparable in weight to two legs of the larger size class. Each sample was processed independently for DNA extraction and metabarcoding library construction but using the same library barcode identifier followed by Gálvez-Reyes *et al.* (2021). For DNA extractions, we treated each bulk-sample with an electric homogenizer (Qiagen TissueLyser II: two times for 1 min at 30 Hz and 30 s at 30 Hz) using platinum beads with Qiagen DNeasy Blood & Tissue Kit and BSA buffer 1x. Each sample was sequenced in Illumina MiSeq 2X300.

Bioinformatic steps. The resulting paired-end reads of the 12 samples were quality filtered following procedures described by Arribas *et al.* (2020) and Gálvez-Reyes *et al.* (2021). Briefly, processing included quality checking, primer removal, pair merging, quality filtering, denoising, and clustering each library independently. The taxonomic identification of each read was done using BLAST against the *nucleotide* NCBI *nt* database (January 04, 2022; blastn -outfmt 5 -evalue 0.001). ANOVAs with post-hoc.kruskal (Bonferroni method) were then used to test for significant differences in alpha diversity among communities through four different traps. These analyses were performed using R package *vegan* (Oksanen *et al.*, 2019) and the plots were visualized with the R package ggplot2 (Wickham *et al.*, 2020).

Results. The richness of arthropods at different clustering levels by type of sampling trap allowed us to know that the trap that yielded higher richness within communities (alpha diversity by trap) was with pitfall. This trap showed a greater richness in Diptera, Arachnida,

Hemiptera, Hymenoptera and Coleoptera (Figure S1). Diptera and Arachnida presented significant differences in richness among traps, but patterns were consistent across the other groups (Figure S1).

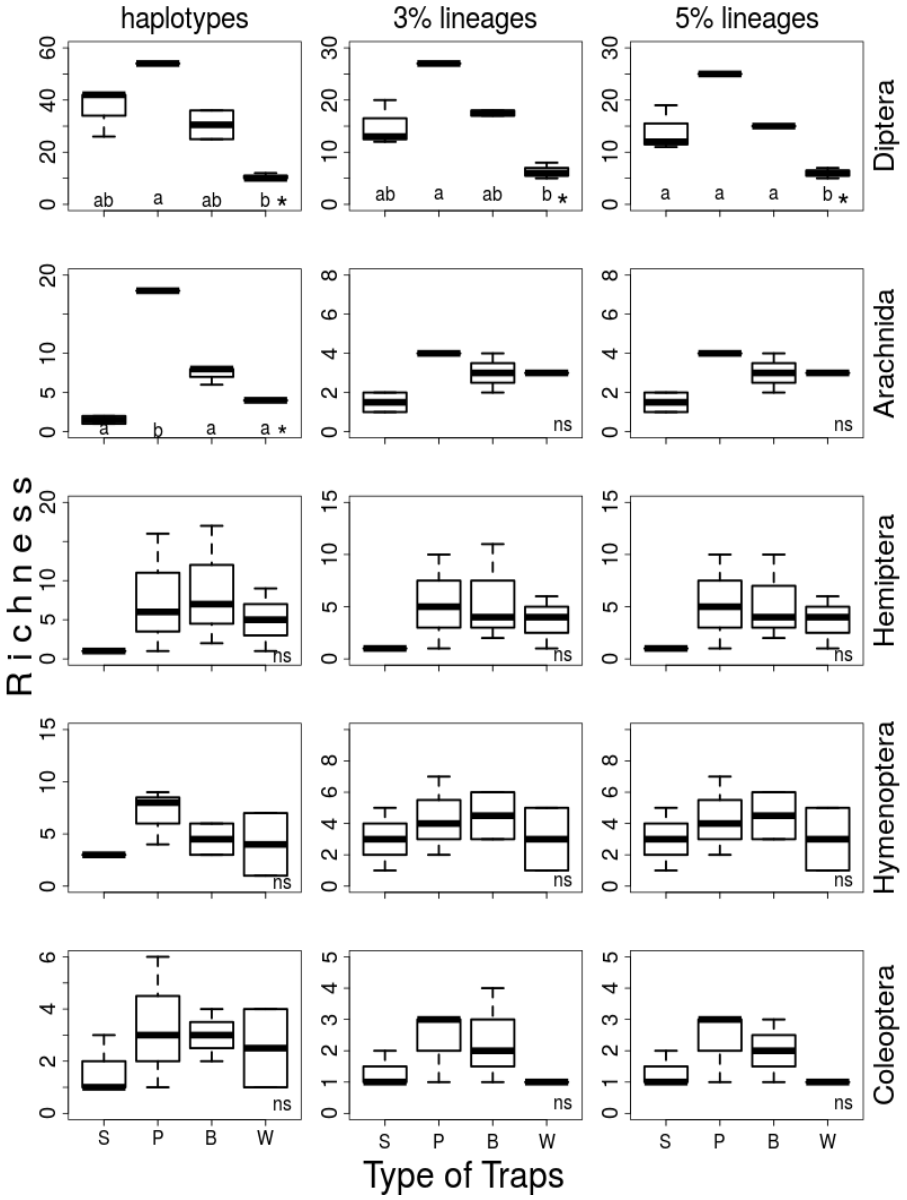

Figure S1. Arthropods richness at different clustering levels by type of sampling trap. S=Soapy-water, P=Pitfall traps, B=Beating tray, and W=Winkler traps. Sampling done in alpine grasslands and *Abies religiosa* forests. Same letters indicate that there are no significative differences among those sites. Significance codes *:p<0.05, ns: non-significant.

Supplementary material 2. **Testing effect of tissue size of DNA extraction on alpha and beta diversity.**

Sampling. We collected samples of arthropods using pitfall traps at >3,000 m.a.s.l. in *Abies religiosa* forests in the Nevado de Toluca volcano within TMVB. Sampling was performed during the rainy season of mid-August and September 2015. The distribution of sampling plots (20 traps) was the same described above, also leaving them in the field for 15 days, but in this case placing three plots of 15 x 20 m per sampling block (Gálvez-Reyes *et al.*, 2021).

Design experiment and Wet-Lab. To test the PCR bias, ligation and if sequencing was lower by performing sample processing, cleanups, tissue size separation and replicates as well as sequence filtering. The samples were from large specimens (L), small specimens (S), and a mix of large and small specimens (LS) pooled together after DNA extraction. For this, samples were processed independently for DNA extraction (bulk-samples from large and small specimens, independently) following the protocol for separating specimens in sizes according Gálvez-Reyes *et al.* (2021). For DNA extractions, we treated each bulk-sample with an electric homogenizer (Qiagen TissueLyser II: two times for 1 min at 30 Hz and 30 s at 30 Hz) using platinum beads with Qiagen DNeasy Blood & Tissue Kit and BSA buffer 1x. Library construction was done using different barcodes identifiers depending on the size of the specimens and were sequenced in Illumina MiSeq 2X300.

Bioinformatic tools. Paired-end reads of each sample were quality filtered following procedures described by Arribas *et al.* (2020) and Gálvez-Reyes *et al.* (2021). Briefly,

processing included quality checking, primer removal, pair merging, quality filtering, denoising, and clustering each library independently. The taxonomic identification of each read was done using BLAST against the *nucleotide* NCBI *nt* database (January 04, 2022; blastn -outfmt 5 -evalue 0.001). ANOVAs with post-hoc.kruskal (Bonferroni method) were then used to test for significant differences in alpha diversity among communities with different sizes of arthropods. These analyses were performed using the *vegan* (Oksanen *et al.*, 2019) package of R (RStudio Team, 2020). The analysis of non-parametric multidimensional scaling ordination (NMDS) based on Sorensen similarity was done using the *R* package *vegan* (Oksanen *et al.*, 2019) and community composition matrices. Plots were created with the *ordispider* option to visualize the compositional ordination of the communities among samples. An analysis of similarity (ANOSIM) was performed for each taxonomic group to compare arthropod community composition among samples. Plots were visualized with the R package ggplot2 (Wickham *et al.*, 2020).

Results. Our results for richness and compositions showed that soups that come from large and small specimens that were sequenced with different barcodes do not have significant differences within the pool soups, which indicates that pools can be made from the samples, as long as the extractions are done independently, and the pool is done before for the sequencing. Although large specimens (L) and small specimens (S) soups tend to have fewer ASVs, the difference with the pool (L+S) is not significant (Figure S2). Also, we observed more haplotypes in the Diptera than lineages 3% because the diversity of haplotypes is at an intraspecific level. Diptera presented the highest diversity in our samples. However, we expected that the NMDs analyzed without structure because the samples were similar and tagged with different barcodes.

