Breaking ground

Plant diversity assessments through soil eDNA metabarcoding

María de los Ángeles Ariza Salazar

Dissertation presented for the degree of
Philosophiae Doctor (PhD)
2023

Natural History Museum &
Faculty of Mathematics and Natural Sciences
University of Oslo

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2620

ISSN 1501-7710

Cover: UiO.
Print production: Graphic Center, University of Oslo.

One cannot separate the plant
-neither physically nor metaphysically-
from the world that accommodates it.
It is the most intense, radical, and paradigmatic
form of being in the world.
To interrogate plants means to understand
what it means to be in the world.

The Life of Plants: A Metaphysics of Mixture
Emmanuelle Coccia

Preface

I was submerged in what seemed an enchanted scenario from the movies. I was jumping on moss mats like a fairy that bounces from cloud to cloud. I was walking by the side of little streams, climbing up the mountains, and from the top, I was sliding down. I was surrounded by brown, blue, and green colors. I was crossing mighty rivers and slipping into mires. I was singing "Agua" by Lido Pimienta and "Gracias a la Vida" by la Negra Sosa. I was never fully warm or cold. My pockets were full of mushrooms and my lips were blueberry blue. Looking to the ground I found newborns and fallen treasures from the giants. I was burying my fingers in the soil and thinking about mi abuela mama Carmen's hands. I was touching the world below and seeking secrets kept for those who dig deep. I was collecting ground truths settled in time and thinking *what goes around comes around* and then comes down. I was packing pictures from the underground to create an album of nature's presence.

These are just a few poetic memories from my fieldwork at the Solhomfjell forest reserve, southern Norway, where I started the collection of soil samples for the research presented in this dissertation. For someone with fieldwork experience mostly in tropical ecosystems, this first experience in a Norwegian temperate forest was, to say the least, unforgettable. I considered myself lucky to see this forest at its splendor before the fast northern autumn stripped the green and the long winter painted it white. Little did I know that by breaking ground with a hammer I would be able to see this forest many times more but through the eyes of the ground: DNA.

Acknowledgments

Daydreaming in front of the lake Atitlán and its astonishing mountains and volcanoes, I started to draft thank you notes on my phone so I wouldn't forget all the amazing people that contributed, in one way or another, to this thesis. The year 2022 had just started and I was already fantasizing about the year ahead: I will start a new job in science communication, I would finally finish my PhD and do all those things that I had suspended by saying "after the PhD because I don't have time now". Little did I know that 2022 would be one of the most challenging years of my life, and that I wouldn't see the end of my PhD until the start of spring 2023. Many times, I wanted to give up and just "hang my research gloves" as I didn't think I had what it takes to be a researcher or the courage to be so competitive and climb the academic ladder. But Hugo, my supervisor, helped me to see it through by giving me my space when I needed but also a push to overcome it. Hugo, gracias, for keeping me on track, for letting me pursue my ideas, even the dumb ones, and for being there for me every step of the way. I haven't been an easy-peasy student over the course of the four years, so thank you for your incessant support.

The research presented in this thesis wouldn't have been possible without the brilliance of Inger and Rune. Takk! You two are my references of true plant ecologists. If I gain a quarter of your knowledge and capacities throughout my career, I will be satisfied. Special thanks to mi amiga Eva Lieungh for her devoted support during the weeks prior to the submission of this thesis, you rock! To Quentin, Lovisa and the rest of EDGE group for your encouragement during the PhD years. I want to thank Andreas Wollan and Mari Engelstad for sharing with me beautiful days collecting soil samples in Solhomfjell and Hvaler forest. Also big thanks to Jarl Andreas, Lisbeth and Audun, from the NHM DNA lab for their support and guidance. Big shout out to all the NHM PhD and master students with whom I shared my PhD journey, because suffering together is always better!

I want to thank the Plant.ID crew for giving me this great research opportunity, to my co-supervisors Bengt and Inger for being so supportive while exploring my own ideas and making mistakes, and special thanks to Brecht and Marcella for their love and care for me from day one. One of the biggest gifts of Plant.ID is that it drove me closer to three beautiful souls that now have a piece of my heart. Nataly, gracias por tu amor y todos esas aventuras en los viajes, sos pura luz en mi vida negrita. Marcel, estas buenísimo joder! lol Es mucho lo que te quiero and your friendship is a big treasure in my life. Anneke, who drew the beautiful illustrations presented in this thesis, thank you for being there ready to dance and hang-out when I most needed.

I wouldn't have kept up with academia if it wasn't by the backup of Sole and the wonderful team at Transmitting Science, and the heartfelt accompaniment of Ana Pineda and the Sustain community. Thank you all for showing me what a mindful and committed scientist is and for pushing the boundaries towards a more fair and egalitarian academia. I also wouldn't have kept my sanity without the reassurance and care of the IMAE family and the waves of love that mi mamá, papá, Luisar, Marta, Raul, Isabelita and Delfi send from Guatemala. Big thanks to Marit, Svein Olav and their lovely family and friends, who became my family in Oslo and kept me afloat with lovely music and Norwegian folklore. Last but never least, Bertrand, who has been there for me since I started my academic journey and never stops believing in me. You and Obsi are my grounding rocks, les amo.

Table of Contents

List of Papers

Paper I. Environmental and biodiversity assessments through eDNA analysis

Ariza, M., Garcés-Pastor, S., de Boer, H.J., 2022. Environmental and biodiversity assessments, in: de Boer, H.J., Rydmark, M.O., Verstraete, B., Gravendeel, B. (Eds.), *Molecular Identification of Plants: From Sequence to Species.* Advanced Books, Pensoft Publishers, pp. 354–371.

Paper II. DNA from soil: considerations and study design

Ariza, M., Garcés-Pastor, S., de Boer, H.J., 2022. DNA from Soil, in: de Boer, H.J., Rydmark, M.O., Verstraete, B., Gravendeel, B. (Eds.), *Molecular Identification of Plants: From Sequence to Species.* Advanced Books, Pensoft Publishers, pp. 354–371.

Paper III. Plant biodiversity assessments through soil eDNA metabarcoding reflects local and regional diversity

Ariza, M., Fouks, B., Mauvisseau, Q., Halvorsen, R., Alsos, I.G., de Boer, H., 2023. Plant biodiversity assessment through soil eDNA reflects temporal and local diversity. *Methods in Ecology and Evolution* 14 (2), 415-430. doi:10.1111/2041-210X.13865

Paper IV. Evaluating the feasibility of using plant-specific metabarcoding to assess forest types from soil eDNA.

Ariza, M., Engelstad, M., Lieungh, E., Laux, M., Ready, J., Mauvisseau, Q., Halvorsen, R., de Boer, HJ. Evaluating the feasibility of using plant-specific metabarcoding to assess forest types from soil eDNA. *Manuscript in Preparation.* Targeted Journal: Applied Vegetation Science.

Summary

Biodiversity assessments are the foundation upon which insights in composition, state, and threats of nature are based. Plants are central to most biodiversity assessments as they are virtually everywhere and indicate the presence of other diversity, environmental and anthropogenic features, and ecosystem types. Traditional assessments of plant diversity are usually carried out through morphological identification of plant traits, e.g., leaf and flower shape, that circumscribe species, genera, families, etc. However, this task is a rigorous and time-consuming process that requires specialized skills and can only be carried in the seasons when these traits are available. As threats to nature continue unabated, sound and reliable operational approaches have never been so urgent to discover, monitor and protect plant biodiversity. The analysis of DNA from environmental samples and substrates, i.e., air, bulk, faecal, sediments, soil, and water, has opened new avenues for expedited identification of biodiversity, as many organisms can be detected in parallel in a single and cost-effective effort. Nevertheless, the spatial and temporal signals of plant environmental DNA (eDNA) present in environmental samples remain poorly understood. This limitation obscures the conclusions that can be derived from eDNA assessments and hinder its applications for sound and complete plant assessments. In this context, the overarching aim of this thesis is to advance eDNA plant assessments. As soil is, presumably, the substrate where most plant DNA accumulates, this thesis places the spotlight on soil eDNA samples. Four papers are presented, from which two are peer-reviewed book chapters that review our current knowledge about plant eDNA samples and DNA from soil for its application to complete and sound biodiversity assessments, and two original research articles that evaluate the power of soil eDNA assessments to monitor plant diversity and determine ecosystems types. By breaking ground -literally and metaphorically- this thesis particularly shows how soil eDNA samples can serve to diagnose local, regional, past and present plant diversity and aid the characterization of the main features of forest types. The applications of these results are broad, from site-specific assessments to land-cover mapping, as they provide a baseline for decision-making in soil eDNA studies. Altogether, this thesis highlights how DNA-based identification is a key tool to help meet the biodiversity challenges of the twenty-first century.

Sammendrag

Vurderinger av biodiversitet er grunnlaget for innsikt i sammensetning, tilstand og endring i naturen. Planter er sentrale i de fleste vurderinger da de er nesten overalt og indikerer tilstedeværelsen av annen diversitet, miljø- og antropogene egenskaper og økosystemtyper. Tradisjonelle vurderinger av plantemangfold utføres vanligvis ved morfologisk identifikasjon av plantenes egenskaper, for eksempel blad- og blomsterform, for å identifisere arter, slekter, familier osv. Denne oppgaven er imidlertid en både grundig og tidkrevende prosess som krever spesialisert kunnskap og kan bare utføres i de årstidene når disse egenskapene er tilgjengelige. Da trusler mot naturen fortsetter, har pålitelige og solide metoder aldri vært så presserende for å oppdage, overvåke og beskytte plantebiodiversitet. Analysen av DNA fra miljøprøver og substrater, det vil si luft, bulk, avføring, sedimenter, jord og vann, har åpnet nye muligheter for rask identifikasjon av biodiversitet, da mange organismer kan oppdages parallelt i en enkelt og kostnadseffektiv metode. Imidlertid er signalene fra planters miljø-DNA (eDNA) i tid og rom dårlig forstått. Denne begrensningen skjuler konklusjoner som kan trekkes fra eDNA-vurderinger og hindrer dens anvendelser for solide og komplette vurderinger av planter. I denne konteksten er det overordnede målet med denne avhandlingen å fremme eDNA-vurderinger av planter. Siden jord antagelig er substratet der mest plante-DNA akkumuleres, fokuserer denne avhandlingen på jord eDNA-prøver. Fire arbeid presenteres, hvorav to er fagfellevurderte bokkapitler som presenterer vår nåværende kunnskap om plante-eDNA-prøver og DNA fra jord for dets anvendelse i vurderinger av biodiversitet, og to forskningsartikler som evaluerer kraften til jord-eDNA-vurderinger for å overvåke plantemangfold og bestemme økosystemtyper. Ved å bryte bakken - bokstavelig talt og metaforisk - viser denne avhandlingen spesielt hvordan jord eDNA-prøver kan brukes til å diagnostisere lokalt, regionalt, fortid og nåtid plante mangfold og hjelpe til med å karakterisere hovedtrekkene ved skogtyper. Bruksområdene for disse resultatene er brede, fra stedsspesifikke vurderinger til kartlegging av landdekke, da de gir en grunnlinje for beslutningsprosesser i jord-eDNA-studier. Samlet sett fremhever denne avhandlingen hvordan DNA-basert identifikasjon er et nøkkelverktøy for å møte biodiversitetsutfordringene i det 21. århundre.