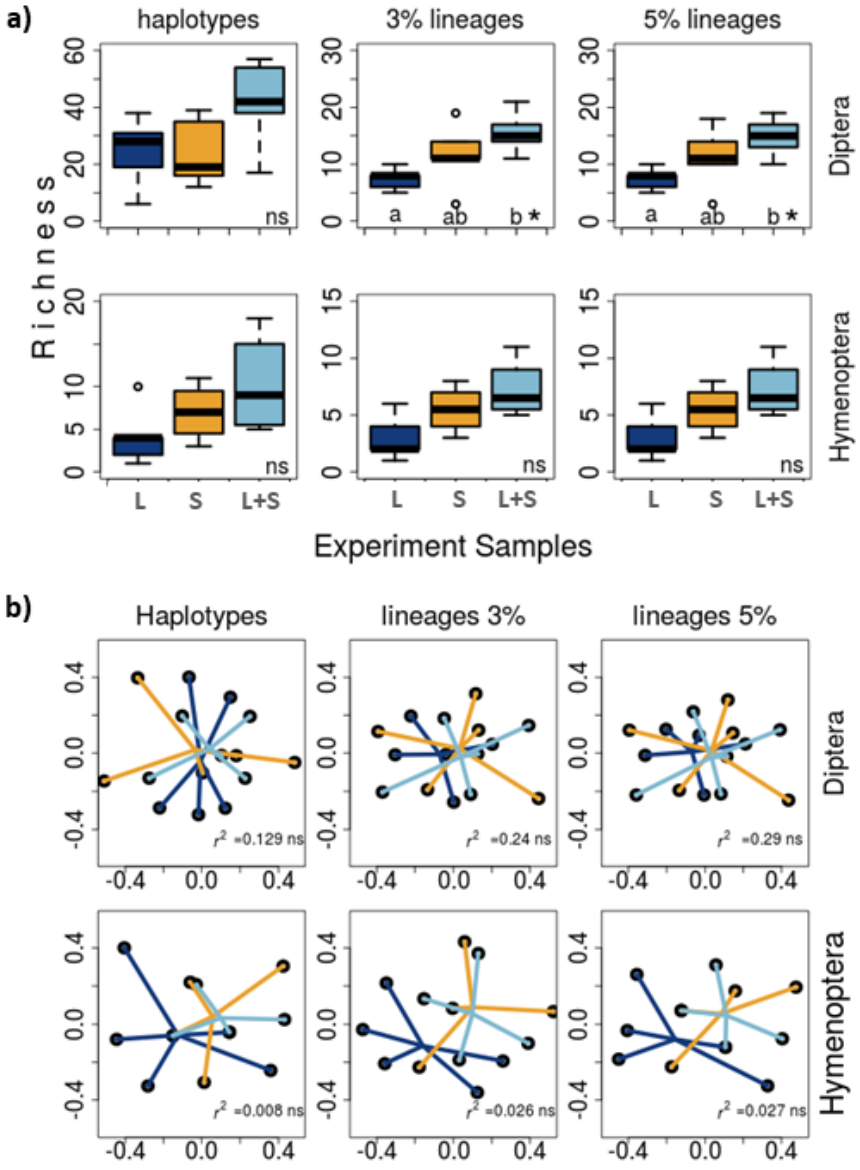

Figure S2. Effect of tissue size selection at the DNA extraction step on alpha and beta diversity. a) Diptera and Hymenoptera richness at clustering levels by different sizes of tissue. L=large specimens, S=small specimens, and L+S=Soup of large and small specimens pooled together after DNA extraction. b) Non-Metric Multidimensional scaling (NMDS) ordinations of community similarity (Simpson index, βsim) from haplotype to 5% lineages for Diptera and Hymenoptera groups. Blue=large specimens, yellow=small specimens, and

light-blue=Pool of large and small specimens. ns=non-significant differences of community structure estimated with ANOSIM.

Supplementary material 3. **Ecological statistics of ASV-based data and DNA metabarcoding.**

5.1 Presence-absence matrix. DNA metabarcoding offers numerous advantages, including high taxonomic resolution and the ability to analyze bulk and complex environmental samples, it is important to acknowledge that it may not be 100% reproducible and faces limitations in quantifying taxon abundance-based composition metrics (Elbrecht and Leese, 2015; Piñol *et al.*, 2015; Creedy et al., 2019). This limitation makes it difficult to establish a positive relationship between sequencing reads and species biomass (Bista *et al.*, 2018; Lamb *et al.*, 2019). Various factors, including taxa biomass, PCR amplification bias, reaction stochasticity, primer affinity, DNA quantity, and sequencing stochasticity (Elbrecht, Peinert and Leese, 2017; Creedy et al., 2019), can affect the abundance of sequence reads, disrupting correlations with species biomass and abundance (Yu *et al.*, 2012; Elbrecht and Leese, 2015; Creedy et al., 2019). Consequently, it's recommended to focus on binary presence-absence (incidence) data instead of abundances (Koleff, Gaston and Lennon, 2003; Yu *et al.*, 2012).

To generate presence-absence data, a presence-absence matrix (PAM) is created, using binary values (1 for presence absence) to indicate species presence in each site (Arita *et al.*, 2012). Recording data in the form of haplotype-level (ASVs) community matrices improves the comparability between biodiversity surveys (Arribas *et al.*, 2020), as haplotypes are directly comparable across data sets (Callahan *et al.*, 2016). To create a haplotype-level community matrix, ASVs from different samples are filtered to remove pseudogenes likely

to be nuclear mitochondrial (numts) using a protocol based on the abundance of distributed reads (Andújar *et al.*, 2021). Then, a community table is generated, displaying the abundance of each haplotype in each sample (Arribas *et al.*, 2020). Haplotypes with less than a low number of reads are removed from each library, and those that contribute less than 1% of the total reads are removed (Arribas *et al.*, 2020). The remaining haplotypes are transformed into incidence (presence/absence) data and combined with the filtered haplotypes for further analysis (Arribas *et al.*, 2020). This approach holds great promise for biomonitoring species diversity and genetic diversity, as indicated by several studies (Arribas *et al.*, 2016, 2020; Elbrecht *et al.*, 2017, 2018; Andújar *et al.*, 2017; Gálvez-Reyes *et al.*, 2021; Noguerales *et al.*, 2021).

5. 2 Richness, alpha and beta diversity, structure, and composition. As with other estimates of biodiversity, the use of DNA sequencing has provided new challenges to estimate biodiversity (Alberdi and Gilbert, 2019). Diversity is a measure of the compositional complexity of a community (Chao and Jost, 2012) and plays a crucial role in underlaying a community's composition, structure, functionality, and evolutionary history in space (Whittaker, 1975). Diversity can be assessed using Whittaker's (1975) definitions with the decomposition of biological diversity into local (α), regional (β), and global (γ) components. Alfa diversity represents species diversity within an ecosystem, while β-diversity measures changes in diversity between two sites in terms of turnover and nestedness (Magurran, 2004; Piñero, 2005; Baselga, 2010).

Alpha diversity is typically quantified as the observed number of operational taxonomic units (OTUs) per site. Researchers also utilize subsets of ASVs to analyze community diversity and composition among sampling sites, using ANOVAs with post-hoc

Kruskal (Bonferroni method) tests to assess significant differences in alpha diversity (Arribas *et al.*, 2020; Gálvez-Reyes *et al.*, 2021; Noguerales *et al.*, 2021). Beta diversity is calculated using various Sorensen and Simpson indices with data from all sites (Koleff, Gaston and Lennon, 2003). The total β-diversity (Sorensen index, βsor), additive turnover (Simpson index, βsim; which quantifies species replacement without accounting for changes in richness), and nestedness (Sorensen–Simpson index, βsne; which measures pure richness effect) components can be estimated to assess community composition (Baselga, 2010).

For diversity estimates, raw data can be based on incidence or abundance of OTUs. Bias in the estimates depend on the degree of regularity of abundances of OTUs (Alberdi and Gilbert, 2019) which should be considered either with simulations or real data. Hill numbers, denoted as "qD," are a diverse family of biodiversity indices that provide a unified framework for quantifying diversity at various levels (Alberdi and Gilbert, 2019). The parameter "q" represents the order of diversity, while "D" stands for diversity itself. Hill numbers are particularly valuable in DNA-based diversity analyses, as they are sensitive to rare species or DNA sequences often overlooked by traditional biodiversity indices (Alberdi and Gilbert, 2019). Hill numbers are used when sequencing environmental samples like soil or water to gain insights into community structure and function. The choice of Hill number order depends on the research objectives: high-order Hill numbers (e.g., q = 2 or 3) are useful for studying the impact of rare species on ecosystem stability or microbial function, while low-order Hill numbers (e.g., q = 0 or 1) focus on traditional diversity aspects like species richness and evenness (Alberdi and Gilbert, 2019). For selecting the appropriate Hill number order based on their research question and the characteristics of the community under investigation.

ASV-based datasets provide the opportunity to test for fractal patterns in the accumulation diversity, which refers to self-similar systems (Baselga, Gómez-Rodríguez and

Vogler, 2015). To perform this analysis, a log-log Pearson correlation is calculated between haplotype level and the number of lineages, initial similarity (i.e., intercept), and mean similarity. High correlation values indicate self-similarity in lineage branching (i.e., number of lineages) and/or spatial geometry of lineage distributional ranges (i.e., initial and mean similarity), thus supporting the presence of fractal patterns, suggesting neutral community evolution processes (Arribas *et al.*, 2020). Detecting these patterns offers valuable insights into the mechanisms shaping community structure and diversity (Baselga, Gómez-Rodríguez, and Vogler, 2015).

5.3 Network analyses, phylogenetic diversity and phylogeography. Network analysis is a valuable tool for studying ecological communities, offering insights into species interactions and ecological processes. Different types of network analyses are available, each with its own set of data requirements and analytical methods (Beckett, 2016). One type of network analysis is the *bipartite network,* to understand the degree of specialization of arthropods assessing the degree of interactions between two distinct node classes (Dormann *et al.*, 2009; Beckett, 2016). Bipartite networks are commonly used to analyze mutualistic and antagonistic interactions (Clare *et al.*, 2019; Encinas-Viso *et al.*, 2022) and can assess specialization in arthropods using regression analysis based on presence/absence data (Bittleston *et al.*, 2016).

Another type of network analysis is the *co-occurrence network*, which provide insights into community structure, highlighting potential interaction networks, and shared niches among community members (Ma *et al.*, 2016). Co-occurrence patterns help understanding the impact of anthropogenic activities on taxa distribution (Lanzén *et al.*, 2016). Also, co-occurrence networks can aid in identifying potential interaction networks

and revealing niche spaces shared by community members and in exploring distribution reveal trophic relationships in the benthic food webs, i.e. *trophic networks* (Lanzén *et al.*, 2016), aiding in the study of energy and nutrients flow and the identification of key species. For example, a trophic network analysis uncovered novel hypotheses about trophic interactions within a community (Lanzén *et al.*, 2016). Additionally, stability analyses of site indicator ASVs have been conducted, suggesting the potential of terrestrial arthropod fauna as ecological integrity indicators in forest systems (Porter *et al.*, 2019; Arribas *et al.*, 2020). Different network analyses address various ecological questions, with bipartite networks informing about specialization, co-occurrence networks revealing interaction networks and shared niches and stability analyses aiding in ecological integrity assessment (Dormann et *al.*, 2009; Beckett, 2016; Clare *et al.*, 2019; Encinas-Viso *et al.*, 2022; Ma *et al.*, 2016; Bittleston *et al.*, 2016; Lanzén *et al.*, 2016; Porter *et al.*, 2019; Arribas *et al.*, 2020).

Phylogenetic diversity measures evolutionary diversity in communities (Faith and Baker, 2007; Gerhold *et al.*, 2015). According to Gerhold *et al.* (2015), while the phylogeny of a community influences its assembly, phylogenetic diversity is related to, but not equivalent to, species richness (Faith and Baker, 2007; *et al.*, 2014). High phylogenetic diversity can exist in areas with few species (Kress *et al.*, 2014), but a disparity arises when closely related species are concentrated. Tucker *et al.* (2017) quantified phylogenetic diversity using richness (accumulated phylogenetic difference), divergence (average differences between taxa), and regularity (variation between taxa). Despite concerns raised by Gerhold *et al.* (2015) about using phylogenetic patterns as proxies for community assembly, such patterns are employed to test hypotheses regarding local coexistence's role

in macroevolution within habitat reserve lineage through competition among close relatives, leading to trait displacement and diversification.