Resumen

Las evaluaciones de la biodiversidad son la base para el desarrollo de conocimientos sobre la composición, el estado y las amenazas de la naturaleza. Las plantas son fundamentales para dichas evaluaciones, ya que están prácticamente en todas partes y son indicadores de la presencia de otras especies de organismos, tipos de ecosistemas y características ambientales y antropogénicas. Las evaluaciones tradicionales de la diversidad taxonómica en plantas generalmente se llevan a cabo mediante su identificación morfológica, por ejemplo, la forma de las hojas y las flores. Sin embargo, esta tarea es un proceso riguroso y lento que requiere habilidades especializadas y solo se puede llevar cuando dichos rasgos son visibles.

A medida que las amenazas a la naturaleza continúan, se hace más urgente la creación de enfoques operativos sólidos y confiables para descubrir, monitorear y proteger la biodiversidad de las plantas. El análisis de ADN de muestras ambientales y sustratos, como el aire, granel, heces, sedimentos, suelo y agua, ha abierto nuevas vías para la identificación rápida de la biodiversidad, ya que se pueden detectar muchos organismos en un solo análisis. Sin embargo, las señales espaciales y temporales de las plantas en las muestras de ADN ambiental (eDNA, por sus siglas en inglés) siguen siendo poco conocidas. Esta falta de conocimiento limita las conclusiones que se pueden derivar de las evaluaciones de eDNA y dificulta sus aplicaciones. En este contexto, el objetivo general de esta tesis es avanzar el conocimiento acerca de las evaluaciones del eDNA para identificar plantas. Este estudio se enfoca en las muestras de eDNA del suelo, ya que es el sustrato donde se acumula la mayor parte del ADN vegetal. Incluye cuatro capítulos, de los cuales dos se refieren a artículos de revisiones bibliográficas del conocimiento actual sobre la aplicación de muestras de eDNA para la evaluación de las plantas y del ADN de plantas en el suelo, y dos se refieren a artículos de investigación originales que evalúan el poder de las evaluaciones de eDNA del suelo para monitorear la diversidad vegetal y determinar los tipos de ecosistemas. Esta tesis es la primera en su género y muestra en particular cómo las muestras de eDNA del suelo pueden ser útiles para diagnosticar tanto la diversidad vegetal tanto en términos espaciales como en términos temporales, y así ayudar a la caracterización de las principales características de los tipos de bosques. Los resultados que aquí se muestran

tienen amplias aplicaciones,desde evaluaciones específicas de la biodiversidad de un sitio hasta el mapeo de la cobertura terrestre, ya que proporcionan información básica para futuros estudios de eDNA del suelo. En resumen, esta tesis destaca que la identificación de la biodiversidad basada en ADN es una herramienta clave para ayudar a enfrentar los desafíos sobre este tema, en el siglo XXI.

Aims and structure of the thesis

Plants are central to most biodiversity assessments as they are virtually everywhere and indicate the presence of other diversity, environmental and anthropogenic features, and ecosystem types. Environmental and biodiversity assessments are the foundation upon which the state of, and threats to, nature are recognized and monitored. The rapid decline of biodiversity stresses the urgency for innovative and rapid operational approaches to identify, map and protect life on Earth. The analysis of DNA from environmental substrates i.e., eDNA-based assessments from e.g., air, water, and soil, has attracted interest by facilitating parallel identification of many organisms from a single sample. As soil is, presumably, the substrate where most plant DNA accumulates, soil eDNA analysis holds great potential to study plant diversity . In this context, the overarching aim of this thesis is to explore the potential and limitations of using DNA extracted from soil to assess plant biodiversity, and to develop a more comprehensive understanding of the spatial and temporal plant signals present in soil samples. Through literature review and empirical investigations, the aims of the thesis are two-fold. It aims to determine best practices for soil sample collection, processing, DNA extraction, and barcode amplification. Additionally, it seeks to evaluate the ability of soil eDNA assessments to accurately detect and characterize plant communities, composition gradients and ecosystem types, and to compare these findings to traditional vegetation surveys. Ultimately, this work aims to contribute to the development of more efficient and accurate methods for assessing plant biodiversity in natural ecosystems.

The specific aims for each of the papers presented in this thesis are as follows. **Paper I**: Review studies employing substrates from which plant DNA can be extracted to inform about the applications, power and limitations of plant eDNA assessments. **Paper II:** Review literature on the state and composition of plant DNA present in soil to provide an overview of pertinent considerations for sample collection in the field, sample processing, DNA extraction and amplification, as well as to identify current knowledge gaps from soil eDNA assessments. **Paper III**:Investigate the temporal and spatial resolution of soil eDNA plant assessments and the relationship between plant DNA detection, aboveground plant biomass and read sequence counts. Evaluate the power

and limitations of soil eDNA assessments relative to traditional vegetation surveys. **Paper IV**: Investigate if representative plant composition can be detected from soil eDNA. Evaluate to what extent the plant composition determined from soil eDNA characterizes ecosystem types defined by their theoretical composition. Assess similarities and differences between gradients in species composition identified by soil eDNA assessments and visual assessments.

The thesis is structured as a kappa and articles. The kappa presents the broader context of the research as well as a synthesis and conclusion of its contribution to the field. The four articles are two peer-reviewed chapters (Paper I and Paper II) from the book "Molecular Identification of Plants: From Sequence to Species" (2023, *Advanced Books*) and two original research papers (Paper III, a published paper, and Paper IV, a manuscript). The two chapters provide a framework and contextual reference for the subsequent original research papers. Additionally, a contribution was made to the book chapter on amplicon metabarcoding from the same book (2023, *Advanced Books*), where the main advantages and disadvantages of plant DNA metabarcoding are discussed together with the most common markers used. Original research papers deal with more applied aspects of soil eDNA plant assessments i.e., monitoring through time and space and ecosystem mapping, and highlight the power and limitations of this tool.

Introduction

The global biodiversity crisis and the detrimental impacts of humanity on nature are now indisputable and, regrettably, evident[1]: Net declines in the conditions of ecosystems, communities, and populations have been observed over recent decades (Cardinale et al., 2012; Hoskins et al., 2020). Less than a quarter of Earth's surface remains without any sort of human print (Jones et al., 2018; Watson et al., 2016), and the most accessible biomes have been substantially modified to the point that now are considered "anthromes" (Blondel, 2006; Ellis and Ramankutty, 2008). Accelerating land- and sea-use change, climate change, pollution, over-exploitation of natural resources, and alien-species invasion are driving the extinction of more than one million animal and plant species across the globe (Brondízio, 2019). More than five precautionary 'Planetary Boundaries' have been crossed (Steffen et al., 2015) and most of the Aichi Biodiversity Targets to mitigate its loss have not been met (Meehan et al., 2020; Shepherd et al., 2016).

Environmental and biodiversity assessments are the cornerstone upon which the above processes have been identified, described, monitored, and reported. These usually rely on inventorying diversity, i.e., counting the number of species, their abundance, and distribution, to draw trends over time and space in relation to environmental and anthropogenic factors (Lughadha et al., 2016). The discovery and recognition of organisms has historically focused on the evaluation of morphological characters that circumscribe the identity of e.g., species, genera, families, etc (Hammond, 1992). This task is a rigorous and time-consuming process that requires specialized skills and thus is often commissioned to taxonomists. While approximately 18,000 species are described each year (Wheeler and Pennak, 2011), many others are disappearing before they are discovered as a consequence of e.g., the accelerated pace of extinction (Ceballos et al., 2015). This race against time is further hampered by the increasing shortage of taxonomists and global disparities in science resources and access (Engel et al., 2021; Moura et al., 2018).

[1] While the root cause of the current state of biodiversity is humans, not all individuals or societies bear the same contribution and/or responsibility and, therefore, are not equally accountable (Leigh et al. 2019; Liao et al. 2022; Chancel 2022; Sponsel 2013).

As threats to nature continue unabated, sound and reliable operational approaches have never been so urgent to discover, monitor and protect biodiversity. The analysis of DNA traces from environmental samples, i.e., air, soil, water, etc., has opened new avenues for expedited identification of biodiversity, as many organisms can be detected in parallel in a single and cost-effective effort (Beng and Corlett, 2020; Bohmann et al., 2014; Thomsen and Willerslev, 2015). Moreover, the identification of plant diversity can be of advantage as it is a predominant component across most ecosystems and signals the presence of broad taxonomic realms together with environmental and anthropogenic features (Brummitt et al., 2020; Brunbjerg et al., 2018).

Plant diversity: state, threats, and assessment

One cannot escape plants[2], they are virtually everywhere[3]: in the air we breathe i.e., in the form of pollen, in the food we eat and the clothes we wear, in our medicines, building materials, and fuels (Antonelli et al., 2020; Molina-Venegas et al., 2021). Throughout history, people and plants have been tightly interconnected, e.g., cultures configured plant diversity by domesticating and translocating plant species, and vice versa, plant diversity shaped cultural practices, e.g., healthcare (Teixidor-Toneu et al., 2018). Moreover, plants are paramount to life on Earth as they are responsible for primary production and carbon sequestration, i.e., carbon dioxide conversion to oxygen and glucose; water cycling, temperature regulation, and food chain stabilization (Díaz and Cabido, 2001).