Also, an integrated *phylogeography* approach involving multiple species a valuable tool for investigating landscape-level processes, including natural and anthropogenic factors (Elbrecht *et al.*, 2018; Turon *et al.*, 2020). Phylogenetic networks, as demonstrated by De Luca et al. (2021), can facilitate this approach. DNA metabarcoding enables phylogeographic and phylogenetic analyses across various taxa to assess the effectiveness of conservation strategies for arthropod biodiversity. Population demographic changes can be studied using COI sequences mismatch distribution and Tajima's D statistic (Shum and Palumbi, 2021), while DNA metabarcoding can simultaneously reveal predators diets and population structures (Bohmann *et al.*, 2018). Additionally, *landscape community* connectivity, considering geographic distance, landscape features, and environmental filtering, has been explored (Gálvez-Reyes *et al.*, 2021; Noguerales *et al.*, 2021). A recent study introduced a novel analytical framework, metaphylogeography, which employs DNA metabarcoding to detect biogeographic and phylogeographic structures of multiple species from haplotype data (Antich *et al.*, 2023).

References

Alberdi, A. and Gilbert, M.T.P. (2019) 'A guide to the application of Hill numbers to DNA-based diversity analyses', *Molecular Ecology Resources*, *19*(4), 804-817. https://doi.org/10.1111/1755-0998.13014

Andújar, C., Arribas, P., Gray, C., *et al.* (2017) 'Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill', *Molecular Ecology*, 27(1), pp. 146–166. https://doi.org/10.1111/mec.14410.

Andújar, C. *et al.* (2021) 'Validated removal of nuclear pseudogenes and sequencing artefacts from mitochondrial metabarcode data', *Molecular Ecology Resources*, 21(6), pp. 1772–1787. https://doi.org/10.1111/1755-0998.13337.

Antich, A. *et al.* (2023) 'Metabarcoding reveals high-resolution biogeographical and metaphylogeographical patterns through marine barriers', *Journal of Biogeography*, 50(3), pp. 515–527. https://doi.org/10.1111/jbi.14548.

Arita, H.T. *et al.* (2012) 'The presence-absence matrix reloaded: The use and interpretation of range-diversity plots', *Global Ecology and Biogeography*, 21(2), pp. 282–292. https://doi.org/10.1111/j.1466-8238.2011.00662.x.

Arribas, P. *et al.* (2016) 'Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil', *Methods in Ecology and Evolution*, 7(9), pp. 1071–1081. https://doi.org/10.1111/2041-210X.12557.

Arribas, P. *et al.* (2020) 'The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding', *Molecular Ecology*, 30(1), pp. 48–61. https://doi.org/10.1111/mec.15591.

Baselga, A. (2010) 'Partitioning the turnover and nestedness components of beta diversity', *Global Ecology and Biogeography*, 19(1), pp. 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x.

Baselga, A., Gómez-Rodríguez, C. and Vogler, A.P. (2015) 'Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes', *Global Ecology and Biogeography*, 24(8), pp. 873–882. https://doi.org/10.1111/geb.12322.

Beckett, S.J. (2016) 'Improved community detection in weighted bipartite networks', *Royal Society Open Science*, 3(1), p. 140536. https://doi.org/10.1098/rsos.140536.

Bista, I. *et al.* (2018) 'Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples', *Molecular Ecology Resources*, 18(5), pp. 1020–1034. https://doi.org/10.1111/1755-0998.12888.

Bittleston, L.S. *et al.* (2016) 'Metabarcoding as a tool for investigating arthropod diversity in Nepenthes pitcher plants', *Austral Ecology*, 41(2), pp. 120–132. https://doi.org/10.1111/aec.12271.

Bohmann, K. *et al.* (2018) 'Using DNA metabarcoding for simultaneous inference of common vampire bat diet and population structure', *Molecular Ecology Resources*, 18(5), pp. 1050–1063. https://doi.org/10.1111/1755-0998.12891.

Callahan, B.J. *et al.* (2016) 'DADA2: High-resolution sample inference from Illumina amplicon data', *Nature Methods*, 13(7), pp. 581–583. https://doi.org/10.1038/nmeth.3869.

Chao, A. and Jost, L. (2012) ‘Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size’, *Ecology*, 93(12), pp. 2533–2547. https://doi.org/10.1890/11-1952.1.

Creedy, T.J., Ng, W.S. and Vogler, A.P. (2019) ‘Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy’, *Ecology and Evolution*, 9(6), pp. 3105–3116. https://doi.org/10.1002/ece3.4839.

Dormann, C.F. *et al.* (2009) ‘Indices, Graphs and Null Models: Analyzing Bipartite Ecological Networks’, *The Open Ecology Journal*, 2(1). https://benthamopen.com/ABSTRACT/TOECOLJ-2-1-7 (Accessed: 15 May 2023).

Elbrecht, V. *et al.* (2017) ‘Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring’, *Methods in Ecology and Evolution*, 8(10), pp. 1265–1275. https://doi.org/10.1111/2041-210X.12789.

Elbrecht, V. *et al.* (2018) ‘Estimating intraspecific genetic diversity from community DNA metabarcoding data’, *PeerJ*, 6, p. e4644. https://doi.org/10.7717/peerj.4644.

Elbrecht, V. and Leese, F. (2015a) ‘Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass - sequence relationships with an innovative metabarcoding protocol’, *Peer J* [Preprint].

Elbrecht, V., Peinert, B. and Leese, F. (2017) ‘Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding’, *Ecology and Evolution*, 7(17), pp. 6918–6926. https://doi.org/10.1002/ece3.3192.

Encinas-Viso, F. *et al.* (2022) ‘Pollen DNA metabarcoding reveals cryptic diversity and high spatial turnover in alpine plant–pollinator networks’, *Molecular Ecology*, n/a(n/a). https://doi.org/10.1111/mec.16682.

Faith, D.P. and Baker, A.M. (2007) ‘Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges’, *Evolutionary Bioinformatics Online*, 2, pp. 121–128. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2674678/ (Accessed: 15 May 2023).

Gálvez-Reyes, N. *et al.* (2021) ‘Dispersal limitations and long-term persistence drive differentiation from haplotypes to communities within a tropical sky-island: Evidence from community metabarcoding’, *Molecular Ecology*, 30(24), pp. 6611–6626. https://doi.org/10.1111/mec.16195.

Gerhold, P. *et al.* (2015) ‘Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better)’, *Functional Ecology*, 29(5), pp. 600–614. https://doi.org/10.1111/1365-2435.12425.

Koleff, P., Gaston, K.J. and Lennon, J.J. (2003) 'Measuring beta diversity for presence-absence data', *Journal of Animal Ecology*, 72(3), pp. 367–382. https://doi.org/10.1046/j.1365-2656.2003.00710.x.

Kress, W.J. *et al.* (2014) 'DNA barcodes for ecology, evolution, and conservation', *Trends in Ecology & Evolution*, 30(1), pp. 25–35. https://doi.org/10.1016/j.tree.2014.10.008.

Lamb, P.D. *et al.* (2019) 'How quantitative is metabarcoding: A meta-analytical approach', *Molecular Ecology*, 28(2), pp. 420–430. https://doi.org/10.1111/mec.14920.

Lanzén, A. *et al.* (2016) 'High-throughput metabarcoding of eukaryotic diversity for environmental monitoring of offshore oil-drilling activities', *Molecular Ecology*, 25(17), pp. 4392–4406. https://doi.org/10.1111/mec.13761.

Ma, B. *et al.* (2016) 'Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China', *The ISME Journal*, 10(8), pp. 1891–1901. https://doi.org/10.1038/ismej.2015.261.

Magurran, A.E. (2004) *Measuring Biological Diversity*. Wiley.

Noguerales, V. *et al.* (2021) 'Community metabarcoding reveals the relative role of environmental filtering and spatial processes in metacommunity dynamics of soil microarthropods across a mosaic of montane forests', *Molecular Ecology*, n/a(n/a), pp. 1–19. https://doi.org/10.1111/mec.16275.

Piñero, D. (2005) 'Similitudes y diferencias entre los conceptos y los patrones de diversidad beta y diferenciación genética: aplicaciones en bosques mexicanos de coníferas', in *Sobre diversidad biológica: el significado de las diversidades alfa, beta y gamma, 2005, ISBN 84-932807-7-1, págs. 53-62*. https://dialnet.unirioja.es/servlet/articulo?codigo=1395357 (Accessed: 15 May 2023).

Piñol, J. *et al.* (2015) 'Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods', *Molecular Ecology Resources*, 15(4), pp. 819–830. https://doi.org/10.1111/1755-0998.12355.

Porter, T.M. *et al.* (2019) 'Variations in terrestrial arthropod DNA metabarcoding methods recovers robust beta diversity but variable richness and site indicators', *Scientific Reports*, 9(1), p. 18218. https://doi.org/10.1038/s41598-019-54532-0.

Shum, P. and Palumbi, S.R. (2021) 'Testing small-scale ecological gradients and intraspecific differentiation for hundreds of kelp forest species using haplotypes from metabarcoding', *Molecular Ecology*, 30(13), pp. 3355–3373. https://doi.org/10.1111/mec.15851.

Tucker, C.M. *et al.* (2017) 'A guide to phylogenetic metrics for conservation, community ecology and macroecology', *Biological Reviews*, 92(2), pp. 698–715. https://doi.org/10.1111/brv.12252.

Turon, X. *et al.* (2020) 'From metabarcoding to metaphylogeography: separating the wheat from the chaff', *Ecological Applications*, 30(2), p. e02036. https://doi.org/10.1002/eap.2036.

Uscanga, A. *et al.* (2021) 'Evaluating species origins within tropical sky-islands arthropod communities', *Journal of Biogeography*, 48(9), pp. 2199–2210. https://doi.org/10.1111/jbi.14144.
Uscanga A (2016) Filogeografía comparada de artrópodos de alta montaña de la Faja Volcánica Transmexicana [UNAM]. Retrieved from http://132.248.9.195/ptd2016/junio/0745798/I

RStudio Team (2020) RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Whittaker, R.H. (1975) 'Communities and ecosystems, 2nd edn Macmillan Publishing Company', *New York, NY* [Preprint].