Along ~475 million years, plants diversified from single-celled algae into a myriad of land plants, i.e., green-algae, bryophytes, lycophytes, pteridophytes, gymnosperms, and angiosperms (Fiz-Palacios et al., 2011; Silvestro et al., 2015). Angiosperms are both the most diverse and important group of plants for humans. By 2021, an estimated 340,000–390,000 vascular plant species were known to science and yearly new discoveries amount to 2,000 (Cheek et al., 2020; Govaerts et al., 2021). Nevertheless, two in five plants are estimated to be at risk of extinction whilst many remain unassessed (Antonelli et al., 2020).

[2] However, plants are largely ignored, considered as the background of nature and thought less than animals. Wandersee and Schussler (1999) coined the term "plant blindness" to describe this trend, which now is institutionalized throughout society and represents a big challenge to the conservation of plant diversity.

[3] Indeed, amongst life kingdoms, plants account for most of Earth's biomass i.e., ~450 gigatons of carbon (Bar-On et al. 2018).

The primary driver of plant diversity loss globally is habitat conversion to either agricultural or urban areas, yet, the over-exploitation of specific plant species encompasses many other threats (Corlett, 2020; Vellend et al., 2017). Tropical hotspots of plant diversity, e.g., Amazonia and Congo Basin are focal points for the extraction of plant resources and therefore are in particular danger (Souza and Prevedello, 2020). Plant species within boreal and temperate forest areas are the most managed, yet, these habitats are increasingly degraded because of their susceptibility to climate change and increasing fire risk (IPBES, 2019; Machado Nunes Romeiro et al., 2022). Finally, habitat fragmentation, climate change and illegal trade are changing the distribution of more than 13,000 plant species (van Kleunen et al., 2010), of which 6,000 are known to be invasive in any part of the world (IPBES, 2019).

Traditional plant assessments, i.e., morphological identification of plants (namely vegetation surveys or visual assessments throughout this thesis), are the pillar upon which our current knowledge of plant diversity has been built and the correspondence between species and environmental gradients understood (Hagen, 2001; Keddy, 2005). Moreover, biodiversity monitoring and conservation planning often draw trends and conclusions from the assessment of vascular plant diversity, while relying on land-cover maps that describe the spatial distribution of vegetation communities and abiotic features (Halvorsen et al., 2020; Saah et al., 2019). In the field, plant assessments are usually[4] carried out through systematic inventorying of taxa, i.e., floristic inventories of quadrats or transects, together with their attributes, e.g., abundance and traits. This often results in long days in the field as species or other circumscribed broader groups, e.g, grasses, sedges and mosses, are identified one by one through morphological characters. The identification of plant species is based on the evaluation of floral and leaf characteristics. Characters are often only available at particular phenological time windows. In addition, not all morphological characters can be inspected in the field, and this makes it necessary to collect plant material for identification using proper equipment. Furthermore, morphological identification of plant species can be challenging as (1) character plasticity can lead to incorrect identifications; (2) morphologically cryptic taxa, which are common in many plant groups, can be neglected; (3) morphological keys

[4] Another searching strategy is simply following the "botanic internal algorithm", which consists of a combination of ability, experience, expertise, and intuition (Palmer et al. 2002).

are not available for all diversity thus many individuals cannot be identified, and in addition, (4) require a high level of expertise (Hebert et al., 2003). Molecular identification of species, i.e., chemical profiles and DNA, resolves some of these challenges, yet, no single method provides an optimal identification across the Tree of Life.

DNA-based plant identification

Deoxyribonucleic acid (DNA) is the underlying molecule of life and carries genetic instructions for the development, functioning, growth and reproduction of all known organisms. In a nutshell, DNA is composed of four basic units, the nucleotide: Adenine, Cytosine, Guanine, and Thymine; the sequence (primary structure) and bonds between these (secondary structure) make up a double-strand helix (tertiary structure) that encodes the instructions (genes) for protein formation (Watson and Crick, 1953). In this way, the DNA sequence of an organism (order of nucleotides in a particular way) is the blueprint that describes singular genetic characteristics. DNA-based taxon identification scans this diversity amongst DNA sequences to diagnose organisms (Hebert et al., 2003), similar to that of barcodes in a supermarket[5].

Like most eukaryotic cells, the majority of plant DNA is located in the nucleus. Mitochondria, chloroplasts and other plastids also contain DNA[6] but their genomes are much smaller than that of the nucleus (Robles and Quesada, 2021). There are multiple plastids per cell, and plant DNA extracts are replete with a high copy number of plastid genomes (Twyford and Ness, 2017). Moreover, plastid genomes have substantial rates of nucleotide substitution across phylogenetic clades which make them suitable markers for plant discrimination (Petit and Vendramin, 2007). The Plant Working Group of the Consortium for the Barcode of Life (CBOL) reviewed candidate plastid DNA regions (*atpF–atpH* spacer, *matK* gene, *rbcL* gene, *rpoB* gene, *rpoC1* gene, *psbK–psbI* spacer, and *trnH–psbA* spacer) and recommended both *rbcL* + *matK* as the plant barcodes because their recoverability, sequence quality,

[5] Hebert et al. (2003) in their seminal paper coined the term "DNA barcodes", now widely used to refer to DNA-based identification.
[6] The organization and inheritance patterns of these organellar DNA are quite different from that of nuclear DNA.

and levels of taxon-interspecific variance (CBOL Plant Working Group, 2009). In addition, the p6 loop of the trnL (UAA) intron has gained attention as its short and variable sequence can be easily recovered, even from fragmented DNA (Taberlet et al., 2007). Moreover, the internal transcribed spacer 2 (ITS2) region of nuclear ribosomal DNA has been also proposed because of its sufficient variability to discriminate closely related species (Yao et al., 2010). As no single barcode provides resolution for all taxa (Fazekas et al., 2012), DNA-based plant assessments often use a combination of barcodes e.g., *rbcL*, *matK*, *trnL*, and ITS2 are routinely used (Kress, 2017).

The process of identifying organisms through DNA entails the availability of reference sequence libraries where unknown sequences can be matched for identification (Kress, 2017). The remarkable advancement of DNA sequencing technologies over the last decade has not only increased the number of organisms for which reference sequences are available but also prompted the proof-of-concept of DNA barcoding to its application in species discovery (Kress et al., 2009), forensics (Johnson et al., 2014), natural product authentication (Raclariu et al., 2017), early detection of invasive species and illegal trade (Williamson et al., 2016). Moreover, the parallel amplification of DNA barcode(s) of many organisms i.e., metabarcoding from environmental DNA (eDNA) i.e., DNA traces present in bulk samples[7], feces, sediments, soil, water, or after filtering air or water (Taberlet et al., 2012), have uncovered hidden and past diversity (Anderson-Carpenter et al., 2011; Hartvig et al., 2021), airborne and water communities (Foster et al., 2020; Johnson et al., 2019), new species interactions and herbivore food webs. (García-Robledo et al., 2013). Nevertheless, eDNA-based assessments remain limited to inform about the biomass of species, and the ecological and conservation status (Beng and Corlett, 2020). In addition, PCR-errors can yield false positives and mix signals from different samples (Bellemain et al., 2010). While the application costs of eDNA metabarcoding are becoming cheaper, the added value of this tool in comparison to traditional varies amongst the Tree of Life (Bohmann et al., 2014).

[7] Refers to the collection of raw material e.g., the content of a Malaise trap. In the context of plant bulk samples, however, there are natural bulk samples i.e., pollen samplers or that from pollinating vectors, and those that are artificially assembled such as collected roots, leaves, or flowers.

Material and methods

An overview of all materials and methods used in the experimental chapters (Paper III, Paper IV) of this thesis is outlined in this section. Detailed information can be found in the manuscripts and supplemental materials.

Study region

Study sites for Papers III and IV are the Solhomfjell Forest Reserve (SFR) and Hvaler archipelago, in southwest Norway, respectively (Figure 1). SFR is situated in the Solhomfjell area, Gjerstad, Aust-Agder County. The landscape is hilly and surrounded by deep valleys at all margins (350–480 m.a.s.l). Over the last one hundred years, the area had little human intervention. Hvaler archipelago is within the jurisdiction of Østfold County, and situated in the outer Oslo fjord. The terrain is dominated by rounded hills interrupted by rift valleys with a strong legacy of farming, grazing and haymaking. Both locations are within the boreo-nemoral[8] bioclimatic zone, which is characterized by abundant rainfall (750 to 1500 mm per year; Moen (1999), and fall within the distribution of boreal and temperate broadleaf forest biomes (Dinerstein et al. 2017; Figure 1). In addition, these sites are marked by the presence of forests[9] dominated by deciduous trees i.e., oak (*Quercus* spp.), maple (*Acer platanoides* L.), beech (*Fagus sylvatica* L.), aspen (*Populus tremula* L.), rowan (*Sorbus aucuparia* L.) and/or conifers i.e., spruce (*Picea abies* L.) and pine (*Pinus sylvestris* (L.) H.Karst.). Common understory species include bilberry (*Vaccinium myrtillus* L.), lingonberry (*Vaccinium vitis-idaea* L.), wood anemone (*Anemone nemorosa* L.), and wood sorrel (*Oxalis acetosella* L.). In addition, lichen and mosses, e.g., haircap moss (*Polytrichum commune* Hewd.), sphagnum moss (*Sphagnum* spp.), feather Moss (*Pleurozium schreberi* (Willd. ex Brid.) Mitt.), can be found covering the understory floor mats (Eriksen, 2017; Økland, 1995).

[8] A term to describe biome transition areas where temperate deciduous forest and boreal forest intergrade.

[9] According to the definition of Nature in Norway (NiN), forests are defined as natural ground that is strongly marked by the prolonged influence of trees (Halvorsen et al. 2020).

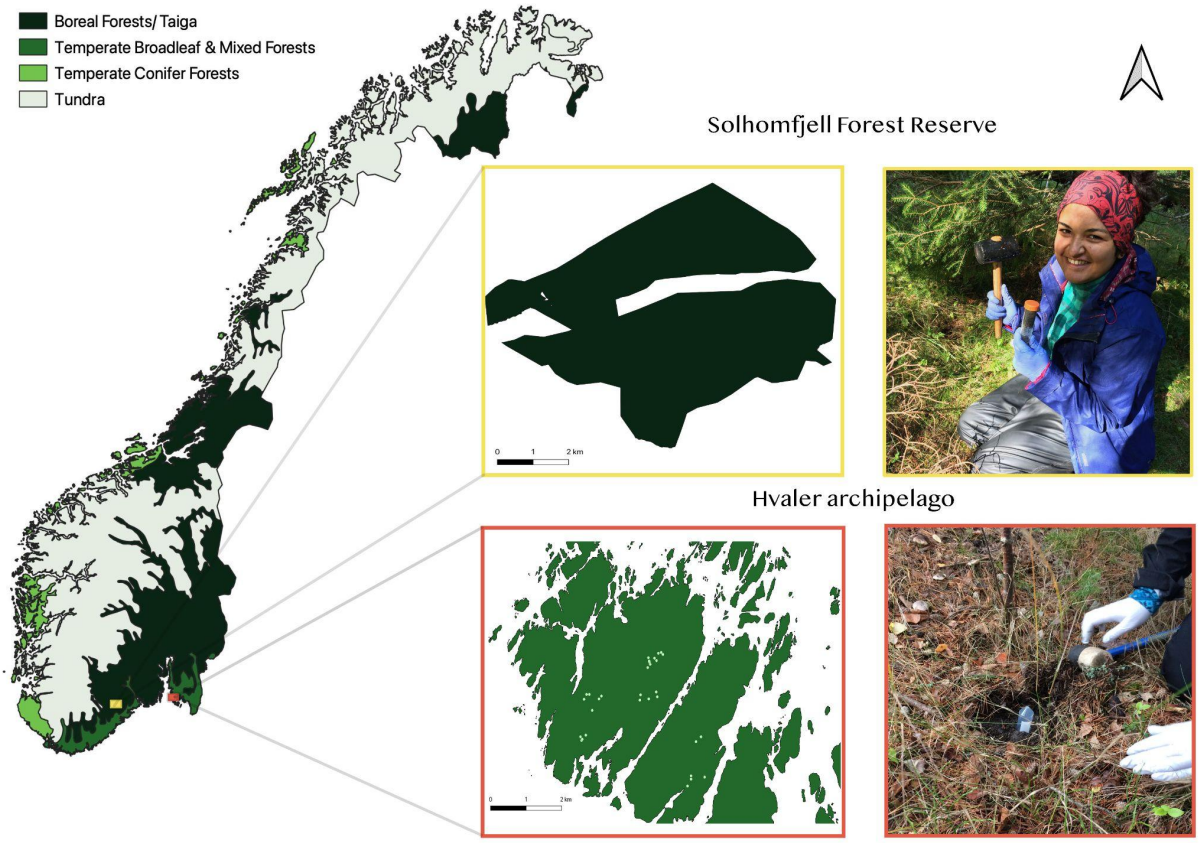

Figure 1. Study sites and soil sampling locations for Papers III and IV

Biome names and distribution in Norway follow Dinerstein et al. 2017. For visualization purposes, study sites are represented by yellow and orange squares, but these do not correspond to the actual size of the areas. Points at the Hvaler site map represent sampling locations.

Soils in these locations are usually podzols formed in situ and are generally characterized by having (a) a thick[10] organic surface layer (O horizon) rich in roots that, in the lower parts of the landscape, can be accompanied by a thin layer of peat, (b) diffused distinction between the O horizon and mineral soil layers, especially when there is more bioturbation, and (c) cooler temperatures than adjacent non-forest soils (Starr and Vasander, 2021; Figure 2). Soil conditions, e.g., clay, sand, and water content, play a role in the distribution of both coniferous and deciduous tree stands. Particularly, SHF spruce and pine stands are differentiated by soil nutrient factors and soil depth (Økland, 1996).

[10] Soils deeper than 50 cm are rarely encountered.

Figure 2. Common soil profile encountered at Solhomfjell Forest Reserve

This particular profile is near the transition from spruce to pine forest. Organic layers (O horizon) mix loose and partly decayed organic matter with mineral particles. Boreo nemoral forests are usually deeper than 10 cm. The light-colored layer corresponds to the mineral layer, a zone of eluviation and leaching.

Soil sampling

For Paper III, soil samples were collected at the SFR across eight transects that represent edaphic gradients. In these, one hundred 1 m^2 plots have been established since 1988 and are monitored every five years to study vegetation responses to climate change (Økland and Eilertsen, 1993). A single soil sample was collected at the center of each plot (n=100) when possible. For Paper IV, soil was sampled at 31 sites across the Hvaler Archipelago in which NiN forest types have been identified (Eriksen, 2017). Site locations were identified by geographical coordinates reported in the latter study.

For both Papers III and IV soil samples were collected in late summer (August-September) of 2018, as follows: First, debris and living plants were manually removed to expose the soil. Second, a 50 ml Eppendorf tube was hammered down into the soil to the largest extent possible and then manually pulled out from the soil. Third, the tube was immediately capped and excess dirt was cleaned using a wipe. The plot/site number was labeled on the tube together with the date. In order to avoid cross-contamination, unique nitrile gloves and wipes were used at each plot/site. Tubes were stored individually in sealable plastic bags for transportation.

In order to stop the decomposition of organic matter by bacterial activity and remove excess water, samples were freeze-dried under vacuum. Soil samples were stored at −20 °C until DNA extraction, and registered in the DNA collection of the Natural History Museum, University of Oslo.

DNA extraction

Each soil sample was homogenized individually by mechanical crushing using ceramic beads. Subsequently, an amount was scooped and weighed. One gram was used for DNA extraction. DNA extraction was carried out for both Paper III and IV in seven rounds[11] and included two steps: (1) incubation pre-treatment with CTAB-buffer and chloroform to increase the separation of the organic phase from the aqueous phase; (2) the E.Z.N.A. soil DNA kit following the manufacturer's protocol (Omega Biotek). A detailed protocol for these steps can be found in the supplementary material of Paper III. Although this protocol mainly targets the extraction of extracellular DNA, the extraction of intracellular DNA cannot be out-ruled as mechanical crushing and heating steps may break plant cell walls and pollen membranes. DNA extracts were stored at 4 °C until barcode amplification. Further considerations for DNA extraction from soil can be found in Paper I.

Amplicon metabarcoding

Metabarcoding extends on the DNA barcoding concept by enabling the identification of multiple organisms from samples with a complex DNA

[11] A DNA extraction negative control, prepared replacing DNA by Milli-Q water, was included in each extraction round (n=7).

mixture or community. In plant barcoding, commonly used markers include the chloroplast regions trnL p6 loop, rbcL and matK, as well as the nuclear ribosomal DNA regions ITS and subregion ITS2. The same markers can be used for DNA metabarcoding of plant communities, however, studies often rely on trnL and the internal transcribed spacer (ITS2). The p6 loop of the chloroplast trnL intron represents a short and highly variable region that has been shown to work well even on samples with highly degraded DNA (Taberlet et al., 2007), and even ancient DNA (Polling et al., 2021). The *g* - *h* primers have better resolution for vascular plants while the *c* - *d* can yield identifications from gymnosperms to angiosperms with lesser resolution (Taberlet et al., 2007). The nuclear ribosomal Internal Transcribed Spacer (nrITS) region has been proposed as a potential barcode for land plants as high species resolution can be obtained (Kress, 2017). The nrITS region consists of two regions: nrITS1 located between 18S and 5.8S, and nrITS2 located between 5.8S and 26S rRNA genes (Chen et al., 2010). The latter has been shown to have higher resolution power and robust reference sequences (Chen et al., 2010), however, successful amplification depends on well-preserved DNA as the region length can vary from 350-500 bp across plant groups. Nevertheless, this nuclear barcode is shared across common DNA communities in the soil i.e., fungi and plants, and targeted amplification of either of these can be achieved by employing specific primers.

In order to harvest the power and counteract the limitations of these barcodes, the trnL (UAA) intron P6 loop (Paper III & IV) and the nuclear internal transcribed spacer (ITS2; Paper IV) barcodes were amplified by PCR using the *g* -*h* and UniPlant primers[12] (Moorhouse-Gann et al., 2018; Taberlet et al., 2007), respectively. Three technical replicates were prepared per sample. Forward and reverse primers were tagged with a unique 12 bp oligonucleotide on the 5' end (Fadrosh et al., 2014). Unique combinations of tagged primers were set up panels for each PCR reaction for 309 samples from SHF (100 samples with 3 PCR replicates each, 5 extractions blanks and 4 PCR negatives) and 97 samples from Hvaler (31 samples with 3 PCR replicates each, 2 extraction blanks and 2 PCR negatives). The PCR negatives had no DNA template and were placed on the 96th well position in each panel.

[12] This primer targets plants specifically.

PCR products were run on a 2% agarose gel and amplicon concentrations were measured via band intensity using ImageLab software (Bio-Rad, California, USA). The lowest concentration (µM) available for all PCR products and its relative volume was identified and the relative concentrations of the PCR products were adjusted to this same concentration. Amplicons from both SHF and Hvaler samples were pooled in one separated library for each barcode using a Biomek 4000 automated liquid handler (Beckman Coulter Life Sciences, Indianapolis, Indiana, USA). The library was cleaned using AMPure XP reagent beads (Beckman Coulter Life Sciences, Indianapolis, Indiana, USA). The length for all amplicons in the library was determined using a Fragment Analyzer (Agilent Technologies, Santa Clara, California, USA). Marker libraries were sequenced at the Norwegian Sequencing Center on an Illumina MiSeq v2 300 cycles (150 bp x 2) for the trnL (UUA) p6 loop fragments and an Illumina MiSeq v3 600 cycles (300 bp x 2) for the ITS fragments.