Oksanen, J, Blanchet, FG, Friendly, M, Kindt, R, Legendre, P, McGlinn, D, Minchin, PR, O'Hara, R B, Simpson, GL, Solymos, P, Stevens, MHH, Szoecs, E, Wagner, H (2019) vegan: Community Ecology Package (2.5-6) [Computer software]. https://CRAN.R-project.org/package=vegan

Wickham, H, Chang, W, Henry, L, Pedersen, TL, Takahashi, K, Wilke, C, Woo, K, Yutani, H, Dunnington, D, RStudio (2020) ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics (3.3.2) [Computer software].

Yu, D.W. *et al.* (2012) 'Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring', *Methods in Ecology and Evolution*, 3(4), pp. 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x.

ANEXO 2: Material suplementario artículo 2

MOLECULAR ECOLOGY

Gálvez-Reyes et. al. Supporting Information.

Supplemental Information for:

Dispersal limitations and long-term persistence drive differentiation from haplotypes to communities within a tropical sky-island: evidence from community metabarcoding

Nancy Gálvez-Reyes[1,2], Paula Arribas[3], Carmelo Andújar[3], Brent C. Emerson[3], Daniel Piñero[1] and Alicia Mastretta-Yanes[4,5].

Table of Contents:

MOLECULAR ECOLOGY
Gálvez-Reyes et. al. Supporting Information.

1. Information on PCR reagents and conditions

Supporting Material S1. Metabarcoding PCR reagents and conditions.

PCR on bulk-sample DNA extracts were performed in 25 µL reaction volumes containing 2 mM MgCl2, 0.1 mM dNTPs (Thermo Scientific), 0.3 mM each primer salt-free, 0.02 mg BSA and 1 U Taq DNA polymerase (Invitrogen™ Platinum™ *Taq* DNA Polymerase High Fidelity) and 5 ng/µL of DNA bulk extraction. We used a cycling profile of 95ºC for 4 min; 28 cycles of 95ºC for 30 s; 48ºC for 30 s; 72ºC for 3 min, and a final extension of 72ºC for 10 min. PCR protocol and conditions were the same for previous study on metabarcoding of arthropods (Arribas et al., 2016). Six PCR replicates were conducted on each bulk-sample and PCR products were pooled before library construction using Illumina XT indexes. For index PCR, we perform limited-cycle PCR on a thermal cycler using the following program: 95°C for 3 minutes, 8 cycles of: 95°C for 30 s, 55°C for 30 s, 72°C for 30 s, 72°C for 5 minutes and hold at 4°C.

2. Sampling and Flotation-Filtration-Stereoscope

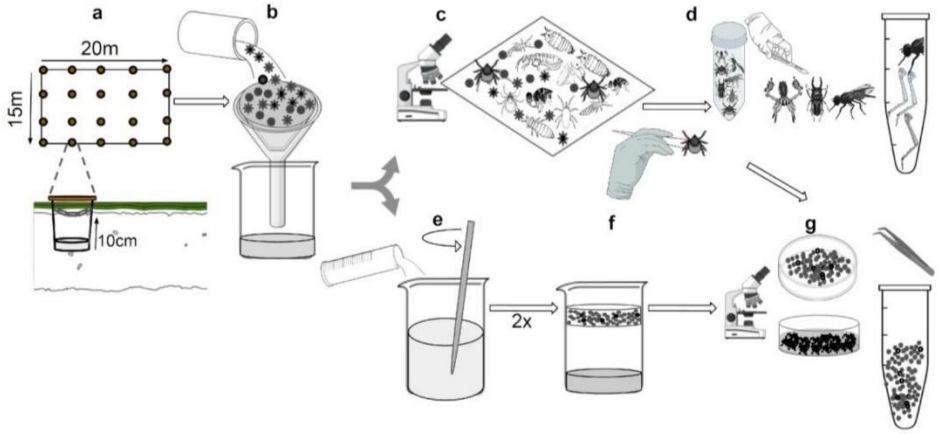

Figure S1. Sampling and Flotation-Filtration-Stereoscope (FFS) protocol for cleaning arthropod bulk-samples. 51 bulk-samples were collected from the TMVB: 42 in Nevado de Toluca and nine between three Natural Protected Areas above *Abies religiosa* forest. (a) We distributed 20 pitfall-traps for each block site (15 x 20 m) with 200 ml (alcohol 70% + glycerin 15 ml). After collecting pitfall traps, bulk-samples were processed following an FFS protocol that allows for the 'clean' extraction of arthropods. The FFS protocol is based on arthropods flotation in water, filtration process and observation in the stereoscope. (b) The Flotation and filtration steps allow the small arthropods to remain on the filter paper, extraction of the large specimens (keep in falcon tube) and the extraction of the organic matter (roots, leaves and soil). (c) After, we observed in a stereoscope and captured small arthropods (<=3 mm) with a small clamp. (d) Samples were segmented according to size, for large-arthropods (>15 mm) only a leg was kept, and the head and thorax for medium-sized-arthropods (>3 mm and <=15 mm). (e) Subsequently, additional water for the last flotation of the sediment portion is placed using 400 ml of water, mixing for 15 seconds and allowing it to stand for 3 minutes. This step was made for two times, depending on the sample's level of dirty. (g) The mixture was decanted on petri dishes and observed in a stereoscope for capturing small specimens. The last step of the FFS protocol was adding ethanol on arthropods for preservation, resulting in 'clean' bulk-sample of specimens ready for DNA extraction.

3. Metabarcoding libraries

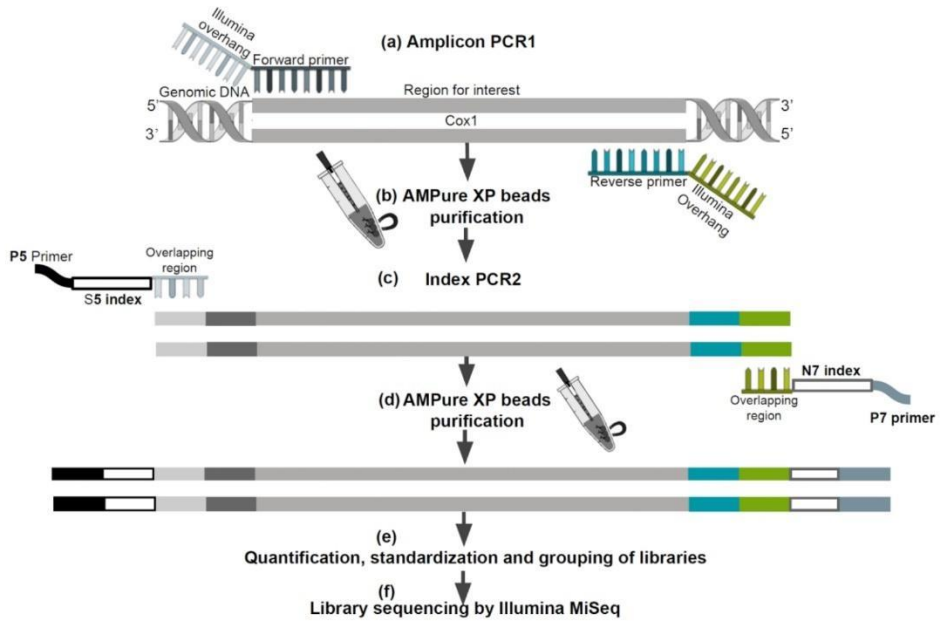

Figure S2. Preparation of amplicon libraries for Metabarcoding of arthropods bulk-samples using Illumina MiSeq with Nextera XT adapters. (a) Amplification by PCR of the amplicon (PCR1): three times for each bulk-sample, using primers from the cox1 region with the overhang adapters. (b) Pooling of the PCR repeats and PCR purification (PCR clean-up) using Ampure XP beads. (c) The indexing of each library by PCR (PCR2) with the primers Nextera XT kit. (d) A second cleaning with magnetic beads from the pool with Ampure XP. (e) Quantify, normalize and pool the libraries. (f) Sequencing on the illumina MiSeq platform 2×300 bp.

4. Supplementary information for analyses

Supporting Material S2. Corrections of the *Abies* forest distribution map. The digital land cover map of Nevado of Toluca used here identifies a high resolution, based on satellite images (SPOT 6/7) of 2015 defining the *Abies religiosa* forest (sampled vegetation), *Pinus hartwegii* forest, alpine grassland, agriculture, urbanization, water in the crater and Nevado de Toluca's crater (González-Fernández et al., 2019). The area surrounding our West sampling points is an area of high topographic complexity where some of the *Abies* forests we sampled had been erroneously classified as *Pinus* forests. After consulting with the authors of the original map, we manually corrected the cells where our sampling coordinates indicated *Abies* forest instead of *Pinus*. Corrections were done using *ArcMap 10.3.1*.

Table S1. Number of haplotypes, lineages at 3% and 5%, putative species by GMYC and GMYC threshold values for eight arthropods orders.

Order	Haplotypes	3% lineages	5% lineages	GMYC	GMYC threshold
Diptera	385	106	95	124	0.0088
Collembola	270	59	52	60	0.029
Arachnida	155	67	61	78	0.01258205
Coleoptera	136	87	81	102	0.007255866
Hymenoptera	133	83	71	94	0.009765845
Hemiptera	116	55	50	64	0.01771803
Myriapoda	51	15	11	20	0.01529633
Lepidoptera	31	24	20	28	0.003
Total	**1,277**	**496**	**441**	**570**	

Table S2. Strategy for testing the effects of increasingly complex models of effective distance on arthropods communities structure. Conductance matrix for Circuitscape corresponding to various tests of landscape resistance of altitude and slope.