Sequence analysis and taxonomic identification

trnL amplicons were curated and analyzed using OBITools 2 (Boyer et al., 2016) following the *wolf* tutorial with adaptations for demultiplexing dual indexes from QUIME3 (Caporaso et al., 2010). ITS2 amplicons were analyzed using different bioinformatic scripts collated in METAPIPE (https://github.com/marlaux/METAPIPE_METAbarcoding_PIPEline), which were used to merge, dereplicate, cluster and assign taxa. Taxonomic assignments of trnL and ITS2 OTUs were carried out by matching to the arctic boreal (*arcborbryo*) reference sequence database for vascular plants and bryophytes (Alsos et al., 2020; Soininen et al., 2015; Sønstebø et al., 2010; Willerslev et al., 2014) and PLANTiTS (Banchi et al., 2020). When different sequences were identified with identical taxa names, a unique entry was retained and the read counts within plots/sites replicates were summed.

Taxa match to reference

Analyses from Papers III and IV are centered around matching DNA-based plant identifications to reference inventories carried through visual assessments (*taxa match* variable). In order to make sound comparisons, inventories were filtered to taxa with reference sequences for both trnL and

ITS2 separately. In addition, taxon names were changed to the highest resolution possible with these barcodes.

For Paper III, plant identifications from soil eDNA are compared to a data time-series of seven vegetation surveys over 30-years of each plot at SHF (Økland and Eilertsen, 1993). This dataset contains data on taxon presence/absence as well as subplot frequency (0– 16) of 157 taxa, including 69 vascular plants and 88 bryophytes. Calculations of *taxa match* to plot composition of each and overall years were made to elucidate the temporal signal of a soil sample.

For Paper IV, plant identifications from soil eDNA are compared to theoretical compositions of NiN forest types reported for each of the sampling sites in the Hvaler archipelago (Eriksen, 2017; Torma, 2022). The dataset contains a list of typical species (up to 280) present in a given forest type with abundance values provided on a semi-quantitative scale (0-6). In order to account for size differences between soil eDNA and NiN theoretical compositions, we calculated the variable *taxa match* as two times the number of identical names between compositions, divided by the sum of total taxa in soil eDNA and NiN theoretical compositions.

Downstream analysis

Papers III further investigates the similarities and differences between the composition registered in vegetation surveys and soil eDNA at the plot scale. Separate analyses are made for vascular plants and bryophytes, and for data subsets from plots located in environments dominated by either spruce or pine. The temporal resolution i.e, survey year, at which the two inventories were most comparable was evaluated by assessing mean values (with 95% confidence) of taxa match across all plots and correlations between taxa matching and dominant (repeated) vegetation over survey years. The spatial resolution was evaluated by categorizing plot taxa detected with soil eDNA by correspondence with the plot and overall visual inventory, as well as to the flora reported for SHF area. Similarly, the representation of vegetation surveyed in soil eDNA inventory was evaluated by categorizing plot taxa by correspondence or not to the plot inventory and to taxa without trnL reference sequence. Finally, the relationship between soil eDNA detection

and both plot taxa abundance and the number of read sequences was assessed by correlating these values from taxa present in both inventories and undetected taxa.

Paper IV makes use of diversity descriptors to show the potential of soil eDNA assessments in describing the main features of vegetation and the composition of given forest types. In addition, the main compositional gradients that describe the composition change from soil eDNA across sites are assessed by performing ordinations using different methods. These gradients are further correlated to gradients established by visual assessments of aboveground composition.

Further paper-specific downstream analyses with relevant methods, R packages and other software are detailed in the papers included in this thesis.

Results and discussion

This section outlines the main results from Papers I-IV and discusses these in context with the objectives and main challenges of both eDNA-based plant biodiversity assessments and traditional plant assessments. To provide a holistic and coherent perspective of the results and discussion across the papers, a synthesis is provided towards the end of this section.

Environmental and biodiversity assessments through eDNA analysis (Paper I)

Plant biodiversity assessments through the analysis of DNA from the environment is a cost-effective method to identify many taxa in parallel from many samples, including rare or hard-to-collect taxa (Taberlet et al., 2018). This educational book chapter provides a review of studies employing substrates from which plant DNA can be extracted to give an overview of the applications, power and limitations of plant eDNA assessments.

The overview shows that studies employing eDNA metabarcoding use airborne and bulk samples together with faecal, soil, sedimentary and water substrates for plant detection. Plant DNA present in airborne samples mainly originates from pollen, which can be transported over large distances, and therefore primarily signals the regional composition of anemophilous plants (Eaton et al., 2018; Johnson et al., 2019). Plant bulk samples can be naturally or artificially assembled, e.g., pollen from pollination vectors or collected leaves, and plants detected usually signal ecological networks (McFrederick and Rehan, 2016; Ritter et al., 2018). Similarly, faecal substrates, i.e., faeces, mucus, and saliva, provide a snapshot of vegetation implicated in trophic interactions and signal plant dietary items (Chua et al., 2021; Valentini, 2007). Moreover, soil and sediments are preferred substrates for plant assessments in both terrestrial and aquatic ecosystems as most plant DNA from active and dormant tissues ultimately deposits in these (Fahner et al., 2016). Plant DNA in soil and sediments mainly mirrors surrounding vegetation, however, the sampling depth of these substrates would indicate the strength of past signals (Rijal et al., 2021; Yoccoz et al., 2012). Although eDNA biodiversity assessments have proliferated in aquatic ecosystems, plant assessments have been widely overlooked compared to that of other organisms and particularly in marine

environments. However, several studies have shown that water samples collected in lakes along vertical strata signal both aquatic and terrestrial plants (Alsos et al., 2018; Drummond et al., 2021).

An analysis of the studies reviewed in the chapter shows that amplicon metabarcoding of trnL and ITS is commonly used across substrates for the detection of plant diversity, particularly that of vascular plants. eDNA is exposed in different ways to biotic and abiotic factors that influence DNA permanence and degradation (Nagler et al., 2018; Pietramellara et al., 2009; Stewart, 2019). Hence, careful consideration of these factors is needed when choosing a barcode given its length and other factors influenced by DNA taphonomy. Furthermore, the assessment of plant abundance through amplicon metabarcoding remains a thorny issue as the relation between biomass and read sequence counts is poorly understood (Beng and Corlett, 2020; Deiner et al., 2017). In addition, PCR cycles and errors may cause stochasticity in the read counts and false positives, respectively (Bellemain et al., 2010). PCR-free methods e.g. metagenomics through shotgun sequencing, promise to alleviate these biases, but the low availability of genomic references across the (Plant) Tree of Life are currently limiting the implementation of this method (Porter and Hajibabaei, 2018).

Taking into account the transportation of DNA through the environment by the detachment of plant parts from the main body, i.e., flowers, leaf debris, pollen, and seeds, is central to the derive sound conclusions as different local, regional, past and present biodiversity signatures could be mixed in eDNA substrates (Rodriguez-Ezpeleta et al., 2021). In this way, the investigation of the temporal and spatial signals present in eDNA samples is crucial to the advancement of eDNA assessments. Moreover, as DNA can remain in the environment after the organism has become locally extinct it is important to recall that eDNA-based detections reflect only the presence of organismal DNA and not of the actual organism (Rodriguez-Ezpeleta et al., 2021). In this way, eDNA-based plant detection provides information about the species pool of a region and can help to monitor the distribution of organisms over large areas.

In addition to providing a comprehensive overview of the state of the art in plant eDNA biodiversity assessments, the chapter provides a horizon scan for future developments in this field. It predicts an increased use of organisms as

eDNA biotic samplers, e.g., aquatic macroinvertebrates that feed both on aquatic vascular plants and debris leached to the environment. Gut analysis of these organisms can provide information about plants involved in biotic interactions, including surrounding terrestrial vegetation.

DNA from soil: considerations and study design (Paper II)

Plants rely on soil as a substrate for anchorage, which typically consists of a mixture of unconsolidated weathered minerals, organic matter, moisture, and air (Binkley and Fisher, 2019; Starr and Vasander, 2021). The surface layer of soils, termed the organic (O) horizon (Fox and Comerford, 1990), typically harbors growing roots and decomposing organic materials and simultaneously accumulates fallen debris, pollen particles, and seeds[13]. Therefore, this layer is particularly rich in plant DNA from the environment. This book chapter provides a comprehensive review of the literature on the state and composition of plant DNA present in soil, including an overview of pertinent considerations for sample collection in the field, sample processing, DNA extraction and amplification. The review shows that DNA from soil is present either in intact cells or free in the environment, i.e., posterior to cell rupture. In addition, intrinsic (e.g., GC content) and extrinsic properties (e.g., soil pH) underline the fate of DNA in soil environments (Pietramellara et al., 2009). As DNA can bind to soil particles and persist, soil can signal the "memory" from past diversity (Foucher et al., 2020).

Key considerations to ensure sound conclusions when designing a soil eDNA study are that at least 10 grams of soil should be sampled, either by scooping out the soil or drilling a core sampler, preferably in flat areas (Dopheide et al., 2019). As roots and leaf fragments are present in the soil, it is recommended to sieve these out from samples to avoid amplification bias towards these and obscuring of DNA signals from rare taxa. Furthermore, as many microorganisms present in soils are active decomposers of organic matter and therefore induce DNA fragmentation, it is recommended to aerate or freeze dry the soil to stop enzymatic activity (Nocker et al., 2012).

The chapter stresses that study aims should determine the DNA fraction that should be targeted, i.e., intracellular or extracellular, the extraction protocol,

[13] An infographic of this process is provided in this book chapter.

and the molecular barcodes to use. An analysis of the literature review shows that metabarcoding of trnL is best suited for soil eDNA as it targets a short region (20-80 bp) that can be amplified even in highly degraded samples (Taberlet et al., 2007). However, ITS2 should be employed when high species resolution is desired (Cheng et al., 2016). Nevertheless, when choosing appropriate barcodes for a soil DNA study it is important to consider the following questions: a) What is the desired taxonomic level of identification? b) What kind of reference library will be used to identify the target barcodes? c) What is the source of reference sequences? Are plant-specific primers available? d) Do the primers amplify the right part of the marker?

Despite an increase in the use of soil eDNA methods for the identification of microorganisms and macrofauna in recent years, plant assessments are broadly overlooked. In this context, this book chapter also aims to identify the main knowledge gaps from soil eDNA assessments. As with most eDNA samples and substrates, the spatial and temporal plant DNA signals present in soil and how these represent vegetation turnover over ecosystem types remains poorly known. This limits the use of soil eDNA in complete and reliable plant assessments that can guide biodiversity monitoring and mapping. As DNA from many vegetative and reproductive plant parts are present in soil, and each of these parts are differentially represented in terms of biomass in the soil, the relation between sequence read counts and biomass must be clarified to understand the power of soil eDNA to assess plant diversity quantitatively.

Plant biodiversity assessment through soil eDNA reflects temporal and local diversity (Paper III)

Soil eDNA assessments are promising means for plant identification and vegetation monitoring as the soil is rich in plant DNA (Kesanakurti et al., 2011). However, it is crucial to determine the temporal and spatial resolution of a soil eDNA assessment to interpret the utility of this approach in these applications. In other words, soil eDNA assessments detect plants from where and when? Furthermore, it is necessary to elucidate the relationship between plant DNA detection, aboveground plant biomass and read sequence counts to understand the power and limitations of soil eDNA for quantitative assessment of diversity. In order to answer these questions, this study

compares a time series of vegetation surveys in a boreal forest carried through visual assessments across one hundred 1 m^2 plots in a 30-year period to a single eDNA assessment of soil samples collected at the same plots on the last year of vegetation survey.

From the 65 vascular plants and 68 bryophytes recorded in the vegetation surveys, soil eDNA detected 53 vascular plants and 11 bryophytes. In addition, soil eDNA detected 51 regional plants that were not recorded in any of the vegetation surveys. Soil eDNA detected vascular and abundant plants much better than bryophytes and rare ones, respectively. A positive relation between plant biomass and sequence read counts was found. The vegetation detected by soil eDNA assessments mainly reflected the composition registered in the last vegetation survey and corresponded to the local species pool rather than to the specific plots, and included the identification of taxa recorded up to 30 years ago in the plots.

Altogether, the results of this study demonstrate that soil eDNA encapsulates mainly local and contemporary plant diversity and provides, in this way, a baseline for decision-making when designing soil eDNA assessments. This notion highlights the contribution of litter to the pool of plant DNA present in soil samples. In addition, this study shows how soil eDNA assessments can aid visual assessments to detect and unveil rare and overlooked plants, including those that persist below ground without above-ground parts. In this way, a combination of soil eDNA and visual assessments can yield complete plant biodiversity assessments at specific, local, and regional scales that are urgent for monitoring vegetation changes in response to climate change. Moreover, a single eDNA sample can be sufficient for the assessment of the main features of vegetation, but when a more detailed assessment is desired, i.e., detection of less abundant and/or rare taxa, collecting multiple eDNA soil samples is preferable.

Nevertheless, this study shows that soil eDNA plant assessments through trnL remain limited in assessing bryophyte diversity. Although the primers used in this study (*g* - *h*) mainly target vascular plants (Taberlet et al., 2007), many bryophytes for which reference sequences were available remained undetected. The low biomass of bryophytes in the forest floor may underline this result (Stefańska-Krzaczek et al., 2022), but further investigation is needed. While the results of this study are applicable to temperate and boreal

ecosystems where vegetation is (relatively) spatially homogenized, tropical soils may signal plant diversity at different spatial and temporal scales. Finally, this study shows how regional plant signals are embedded in a soil sample and recommends further studies to investigate the role of landscape in DNA transport and therefore its contributions to these signals.

Evaluating the Feasibility of Using Plant-Specific Metabarcoding to Assess Forest Types from Soil eDNA (Paper IV)

Vegetation surveys required for habitat typification purposes are limited to the season in which physiognomic characters required for habitat classification are visible and by the variability of habitat types detected amongst surveyors (Naas et al., 2023). This study addresses this challenge by incorporating an observer-free method that is accessible at all times, i.e., metabarcoding of trnL and ITS2 from soil samples, to identify habitat types through their theoretical composition. In addition to evaluating the method's feasibility, this study compares gradients in plant composition identified by soil eDNA assessments to those from vegetation surveys carried by 11 observers (Eriksen et al., 2019).

This study investigates the plant DNA signals of 31 soil samples collected across a gradient of six forest types in Hvaler archipelago. From these soil samples, 70 plant taxa were identified, mostly at the species level. More taxa were identified with trnL than ITS2. From the 254 plant taxa that can be theoretically expected across all assessed forest types and the 70 taxa detected from soil eDNA samples, 29 taxa match. Samples identified on average seven taxa that were not registered in the theoretical composition of any forest type.

After correcting for composition size differences, soil eDNA samples match 26% to 76% of the theoretical compositions in each forest type assessed in this study. Ordinations of soil eDNA compositions revealed two gradients, i.e., axes, without a clear pattern, and these correlated poorly to gradients previously identified by visual assessments.

Altogether, this study shows that the combined use of ITS2 and trnL barcodes provides plant compositional data that can aid the characterization of the

main features of forest types. While there is an incomplete match between soil eDNA and theoretical compositions of a forest type, a complete match cannot be expected as field assessments often detect a subset of this theoretical composition. Nevertheless, ITS2 amplicons yielded fewer identifications than expected. Amplification of the ITS2 flanking regions by the primers used in this study may cause difficulties in initiating local alignments to reference sequences. Although most soil eDNA-based identifications were at the species level, lack of species-resolution data can hinder the determination of habitat types as these are mainly based on specific species.

Furthermore, this study highlights the power of this method to uncover other ecological structures of potential importance. Although the lack of correspondence between the gradients in plant composition identified with both proxies can be underlined by the lack of detection of indicator species, soil eDNA detected some taxa that are not registered in any of the assessed forest types and these can drive the differences observed. More robust reference libraries and the use of barcodes targeting bryophytes may clarify these results and their relations to gradients.

Synthesis

The growth of eDNA metabarcoding studies over the last decades has shown the potential of this tool to assess, monitor and manage biodiversity. The results from the four Papers included in this thesis contribute to developing and advancing the applications and power of this tool in botany, particularly for the assessment of plant diversity from soil eDNA applied to monitoring and habitat mapping. Papers I and II show through reviews of the state of the art that the amplification of trnL and ITS2 barcodes is particularly useful to analyze degraded eDNA samples and to yield plant identifications at species-level resolution. These results informed the choice of barcode markers for Paper III and IV, where it is demonstrated that these barcodes can potentially identify more than 3,000 plants of the Fennoscandian flora. Regardless, Paper I and II emphasizes how limited are the the conclusions from eDNA-based and soil eDNA plant assessments, respectively, when the spatial and temporal plant signals present in these samples and substrates are not known. Paper III fills this knowledge gap by comparing a time series of

vegetation surveys to one single soil eDNA assessment and shows that soil eDNA samples mainly encapsule signals from current and local diversity, and to a lesser extent include signals from past and regional diversity. In this way, these results provide the bases to foresee the applications of soil eDNA metabarcoding explored in Paper IV, i.e., characterisation of habitat types for mapping ecosystems, and simultaneously helping to delineate the power and limitations of this tool for this purpose. While Paper III employs a spatial and temporal framework to compare diversity detected from soil eDNA, the correspondence to older and wider signals cannot be ruled out. Regardless, the calibration in time of plant signals presented in Paper III corroborates the soil "memory" capacity described in Paper II. In addition, the temporal and spatial signals detected in Paper III should only guide soil sampling protocols and study designs in boreal and temperate environments. Plant signals present in soil samples from tropical environments remain widely unknown and thus its applications to assess and monitor diversity are limited.

Moreover, the review from Paper I highlights how eDNA metabarcoding remains limited in quantitative assessment of plant diversity as the relationships between plant biomass, detection and read sequence counts remain poorly known. As shown in Paper II, this challenge is particularly exacerbated in soil eDNA samples as the biomass from plant vegetative and reproductive parts that are deposited in the soil contributes differentially to the DNA present in the soil and therefore some bias towards the amplification of certain plant signals can be expected. Paper III addresses this challenge by exploring the relevance of aboveground plant biomass in soil eDNA plant detection and by correlating this to read sequence counts and shows that, as expected, soil eDNA assessments detect better plants with higher biomass. Nevertheless, Papers III shows that rare plants can also be detected and this result is of particular relevance when deciding how many soil samples should be collected in reference to the study aims. In spite of that, some plants remained undetected in both Papers III and IV, including those plants with medium to high aboveground biomass in Paper III. The stochasticity of these findings raises more questions, particularly on how biomass differences between root and shoot at different life-history stages can potentially underlie detection patterns.

As eDNA-based assessments become more accessible and popular, studies comparing this tool to traditional methods have increased in parallel.

However, a common trend amongst these studies is to place these plant identification proxies in a competition, and as consequence the power and advantages of traditional methods were often neglected. The experimental papers in this thesis (Papers III-IV), which are based on these comparisons, do not follow this trend. Instead, the results from these empirical studies attest to the complementarity of traditional and soil eDNA assessments to reliably and completely inventory plant biodiversity. For instance, these papers together demonstrate that soil eDNA can identify a large number of plant taxa from the Norwegian flora, including liverworts, mosses, ferns, graminoids, forbs, shrubs, trees, and hidden or overlooked diversity. Moreover, Paper IV shows how soil eDNA metabarcoding can contribute to characterize and establish new plant compositional gradients but traditional assessments are still required to confidently detect habitat types through their theoretical composition. With the increasing completeness of reference libraries and the refinement of molecular tools for plant identification, some of the challenges and limitations of the papers results will be lessened. This will contribute to empower and accelerate fast and reliable assessment of nature that is so needed at this time.

In light of the results shown in this thesis and the contributions made to the field of molecular plant assessments and its applications, new challenges and limitations are yet to be overcome. For instance, Paper III shows a calibration in space and time of soil eDNA plant signals, however, these scales are tracked simultaneously, i.e., a plant is present both at time t_x and at space s_x and therefore the contribution of each cannot be assessed. Given this limitation, it is important to conduct studies in controlled settings that allow for tracking of plant signals, thereby facilitating the ability to trace both spatial and temporal signals. Moreover, the inclusion of state-of-the-art methodological approaches for the analysis of eDNA samples and substrates e.g. metagenomics through shotgun sequencing, would require revisiting the questions and aims from all papers in this thesis to delineate the power and limitations of these tools. Particularly, Paper IV indicates that plant signals present in soil eDNA samples portray unprecedented compositional gradients, which are yet to be studied in depth to comprehend the environmental features that underline them.