Raster	**How conductance values were assigned**									
Vegetation type	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
	We assumed no resistance to dispersal (maximal conductance) in A. religiosa forests, very small limited dispersal in agricultural or urbanized lands and no dispersal at the volcano crater. What we varied were the conductances of *P. hartwegii* and alpine grasslands, which are the immediate natural habitats surrounding A. religiosa forests. The conductances were chosen to explore scenarios where these vegetations allow or impede dispersal, as well as its combinations.									
Abies religiosa	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Pinus hartwegii	1.0	1.0	1.0	0.9	0.7	0.5	0.2	0.2	0.2	1.0
Grassland	0.7	0.5	0.2	0.2	0.2	0.2	0.2	0.7	0.2	0.7
Agriculture	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.5
Urbanization	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Water crater	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Nevado de Toluca's crater	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Altitude										
Altitude 2000 m.a.s.l.	Each altitude (2000, 2700, 2800, 2900, etc.) had a conductance value of 0.2 below that altitude, and above or equal to that altitude the conductance was assigned 1. For example: Conductance matrix: 0, 2000, **0.2** 2000, 4457, **1**									
Altitude 2700 m.a.s.l.										
Altitude 2800 m.a.s.l.										
Altitude 2900 m.a.s.l.										
Altitude 3000 m.a.s.l.										
Altitude 3100 m.a.s.l.										
Altitude 3200 m.a.s.l.										
Altitude 3300 m.a.s.l.										
Altitude 3400 m.a.s.l.										
Altitude 3500 m.a.s.l.										
Altitude 3600 m.a.s.l.										
Altitude 3700 m.a.s.l.										
"Mean-elevation peak (truncated)" (Altitude gradient A)	We gave maximum levels of conductance (1) at 25 m above the average (red) of our sampling and 25 m below the average of our sampling, and then decreased it (0.3) until reaching the minimal (green) and maximum altitudes (blue) of our sampling. Altitudes outside of our sampling area but where *Abies* can still be found were assumed to allow conductance to a lesser degree (0.1). Altitudes outside our sampling and *Abies* range were allowed no conductance (0). Conductance matrix: 4000, 4457, **0** 3386, 4000, **0.1**									

	3189, 3386, **0.3** 3164, 3189, **1** 3139, **3164**, **1** 3010, 3139, **0.3** 2600, **3010**, **0.2** 2000, 2600, **0.1** 0, 2000, **0**
"Mean-elevation peak (symmetrical)" (Altitude gradient B)	We gave maximum levels of conductance (1) at 25 m above the average (red) of our sampling and 25 m below the average of our sampling, and then decreased it (0.3) until reaching the minimal (green) and maximum altitudes (blue) of our sampling. Altitudes outside of our sampling area but where *Abies* can still be found were assumed to allow conductance to a lesser degree (0.2). Altitudes outside our sampling and *Abies* range were allowed no conductance (0). Conductance matrix: 4000, 4457, **0** 3386, 4000, **0.2** 3189, **3386**, **0.3** **3164**, 3189, **1** 3139, **3164**, **1** **3010**, 3139, **0.3** 2000, 3010, **0.2** 0, 2000, **0**
"Mean-elevation peak (narrow)" (Altitude gradient C)	We gave maximum levels of conductance (1) from the minimal (3,010) to the maximum altitudes (3,386) of our sampling area but where Abies can still be found were assumed to allow conductance to a lesser degree (0.2). Altitudes outside our sampling and Abies range were allowed no conductance (0). Conductance matrix: 4000, 4457, **0** 3386, 4000, **0.2** **3010**, **3386**, **1** 2000, 3010, **0.2** 0, 2000, **0**
Slope	
Low slopes facilitate dispersal and middle slopes partially impede it (Slope A)	Slope conductance matrix A 0, 6.57, **1** 6.57, 13.14, **0.5** 13.14, 19.6999, **0.2**
Low slopes facilitate dispersal and middle slopes partially allow it (Slope B)	Slope conductance matrix B 0, 6.57, **1** 6.57, 13.14, **0.7** 13.14, 19.6999, **0.2**
Only low slopes (<=6.57) facilitate dispersal (Slope C)	Slope conductance matrix C 0, 6.57, **1** 6.57, 19.6999, **0.2**
Only low slopes (<=9.85) facilitate dispersal (Slope D)	Slope conductance matrix D 0, 9.8499, **1** 9.8499, 19.6999, **0.2**

Table S3. Results of Isolation by resistance in Collembola with hierarchical level and species delimitation using GMYC. Associations with linear regressions between the pairwise effective distances and the pairwise beta diversity. The r values (r^2) are reported in each hierarchical level. All tests were significative ($p < 0.0001$). Underlined cells correspond to the surface with the highest prediction value.

	Hierarchical level															
Collembola community	**haplotype**		**lineages 0.5 %**		**lineages 1.5 %**		**Lineages 2%**		**Lineages 3 %**		**lineages 5 %**		**lineages 7.5%**		**GMYC**	
Surface	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)
A	0.744	-1.668	0.704	-1.388	0.639	-0.965	0.655	-0.925	0.615	-0.854	0.600	-0.746	0.594	-0.691	0.615	-0.857
B	0.747	-1.594	0.706	-1.322	0.637	-0.914	0.654	-0.876	0.615	-0.808	0.600	-0.706	0.592	-0.653	0.615	-0.811
C	0.753	-1.425	0.709	-1.172	0.631	-0.798	0.650	-0.765	0.611	-0.706	0.596	-0.614	0.587	-0.567	0.612	-0.708
D	0.758	-1.385	0.712	-1.134	0.630	-0.767	0.650	-0.736	0.613	-0.679	0.597	-0.590	0.588	-0.545	0.613	-0.681
E	**0.762**	-1.280	**0.716**	-1.042	0.625	-0.693	0.647	-0.667	0.614	-0.617	0.596	-0.534	0.586	-0.492	0.614	-0.619
F	0.756	-1.133	0.711	-0.917	0.612	-0.601	0.636	-0.579	0.608	-0.538	0.588	-0.464	0.577	-0.427	0.608	-0.539
G	0.699	-0.775	0.662	-0.625	0.556	-0.400	0.584	-0.388	0.567	-0.363	0.541	-0.310	0.530	-0.285	0.567	-0.364
H	0.682	-0.920	0.646	-0.747	0.546	-0.483	0.571	-0.467	0.556	-0.438	0.528	-0.374	0.518	-0.344	0.556	-0.439
I	0.698	-0.824	0.661	-0.668	0.560	-0.432	0.586	-0.417	0.569	-0.391	0.542	-0.335	0.532	-0.307	0.569	-0.392
J	0.730	-1.844	0.693	-1.548	0.646	-1.103	0.656	-1.055	0.615	-0.975	0.599	-0.855	0.595	-0.795	0.615	-0.978
2000	0.702	-2.057	0.668	-1.745	0.637	-1.270	0.642	-1.212	0.599	-1.119	0.583	-0.986	0.581	-0.919	0.599	-1.123
2700	0.715	-2.011	0.685	-1.705	0.654	-1.237	0.660	-1.182	0.619	-1.092	0.603	-0.961	0.599	-0.895	0.619	-1.095
2800	0.726	-1.968	0.699	-1.667	0.667	-1.206	0.675	-1.152	0.635	-1.066	0.618	-0.937	0.614	-0.872	0.635	-1.069
2900	0.710	-1.787	0.681	-1.506	0.661	-1.093	0.659	-1.039	0.619	-0.960	0.599	-0.841	0.596	-0.783	0.619	-0.963
3000	0.723	-1.633	0.698	-1.370	**0.684**	**-0.990**	**0.679**	**-0.939**	**0.644**	**-0.869**	**0.622**	**-0.759**	**0.615**	**-0.705**	**0.644**	**-0.872**
3100	0.051	-0.189	0.068	-0.183	0.117	-0.155	0.088	-0.131	0.079	-0.118	0.065	-0.093	0.064	-0.084	0.079	-0.119
3200	0.078	-0.206	0.099	-0.197	0.153	-0.162	0.122	-0.140	0.112	-0.128	0.094	0.094	0.093	-0.094	0.112	-0.129
3300	0.285	-0.326	0.304	-0.294	0.368	-0.232	0.338	-0.214	0.322	-0.199	0.311	-0.173	0.299	-0.158	0.321	-0.199
3400	0.530	-0.413	0.517	-0.357	0.513	-0.264	0.499	-0.248	0.456	-0.227	0.445	-0.199	0.438	-0.185	0.454	-0.227
3500	0.684	-0.429	0.651	-0.365	0.629	-0.268	0.629	-0.256	0.585	-0.236	0.571	-0.208	0.569	-0.194	0.585	-0.236
3600	0.690	-0.423	0.657	-0.360	0.632	-0.263	0.634	-0.251	0.590	-0.232	0.575	-0.205	0.573	-0.191	0.590	-0.232
3700	0.697	-0.417	0.663	-0.354	0.635	-0.258	0.638	-0.246	0.594	-0.228	0.579	-0.201	0.577	-0.187	0.594	-0.228

"Mean-elevation peak (truncated)"	0.618	-0.120	0.589	-0.093	0.564	-0.076	0.577	-0.076	0.544	-0.067	0.513	-0.066	0.524	-0.037	0.546	-0.084
"Mean-elevation peak (symmetrical)"	0.610	-0.514	0.585	-0.435	0.580	-0.320	0.581	-0.306	0.540	-0.282	0.512	-0.246	0.522	-0.230	0.540	-0.283
"Mean-elevation peak (narrow)"	0.194	-0.514	0.184	-0.379	0.192	-0.247	0.189	-0.230	0.169	-0.205	0.135	-0.158	0.155	-0.155	0.171	-0.206
Slope A**	0.380	-0.966	0.398	-0.860	0.454	-0.672	0.431	-0.630	0.403	-0.587	0.383	-0.511	0.383	-0.476	0.404	-0.588
Slope B**	0.562	-1.488	0.561	-1.297	0.589	-0.987	0.576	-0.935	0.538	-0.868	0.521	-0.763	0.520	-0.713	0.539	-0.870
Slope C**	0.085	-0.230	0.103	-0.215	0.155	-0.177	0.129	-0.156	0.118	-0.144	0.104	-0.118	0.104	-0.109	0.119	-0.145
Slope D**	0.131	-0.533	0.125	-0.380	0.149	-0.252	0.117	-0.208	0.129	-0.206	0.135	-0.179	0.109	-0.148	0.129	-0.206
Flat*	0.704	-2.075	0.669	-1.755	0.637	-1.275	0.642	-1.217	0.598	-1.124	0.583	-0.990	0.580	-0.922	0.599	-1.127
Euclidean*	0.799	-0.0002 e-05	0.814	-0.0001 e-05	0.722	-6.649 e-05	0.760	-6.383 e-05	0.750	-5.933 e-05	0.711	-4.954 e-05	0.700	-4.502 e-05	0.749	-5.954 e-05

Table S4. Results of Isolation by resistance in Diptera at different hierarchical levels and species delimitation using GMYC. Associations with linear regressions between the pairwise effective distances and the pairwise beta diversity. The r values (r^2) are reported in each hierarchical level. All tests were significative ($p < 0.0001$). Underlined cells correspond to the surface with the highest prediction value.

	Hierarchical level															
Diptera community	**haplotype**		**lineages 0.5 %**		**lineages 1.5 %**		**lineages 2%**		**lineages 3 %**		**lineages 5 %**		**lineages 7.5%**		**GMYC**	
Surface	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)
A	0.301	-0.412	0.236	-0.288	0.173	-0.245	0.170	-0.243	0.146	-0.215	0.150	-0.215	0.023	-0.080	0.204	-0.276
B	0.300	-0.388	0.236	-0.272	0.172	-0.231	0.170	-0.228	0.146	-0.203	0.150	-0.202	0.022	-0.074	0.203	-0.260
C	0.295	-0.334	0.235	-0.235	0.171	-0.199	0.169	-0.197	0.145	-0.175	0.149	-0.174	0.020	-0.061	0.202	-0.224
D	0.296	-0.321	0.234	-0.224	0.170	-0.190	0.168	-0.188	0.144	-0.167	0.148	-0.166	0.018	0.726	0.201	-0.214
E	0.297	-0.289	0.232	-0.201	0.167	-0.169	0.165	-0.167	0.143	-0.149	0.144	-0.147	0.016	-0.046	0.198	-0.191
F	0.294	-0.250	0.227	-0.173	0.162	-0.145	0.159	-0.143	0.139	-0.128	0.137	-0.125	0.012	-0.036	0.192	-0.163
G	0.273	-0.166	0.203	-0.113	0.140	-0.093	0.138	-0.091	0.125	-0.083	0.115	-0.079	0.007	-0.018	0.168	-0.105
H	0.271	-0.202	0.196	-0.135	0.133	-0.110	0.130	-0.108	0.118	-0.099	0.105	-0.092	0.006	-0.021	0.159	-0.125

I	0.275	-0.180	0.203	-0.121	0.139	-0.099	0.136	-0.098	0.123	-0.089	0.112	-0.084	0.007	-0.020	0.166	-0.113
J	0.301	-0.478	0.236	-0.336	0.171	-0.286	0.170	-0.283	0.145	-0.251	0.149	-0.250	0.029	-0.107	0.202	-0.321
2000	0.293	-0.555	0.229	-0.393	0.167	-0.335	0.166	-0.333	0.142	-0.296	0.146	-0.295	0.036	-0.141	0.196	-0.375
2700	0.307	-0.543	0.240	-0.383	0.174	-0.326	0.173	-0.324	0.148	-0.288	0.153	-0.287	0.039	-0.138	0.204	-0.365
2800	**0.319**	-0.531	0.249	-0.374	0.179	-0.317	0.179	-0.315	0.153	-0.280	0.157	-0.279	0.041	-0.136	0.211	-0.356
2900	0.312	-0.479	0.247	-0.340	0.175	-0.286	0.173	-0.283	0.149	-0.252	0.152	-0.250	0.047	-0.133	0.204	-0.319
3000	0.339	-0.437	0.268	-0.309	0.184	-0.257	0.182	-0.254	0.158	-0.226	0.158	-0.223	0.054	-0.124	0.215	-0.287
3100	0.048	-0.058	0.039	-0.040	0.029	-0.034	0.028	-0.034	0.022	-0.028	0.017	-0.024	0.094	-0.052	0.029	-0.036
3200	0.064	-0.063	0.049	-0.043	0.039	0.039	0.037	-0.036	0.030	-0.031	0.024	-0.027	0.101	-0.051	0.040	-0.039
3300	0.134	-0.091	0.068	-0.052	0.069	-0.052	0.060	-0.048	0.046	-0.040	0.033	-0.034	0.043	-0.036	0.076	-0.056
3400	0.216	-0.111	0.145	-0.073	0.127	-0.068	0.122	-0.066	0.107	-0.059	0.106	-0.058	0.034	-0.031	0.147	-0.075
3500	0.288	-0.118	0.222	-0.083	0.162	-0.071	0.160	-0.071	0.135	-0.062	0.137	-0.062	0.036	-0.030	0.189	-0.080
3600	0.290	-0.116	0.224	-0.082	0.163	-0.070	0.162	-0.069	0.137	-0.061	0.139	-0.061	0.036	-0.030	0.191	-0.078
3700	0.292	-0.113	0.227	-0.080	0.165	-0.068	0.164	-0.068	0.139	-0.060	0.143	-0.050	0.036	-0.029	0.193	-0.077
"Mean-elevation peak (truncated)	0.276	-0.463	0.275	-0.389	0.184	-0.279	0.182	-0.267	0.150	-0.247	0.155	-0.213	0.055	-0.199	0.206	-0.248
"Mean-elevation peak (symmetrical)"	0.286	-0.145	**0.276**	-0.112	**0.207**	-0.097	**0.202**	-0.095	**0.171**	-0.084	**0.176**	-0.084	0.070	-0.051	**0.228**	-0.106
"Mean-elevation peak (narrow)"	0.092	-0.094	0.156	-0.094	0.135	-0.086	0.126	-0.083	0.108	-0.073	0.108	-0.072	0.095	-0.063	0.128	-0.088
Slope A**	0.225	-0.310	0.187	-0.227	0.124	-0.186	0.125	-0.185	0.099	-0.159	0.095	-0.153	0.092	-0.141	0.140	-0.204
Slope B**	0.287	-0.453	0.233	-0.328	0.162	-0.275	0.162	-0.274	0.133	-0.239	0.132	-0.234	0.071	-0.163	0.186	-0.304
Slope C**	0.073	-0.073	0.065	-0.054	0.037	-0.040	0.038	-0.041	0.026	-0.032	0.022	-0.029	0.089	-0.053	0.038	-0.043
Slope D**	0.135	-0.126	0.099	-0.083	0.022	-0.039	0.023	-0.039	0.020	-0.035	0.015	-0.030	0.024	-0.035	0.028	-0.045
Flat*	0.293	-0.558	0.228	-0.394	0.166	-0.336	0.166	-0.333	0.141	-0.296	0.146	-0.295	0.036	-0.141	0.195	-0.376
Euclidean*	0.363	-2.558 e-05	0.288	-1.747 e-05	0.207	-1.454 e-05	0.208	-1.446 e-05	0.170	-1.318 e-05	0.198	-1.324 e-05	0.054	-6.420 e-05	0.247	-1.657 e-05

*To compare the effect of landscape variables against the effect of distance alone, we generated a 'flat' raster landscape which accounts for the finite size of the input landscape being analyzed, so therefore is more appropriate for comparison with the models using other grids (Lee-Yaw et al., 2009). Tests were also performed with Euclidean distances for reference and comparison with other IBD analyses, but these are not comparable with IBR results. **Names of the resistance slope surfaces are in Table S2.

MOLECULAR ECOLOGY
Gálvez-Reyes et. al. Supporting Information.

Table S5. Values of IBR versus IBD (flat) on Collembola (Coll) and Diptera (Dip) community in East and West sampling areas of Nevado de Toluca. Significant values ($p < 0.01$) are highlighted in bold.

			Hierarchical level															
			haplotypes		**lineages 0.5 %**		**lineages 1.5 %**		**lineages 2%**		**lineages 3 %**		**lineages 5 %**		**lineages 7.5%**		**GMYC**	
		Surface	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)	Pseudo-r^2	slope (b)
East	**Coll**	Altitude 3000	**0.489**	-1.185	**0.410**	-0.751	**0.279**	-0.456	**0.286**	-0.442	**0.149**	-0.288	**0.173**	-0.291	**0.140**	-0.255	**0.152**	-0.292
		Flat	**0.485**	-1.22	**0.407**	-0.777	**0.278**	-0.473	**0.286**	-0.459	**0.150**	-0.300	**0.173**	-0.303	**0.141**	-0.266	**0.152**	-0.305
	Dip	"Mean-elevation peak (symmetrical)"	**0.123**	-0.047	**0.254**	-0.078	**0.188**	-0.074	**0.199**	-0.078	**0.146**	-0.065	**0.131**	-0.066	**0.041**	-0.037	**0.197**	-0.078
		Flat	**0.110**	-0.165	**0.210**	-0.265	**0.195**	-0.279	**0.205**	-0.294	**0.157**	-0.248	**0.137**	-0.249	**0.031**	-0.119	**0.200**	-0.291
West	**Coll**	Altitude 3000	**0.543**	-1.020	**0.428**	-0.690	**0.446**	-0.555	**0.389**	-0.430	**0.313**	-0.338	**0.217**	-0.289	**0.218**	-0.279	**0.313**	-0.338
		Flat	**0.544**	-1.359	**0.442**	-0.947	**0.448**	-0.762	**0.414**	-0.609	**0.329**	-0.479	**0.242**	-0.422	**0.240**	-0.405	**0.329**	-0.479
	Dip	"Mean-elevation peak (symmetrical)"	0.022	-0.023	0.011	-0.0215	0.006	-0.017	0.003	-0.013	0.007	-0.006	0.001	-0.007	0.003	-0.013	0.007	-0.021
		Flat	0.020	-0.087	0.006	0.019	0.001	0.033	0.002	0.041	0.004	0.058	0.004	0.055	0.004	0.063	0.001	0.0327

MOLECULAR ECOLOGY
Gálvez-Reyes et. al. Supporting Information.