Future directions

Twenty years after DNA barcoding emerged, the way in which biodiversity is examined and understood has significantly changed. Paper I has evidenced that environmental DNA metabarcoding is now an integral part of the ecologists' toolkit, and alongside its development and applications new exciting horizons are in sight. The recovery of genetic information below the species level i.e., populations, from eDNA samples has opened new avenues to monitor diseases, migrations, and population dynamics (Adams et al., 2019; Sigsgaard et al., 2016). Although this approach is still in its infancy and the available evidence is exclusive to marine and aquatic environments e.g., (Adams et al., 2022; Andres et al., 2021; Sigsgaard et al., 2016), its application in the botanical world promises to aid the estimation of abundance via haplotype variability and the monitoring of plant introgressions and hybridizations within invasive species (Johnson et al., 2023). Similarly, eDNA biotic samplers of plant diversity are poorly explored but evidence from other organisms (e.g. shrimps) shows how these can do the job as water or sediment substrates would (Mariani et al., 2019).

Moreover, the inclusion of other tools that can seize more informative regions for species identification i.e., metagenomics, target capture, can pave the way to a sound understanding of plant assemblages in environmental samples by bringing to light new community compositions (Foster et al., 2021). Similarly, the incorporation of RNA assessment from environmental samples i.e., eRNA, in monitoring surveys can draw back the curtain on the actual presence of organisms in a site together with life-history stages, sexes, or even specific phenotypes within a species that is so needed to design sound conservation strategies (Yates et al., 2021).

Albeit the increasing use of eDNA for biodiversity assessment and monitoring, eDNA research has a lot of ground to cover. A recent study reviewed 4114 eDNA studies from which only 558 (13%) were devoted to the detection of plant species or communities (Banerjee et al., 2022). Although this trend is worrying as it reflects academic plant blindness, it also highlights the enormous potential of eDNA to uncover the hidden treasures of the plant world. Artificial intelligence can help to identify knowledge gaps and priorities to chart a course toward a greener eDNA future.

Towards complete and sound plant assessments from soil eDNA

Papers II to IV of this thesis shed light on the potential of soil eDNA metabarcoding to assess and monitor plant biodiversity. However, the conclusions of these papers are only applicable to temperate and boreal ecosystems. Compared to these, tropical ecosystems harbor relatively higher diversity distributed in space more heterogenially (Stein et al., 2014). Therefore, the aims of Papers II and IV should be revisited in context with tropical environments to thereby derive sound conclusions from soil eDNA assessments in these ecosystems. Nevertheless, the future of this tool lies within the integration of state-of-the-art technologies that can both strengthen its applications and overcome its limitations. For example, target capture of informative genes in soil eDNA samples may produce richer plant inventories that can complement monitoring protocols established in Paper III and that needed for the identification of ecosystem types as shown in Paper IV. In addition, machine learning algorithms, e.g., random forests (Breiman, 2001), can aid the detection of these types from inventories of both soil eDNA and vegetation surveys (Cutler et al., 2007). Moreover, the inclusion of ecological knowledge from Indigenous Peoples and Local Communities (IPLC) can greatly enhance the accuracy and applicability of these assessments (Brondízio, 2019) and complement the temporal signals detected. Today, a future where a synthesis of the state of terrestrial ecosystems is automatized by (soil) eDNA assessments, remote sensing and artificial intelligence, and supported by IPLC knowledge, does not seem far.

Conclusions

This thesis breaks ground (literally and metaphorically) and demonstrates:

- The enormous potential of airborne and bulk samples together with faecal, soil, sedimentary and water substrates to identify, monitor and protect plant diversity across a variety of applications.
- Knowledge gaps regarding the temporal and spatial plant signals present in samples and substrates, the relationship between biomass and detection, including that of alive organisms are constraining the applications and conclusions that can be derived from soil eDNA assessments.
- Key considerations when designing a soil eDNA study are: study aims together with the intrinsic and extrinsic properties of soil DNA will determine the DNA fraction that should be targeted, extraction protocols and barcodes to be used.
- Soil eDNA samples harbor signals from local, regional, past and present plant diversity.
- Plant detection with soil eDNA is positively correlated with aboveground plant biomass and read sequence counts.
- Vegetation surveys through morphological identification of plants are needed to achieve complete plant biodiversity assessments.
- Soil eDNA samples mirror the main components of plant composition across environmental gradients.
- Plant composition determined from soil eDNA assessments can partially aid the characterisation of ecosystem types by their theoretical composition.
- Plant composition gradients identified through soil eDNA metabarcoding correspond poorly to those detected by traditional vegetation surveys.
- Soil eDNA assessments can elucidate plant composition gradients.

Altogether, this thesis highlights how DNA-based identification can contribute to meet the biodiversity challenges of the twenty-first century.

References

Adams, C.I.M., Hepburn, C., Jeunen, G., Cross, H., Taylor, H.R., Gemmell, N.J., Bunce, M., Knapp, M., 2022. EnvironmentalDNA reflects common haplotypic variation. Environmental DNA. doi:10.1002/edn3.352

Adams, C.I.M., Knapp, M., Gemmell, N.J., Jeunen, G.-J., Bunce, M., Lamare, M.D., Taylor, H.R., 2019. Beyond biodiversity: can environmental DNA (edna) cut it as a population genetics tool? Genes (Basel) 10. doi:10.3390/genes10030192

Alsos, I.G., Lammers, Y., Yoccoz, N.G., Jørgensen, T., Sjögren, P., Gielly, L., Edwards, M.E., 2018. Plant DNA metabarcoding of lake sediments: how does it represent the contemporary vegetation. PLoS ONE 13, e0195403. doi:10.1371/journal.pone.0195403

Alsos, I.G., Lavergne, S., Merkel, M.K.F., Boleda, M., Lammers, Y., Alberti, A., Pouchon, C., Denoeud, F., Pitelkova, I., Puşcaş, M., Roquet, C., Hurdu, B.-I., Thuiller, W., Zimmermann, N.E., Hollingsworth, P.M., Coissac, E., 2020. The treasure vault can be opened: large-scale genome skimming works well using herbarium and silica gel dried material. Plants 9, 432. doi:10.3390/plants9040432

Anderson-Carpenter, L.L., McLachlan, J.S., Jackson, S.T., Kuch, M., Lumibao, C.Y., Poinar, H.N., 2011. Ancient DNA from lake sediments: bridging the gap between paleoecology and genetics. BMC Evol. Biol. 11, 30. doi:10.1186/1471-2148-11-30

Andres, K.J., Sethi, S.A., Lodge, D.M., Andrés, J., 2021. Nuclear eDNA estimates population allele frequencies and abundance in experimental mesocosms and field samples. Mol. Ecol. 30, 685–697. doi:10.1111/mec.15765

Antonelli, A., Fry, C., Smith, R.J., Simmonds, M.S.J., Kersey, P.J., Pritchard, H.W., Abbo, M.S., Acedo, C., Adams, J., Ainsworth, A.M., Allkin, B., Annecke, W., Bachman, S.P., Bacon, K., Bárrios, S., Barstow, C., Battison, A., Bell, E., Bensusan, K., Bidartondo, M.I., et al., 2020. State of the World's Plants and Fungi 2020. Royal Botanic Gardens, Kew. doi:10.34885/172

Banchi, E., Ametrano, C.G., Greco, S., Stanković, D., Muggia, L., Pallavicini, A., 2020. PLANiTS: a curated sequence reference dataset for plant ITS DNA metabarcoding. Database (Oxford) 2020. doi:10.1093/database/baz155

Banerjee, P., Stewart, K.A., Dey, G., Antognazza, C.M., Sharma, R.K., Maity, J.P., Saha, S., Doi, H., de Vere, N., Chan, M.W.Y., Lin, P.-Y., Chao, H.-C., Chen, C.-Y., 2022. Environmental DNA analysis as an emerging non-destructive method for plant biodiversity monitoring: a review. AoB Plants 14, plac031. doi:10.1093/aobpla/plac031

Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., Kauserud, H., 2010. ITS as an environmental DNA barcode for fungi: an in silico

approach reveals potential PCR biases. BMC Microbiol. 10, 189. doi:10.1186/1471-2180-10-189
Beng, K.C., Corlett, R.T., 2020. Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. Biodivers. Conserv. 29, 2089–2121. doi:10.1007/s10531-020-01980-0
Binkley, D., Fisher, R.F., 2019. Ecology and management of forest soils. 5th edition. Wiley-Blackwell, New Jersey, USA. doi:10.1002/9781119455745
Blondel, J., 2006. The 'Design' of Mediterranean Landscapes: A Millennial Story of Humans and Ecological Systems during the Historic Period. Hum. Ecol. 34, 713–729. doi:10.1007/s10745-006-9030-4
Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W., de Bruyn, M., 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends Ecol. Evol. 29, 358–367. doi:10.1016/j.tree.2014.04.003
Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., Coissac, E., 2016. obitools: a unix-inspired software package for DNA metabarcoding. Mol. Ecol. Resour. 16, 176–182. doi:10.1111/1755-0998.12428
Breiman, L., 2001. Random Forests. Springer Science and Business Media LLC. doi:10.1023/a:1010933404324
Brondízio, E., 2019. Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Zenodo. doi:10.5281/zenodo.3831674
Brummitt, N., Araújo, A.C., Harris, T., 2020. Areas of plant diversity—What do we know? Plants, People, Planet. doi:10.1002/ppp3.10110
Brunbjerg, A.K., Bruun, H.H., Dalby, L., Fløjgaard, C., Frøslev, T.G., Høye, T.T., Goldberg, I., Læssøe, T., Hansen, M.D.D., Brøndum, L., Skipper, L., Fog, K., Ejrnæs, R., 2018. Vascular plant species richness and bioindication predict multi-taxon species richness. Methods Ecol. Evol. 9, 2372–2382. doi:10.1111/2041-210X.13087
Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336. doi:10.1038/nmeth.f.303
Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., Narwani, A., Mace, G.M., Tilman, D., Wardle, D.A., Kinzig, A.P., Daily, G.C., Loreau, M., Grace, J.B., Larigauderie, A., Srivastava, D.S., Naeem, S., 2012. Biodiversity loss and its impact on humanity. Nature 486, 59–67. doi:10.1038/nature11148
CBOL Plant Working Group, 2009. A DNA barcode for land plants. Proc Natl Acad Sci USA 106, 12794–12797. doi:10.1073/pnas.0905845106
Ceballos, G., Ehrlich, P.R., Barnosky, A.D., García, A., Pringle, R.M., Palmer,