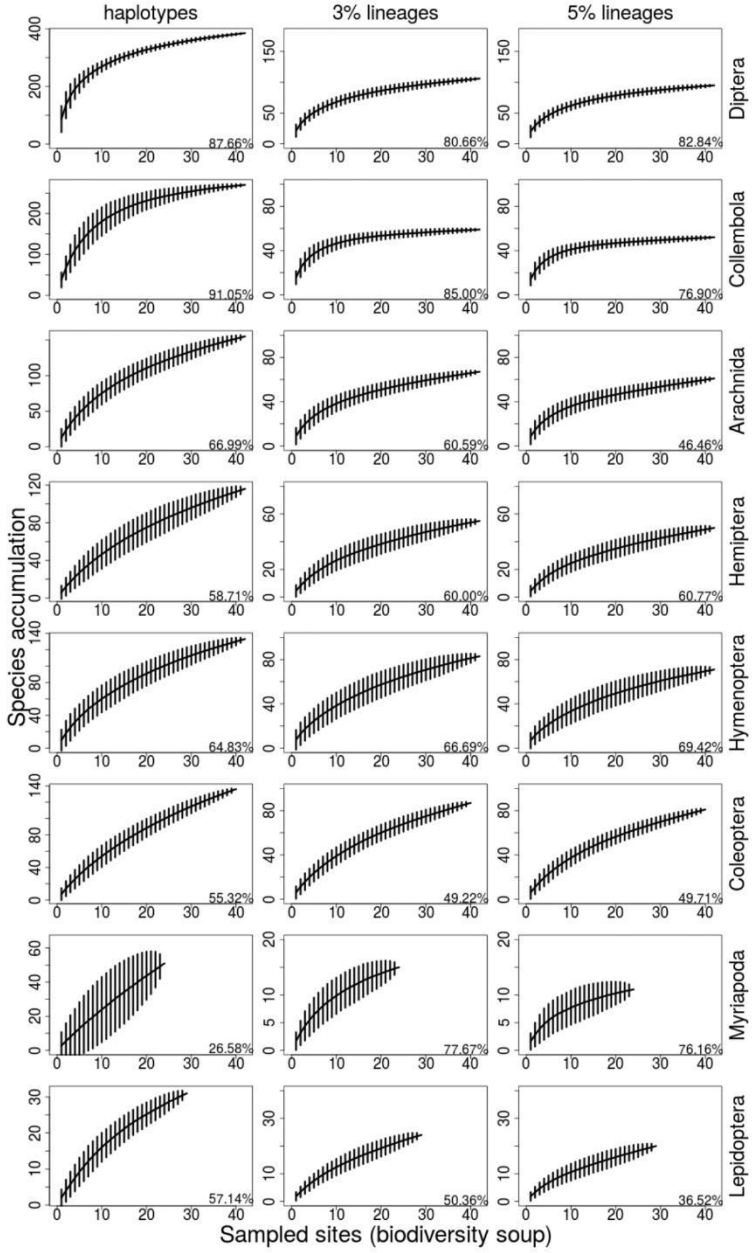

Figure S3. Taxa accumulation curves across sampled sites for each arthropod group and three lineage levels. We used 1,000 permutations, the random method, and calculated estimated completeness (Chao) versus observed completeness with the presence/absence data matrix.

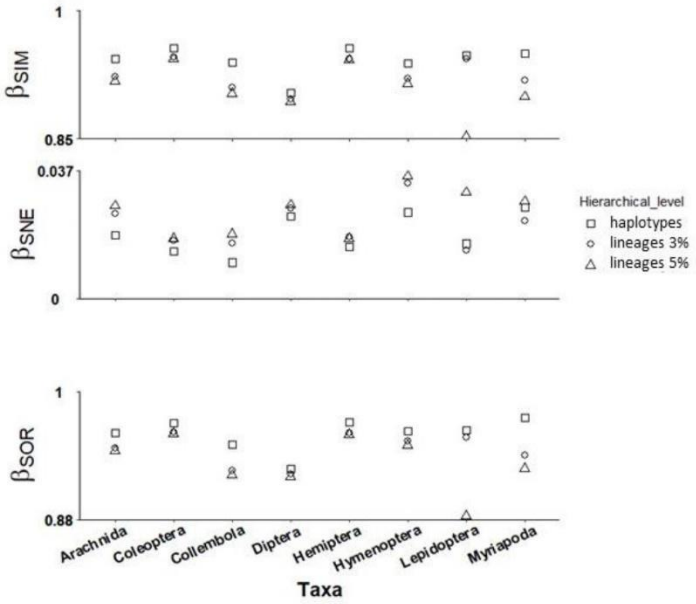

Figure S4. Multiple-site dissimilarity values of species turnover (β_{sim}) and nestedness-resultant (β_{sne}) components in arthropods computed among sites. β_{sor}: Sørensen dissimilarity index, β_{sim}: the Simpson dissimilarity index (turnover), and β_{nes}: the dissimilarity due to nestedness.

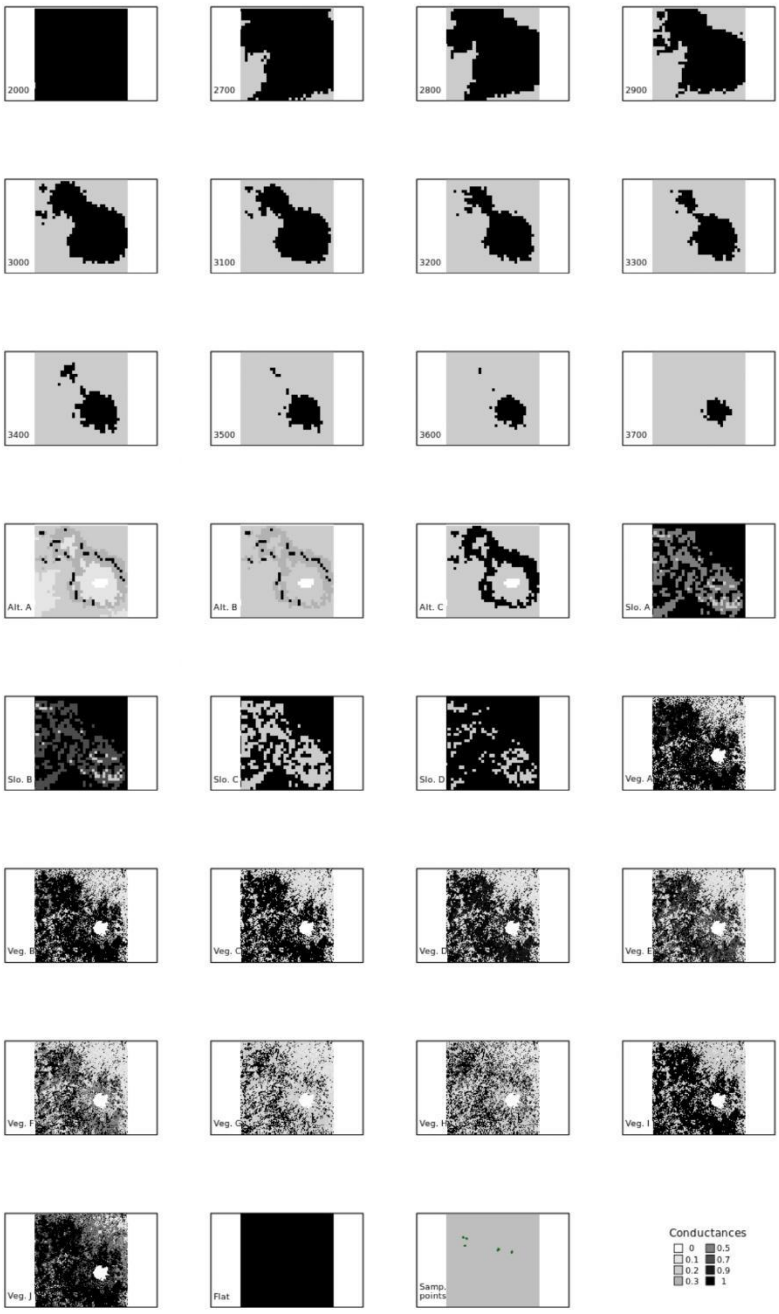

Figure S5. Resistance surfaces used to estimate effective distances among sites. Areas allowing higher dispersion are shown in black. The first four rows show the surfaces using the elevation

data; the last row shows a landscape where all cells have high conductance ('flat' landscape) and the 42 sampling points. Names of the resistance surfaces as in Table S2.

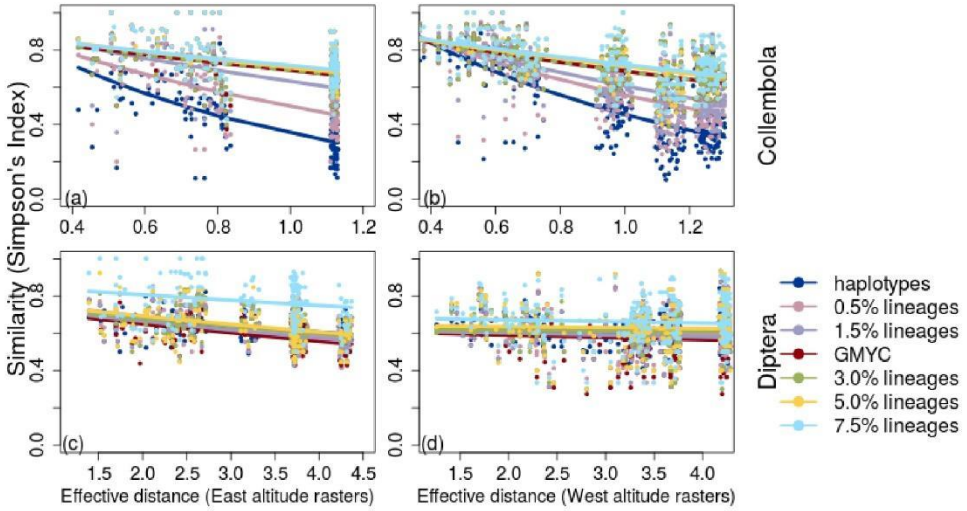

Figure S6. Distance decay of similarity against altitudes within Nevado de Toluca for Collembola (a,b) and Diptera (c,d) at fine geographic distances. Analysis at <5 km (a,c) using the Eastern subset of sampling sites (SJH and TLC) and at <2 km using the Western subset (AAB and ASB). Decay of similarity is shown against altitude rasters. Dots represent pairs of sampling sites for each ASV and lines are the fitted model for each lineage. Altitude rasters are as in Figure S5. Significance levels and r^2 are as in Table S5. Only Collembola showed significant values ($p <$ 0.01).

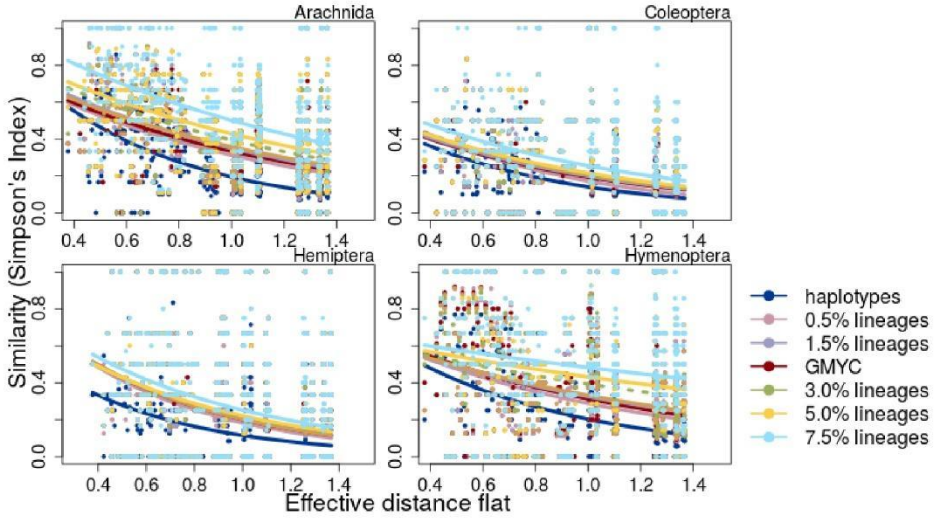

Figure S7. Distance decay of similarity against geographic distances within Nevado de Toluca for Arachnida, Collembola, Hemiptera and Hymenoptera at different hierarchical levels and species delimitation using GMYC. Dots represent pairs of sampling sites for each ASV and lines are the fitted model for each lineage. Significance levels and r^2 are as in Table S7. All tests were significative ($p < 0.01$).

5. References

Arribas, P., Andújar, C., Hopkins, K., Shepherd, M., & Vogler, A. P. (2016). Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. *Methods in Ecology and Evolution*, *7*(9), 1071-1081. https://doi.org/10.1111/2041-210X.12557

González-Fernández, A., Arroyo-Rodríguez, V., Ramírez-Corona, F., Manjarrez, J., Aguilera-Hernández, A., & Sunny, A. (2019). Local and landscape drivers of the number of individuals and genetic diversity of a microendemic and critically endangered salamander. *Landscape Ecology*, *34*(8), 1989-2000.

Lee-Yaw, J. A., Davidson, A., McRae, B. H., & Green, D. M. (2009). Do landscape processes predict phylogeographic patterns in the wood frog? *Molecular Ecology*, *18*(9), 1863-1874. https://doi.org/10.1111/j.1365-294X.2009.04152.x

ANEXO 3: Material suplementario artículo 3

Conservation Biology

Gálvez-Reyes *et al.* Supporting Information

Suplementary Information for: Strong turnover in the composition on forests subjected to management and reforested activities revealed by metabarcoding of arthropod communities

Gálvez-Reyes Nancy*[1,2], Piñero Daniel[1], and Mastretta-Yanes Alicia*[3,4].

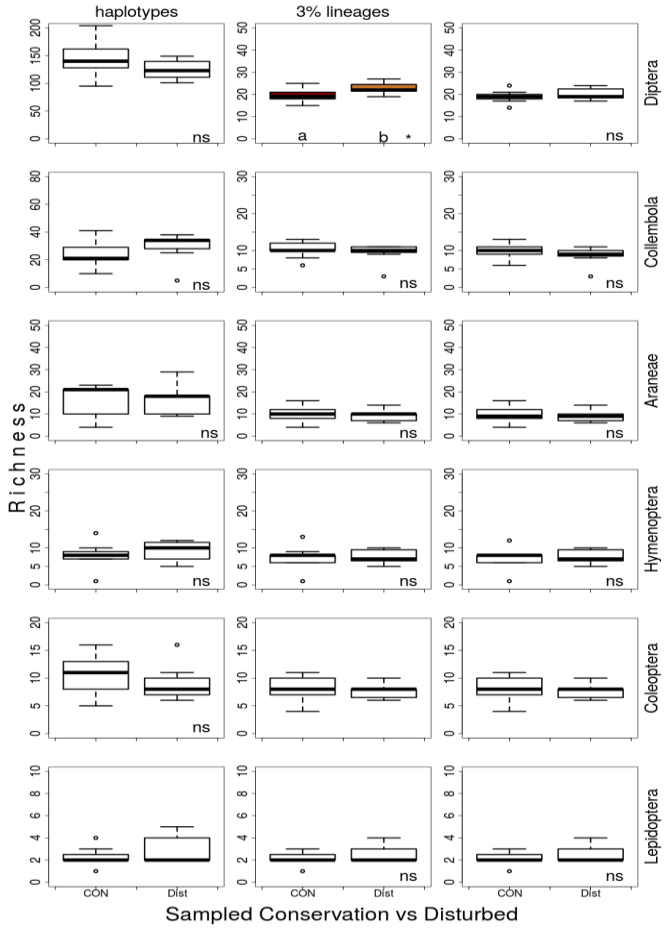

Figure S2. Arthropods richness by sampling Disturbed treatment. CON = Under conservation, and DIS= under disturbed within *Abies religiosa* forest. Same letters indicate that there are no significative differences among those sites $p<0.05$. Significance codes *:$p<0.05$, **:$p<0.01$, ***:$p<0.001$, ns: non-significant.

Table S1. Treatments in forests of *A. religiosa* and *P. hartwegii* with categories of forests (conserved, forest management, disturbance, and plantations).

Treatment	Categories	Geographical sites of mountains (locality)	Land-use change	Forest type	No. Of evaluations	Evaluation sampling sites.	Sample site replicates (No. blocks).	No. trap per block (Area=20x15m) -No. pitfall total	Collection date	masl	Biodiversity soup
(i)	Forest management Vs Forest under conservation	Nevado Toluca -Amanalco: San Bartolo	Forest Management: Chaponeo activity	*A religiosa*	2	3	3	**20** pitfall (2x3x3x20)= (18x20)=**360**	Lluvia (Ago-Sep 2015)	2500 a 3500	7x3x3=**63**
		Nevado Toluca -Amanalco: San Bartolo	Forest under conservation								
		Nevado Toluca -Amanalco: San Bartolo	Forest Management: Cut 7 years before ago		2			**20** pitfall (2x3x3x20)= (18x20)=**360**			
		Nevado Toluca -Amanalco: San Bartolo	Forest under conservation								
		Nevado Toluca -Amanalco: Agua Bendita	Forest Management: Cut 1 years before ago		3			**20** pitfall (3x3x3x20)= (27x20)=**540**			
		Nevado Toluca -Agua Bendita	Forest Management: Cut 10 years before ago								
		Nevado Toluca -Amanalca: Agua Bendita	Forest under conservation								
(ii)	Disturbed forest vs Forest under conservation	Nevado Toluca -Santiago Tlacotepec	Disturbed forest with cow	*A religiosa*	2	3	3	20 pitfall (2x3x3x20)= 18x20=**360**	Lluvia (Ago-Sep 2015)	2500 a 3500	2x3x3=**18**
		Nevado Toluca -Santiago Tlacotepec	Forest under conservation								
(iii)	Plantation forest vs Forest under conservation	Nevado Toluca -Santiago Tlacotepec	Plantation forest 1980	*P. hartwegii*	3	3	3	20 pitfall (3x3x3x20)= (27x20)=**540**	Lluvia (Ago-Sep 2015)	2500 a 3500	3x3x3=**27**
		Nevado Toluca -Santiago Tlacotepec	Plantation forest 2013								
		Nevado Toluca -Zinacantepec	Forest under conservation								
(iv)	Assessing diversity under two forest types	Nevado Toluca -San Juan de las Huertas -Tlacotepec	Forest under conservation	*A religiosa*	2	3	3	20 pitfall (2x3x3x20)= 18x20=**360**	Lluvia (Ago-Sep 2015)	2500 a 3500	2x3x3=**18**
		Nevado Toluca -San Juan de las Huertas -Zinacantepec	Forest under conservation	*P. hartwegii*	2	3	3	20 pitfall (2x3x3x20)= 18x20=**360**		2500 a 3500	2x3x3=**18**

Table S2. Number of haplotypes, lineages at 3% and 5%, putative species by GMYC and GMYC threshold values for eight arthropods orders.

Order		Haplotypes	Lineages 3%	Lineages 5%	GMYC	GMYC threshold
Diptera	2,521	385	106	95	124	0.0088
Collembola	2,370	270	59	52	60	0.029
Arachnida	1,027	155	67	61	78	0.01258205
Coleoptera	708	136	87	81	102	0.007255866
Hymenoptera	445	133	83	71	94	0.009765845
Hemiptera	240	116	55	50	64	0.01771803
Myriapoda	112	51	15	11	20	0.01529633
Lepidoptera	20	31	24	20	28	0.003
Total	7,443	1,277	496	441	570	

Table S3. Results of Isolation by distance in Diptera, Collembola, Arachnida, Coleoptera, Hemiptera and Hymenoptera at different hierarchical levels and species delimitation using GMYC. Associations with linear regressions between the pairwise effective distances and the pairwise beta diversity. The r values (r^2) are reported in each hierarchical level. All tests were significative ($p < 0.01$).

	Hierarchical level															
community	Haplotype		lineages 0.5%		lineages 1.5%		lineages 2%		lineages 3%		lineages 5%				GMYC	
Surface flat	Pseudo-r^2	Slope (b)	Pseudo-r^2	Slope (b)	Pseudo-r^2	Slope (b)	Pseudo-r^2	Slope (b)	Pseudo-r^2	Slope (b)	Pseudo-r^2	Slope (b)	Pseudo-r^2	Slope (b)		
Diptera	0.373	-2.658 e-05	0.298	-1.757 e-05	0.217	-1.554 e-05	0.218	-1.456 e-05	0.180	-1.328 e-05	0.199	-1.334 e-05	0.064	-6.430 e-05	0.257	-1.667 e-05
Collembola	0.819	-0.0003 e-05	0.824	-0.0001 e-05	0.732	-6.749 e-05	0.770	-6.393 e-05	0.760	-5.943 e-05	0.712	-4.964 e-05	0.710	-4.512 e-05	0.759	-5.964 e-05
Araneae	0.333	-1.833	0.165	-1.046	0.171	-0.979	0.155	-0.993	0.179	-0.876	0.164	-0.817	0.299	-0.799	0.161	-0.974
Coleoptera	0.164	-1.667	0.180	-1.494	0.124	-1.192	0.127	-1.282	0.130	-1.194	0.117	-1.118	0.099	-1.049	0.127	-1.215
Hymenoptera	0.182	-1.530	0.112	-0.999	0.086	-0.815	0.085	-0.797	0.054	-0.556	0.046	-0.469	0.036	-0.372	0.096	-0.936