T.M., 2015. Accelerated modern human-induced species losses: Entering the sixth mass extinction. Sci. Adv. 1, e1400253. doi:10.1126/sciadv.1400253
Cheek, M., Nic Lughadha, E., Kirk, P., Lindon, H., Carretero, J., Looney, B., Douglas, B., Haelewaters, D., Gaya, E., Llewellyn, T., Ainsworth, A.M., Gafforov, Y., Hyde, K., Crous, P., Hughes, M., Walker, B.E., Campostrini Forzza, R., Wong, K.M., Niskanen, T., 2020. New scientific discoveries: plants and fungi. Plants, People, Planet 2, 371–388. doi:10.1002/ppp3.10148
Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., Zhou, S., 2016. Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Mol. Ecol. Resour. 16, 138–149. doi:10.1111/1755-0998.12438
Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y., Leon, C., 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE 5, e8613. doi:10.1371/journal.pone.0008613
Chua, P.Y.S., Lammers, Y., Menoni, E., Ekrem, T., Bohmann, K., Boessenkool, S., Alsos, I.G., 2021. Molecular dietary analyses of western capercaillies (*Tetrao urogallus*) reveal a diverse diet. Environmental DNA 3, 1156–1171. doi:10.1002/edn3.237
Corlett, R.T., 2020. Safeguarding our future by protecting biodiversity. Plant Diversity 42, 221–228. doi:10.1016/j.pld.2020.04.002
Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792. doi:10.1890/07-0539.1
Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N., Pfrender, M.E., Bernatchez, L., 2017. Environmental DNA metabarcoding: transforming how we survey animal and plant communities. Mol. Ecol. 26, 5872–5895. doi:10.1111/mec.14350
Díaz, S., Cabido, M., 2001. Vive la différence: plant functional diversity matters to ecosystem processes. Trends Ecol. Evol. 16, 646–655. doi:10.1016/S0169-5347(01)02283-2
Dopheide, A., Xie, D., Buckley, T.R., Drummond, A.J., Newcomb, R.D., 2019. Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity. Methods Ecol. Evol. 10, 120–133. doi:10.1111/2041-210X.13086
Drummond, J.A., Larson, E.R., Li, Y., Lodge, D.M., Gantz, C.A., Pfrender, M.E., Renshaw, M.A., Correa, A.M.S., Egan, S.P., 2021. Diversity metrics are robust to differences in sampling location and depth for environmental DNA of plants in small temperate lakes. Front. Environ. Sci. 9, 617924. doi:10.3389/fenvs.2021.617924
Eaton, S., Zúñiga, C., Czyzewski, J., Ellis, C., Genney, D.R., Haydon, D., Mirzai, N., Yahr, R., 2018. A method for the direct detection of airborne dispersal

in lichens. Mol. Ecol. Resour. 18, 240–250. doi:10.1111/1755-0998.12731
Ellis, E.C., Ramankutty, N., 2008. Putting people in the map: anthropogenic biomes of the world. Front. Ecol. Environ. 6, 439–447. doi:10.1890/070062
Engel, M.S., Ceríaco, L.M.P., Daniel, G.M., Dellapé, P.M., Löbl, I., Marinov, M., Reis, R.E., Young, M.T., Dubois, A., Agarwal, I., Lehmann A., P., Alvarado, M., Alvarez, N., Andreone, F., Araujo-Vieira, K., Ascher, J.S., Baêta, D., Baldo, D., Bandeira, S.A., Barden, P., Zacharie, C.K., 2021. The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. Zool. J. Linn. Soc. 193, 381–387. doi:10.1093/zoolinnean/zlab072
Eriksen, E.L., Ullerud, H.A., Halvorsen, R., Aune, S., Bratli, H., Horvath, P., Volden, I.K., Wollan, A.K., Bryn, A., 2019. Point of view: error estimation in field assignment of land-cover types. Phyto. 49, 135–148. doi:10.1127/phyto/2018/0293
Eriksen, E.L., 2017. Point of View: Uncertainty in Nature Type Identification (Undergraduate thesis). University of Oslo.
Fadrosh, D.W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R.M., Ravel, J., 2014. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. Microbiome 2, 6. doi:10.1186/2049-2618-2-6
Fahner, N.A., Shokralla, S., Baird, D.J., Hajibabaei, M., 2016. Large-scale monitoring of plants through environmental DNA metabarcoding of soil: recovery, resolution, and annotation of four DNA markers. PLoS ONE 11, e0157505. doi:10.1371/journal.pone.0157505
Fazekas, A.J., Kuzmina, M.L., Newmaster, S.G., Hollingsworth, P.M., 2012. DNA barcoding methods for land plants. Methods Mol. Biol. 858, 223–252. doi:10.1007/978-1-61779-591-6_11
Fiz-Palacios, O., Schneider, H., Heinrichs, J., Savolainen, V., 2011. Diversification of land plants: insights from a family-level phylogenetic analysis. BMC Evol. Biol. 11, 341. doi:10.1186/1471-2148-11-341
Foster, N.R., Gillanders, B.M., Jones, A.R., Young, J.M., Waycott, M., 2020. A muddy time capsule: using sediment environmental DNA for the long-term monitoring of coastal vegetated ecosystems. Mar. Freshwater Res. 71, 869. doi:10.1071/MF19175
Foster, N.R., van Dijk, K., Biffin, E., Young, J.M., Thomson, V., Gillanders, B.M., Jones, A., Waycott, M., 2021. A multi-gene region targeted capture approach to detect plant DNA in environmental samples: A case study from coastal environments. BioRxiv. doi:10.1101/2021.07.03.450983
Foucher, A., Evrard, O., Ficetola, G.F., Gielly, L., Poulain, J., Giguet-Covex, C., Laceby, J.P., Salvador-Blanes, S., Cerdan, O., Poulenard, J., 2020. Persistence of environmental DNA in cultivated soils: implication of this memory effect for reconstructing the dynamics of land use and cover changes. Sci. Rep. 10, 10502. doi:10.1038/s41598-020-67452-1

Fox, T.R., Comerford, N.B., 1990. Low-Molecular-Weight Organic Acids in Selected Forest Soils of the Southeastern USA. Soil Sci. Soc. Am. J. 54, NP-NP. doi:10.2136/sssaj1990.03615995005400040051x

García-Robledo, C., Erickson, D.L., Staines, C.L., Erwin, T.L., Kress, W.J., 2013. Tropical plant-herbivore networks: reconstructing species interactions using DNA barcodes. PLoS ONE 8, e52967. doi:10.1371/journal.pone.0052967

Govaerts, R., Nic Lughadha, E., Black, N., Turner, R., Paton, A., 2021. The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. Sci. Data 8, 215. doi:10.1038/s41597-021-00997-6

Hagen, J.B., 2001. History of plant ecology, in: John Wiley & Sons, Ltd (Ed.), ELS. Wiley. doi:10.1002/9780470015902.a0003288.pub2

Halvorsen, R., Skarpaas, O., Bryn, A., Bratli, H., Erikstad, L., Simensen, T., Lieungh, E., 2020. Towards a systematics of ecodiversity: the EcoSyst framework. Global Ecol. Biogeogr. 29, 1887–1906. doi:10.1111/geb.13164

Hammond, P.M., 1992. Species inventory, in: Groombridge, B. (Ed.), Global Biodiversity: Status of the Earth's Living Resources. Springer, Houten, pp. 17–39.

Hartvig, I., Kosawang, C., Kjær, E.D., Nielsen, L.R., 2021. Detecting rare terrestrial orchids and associated plant communities from soil samples with eDNA methods. Biodivers. Conserv. 30, 3879–3901. doi:10.1007/s10531-021-02279-4

Hebert, P.D.N., Cywinska, A., Ball, S.L., de Waard, J.R., 2003. Biological identifications through DNA barcodes. Proc R Soc Lond B 270, 313–322.

Hoskins, A.J., Harwood, T.D., Ware, C., Williams, K.J., Perry, J.J., Ota, N., Croft, J.R., Yeates, D.K., Jetz, W., Golebiewski, M., Purvis, A., Robertson, T., Ferrier, S., 2020. BILBI: Supporting global biodiversity assessment through high-resolution macroecological modelling. Environmental Modelling & Software 132, 104806. doi:10.1016/j.envsoft.2020.104806

IPBES, 2019. Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Zenodo. doi:10.5281/zenodo.3831673

Johnson, M.D., Cox, R.D., Barnes, M.A., 2019. The detection of a non-anemophilous plant species using airborne eDNA. PLoS ONE 14, e0225262. doi:10.1371/journal.pone.0225262

Johnson, M.D., Freeland, J.R., Parducci, L., Evans, D.M., Meyer, R.S., Molano-Flores, B., Davis, M.A., 2023. Environmental DNA as an emerging tool in botanical research. Am. J. Bot. doi:10.1002/ajb2.16120

Johnson, R.N., Wilson-Wilde, L., Linacre, A., 2014. Current and future directions of DNA in wildlife forensic science. Forensic Sci. Int. Genet. 10, 1–11. doi:10.1016/j.fsigen.2013.12.007

Jones, K.R., Klein, C.J., Halpern, B.S., Venter, O., Grantham, H., Kuempel, C.D.,

Shumway, N., Friedlander, A.M., Possingham, H.P., Watson, J.E.M., 2018. The location and protection status of earth's diminishing marine wilderness. Curr. Biol. 28, 2506-2512.e3. doi:10.1016/j.cub.2018.06.010

Keddy, P.A., 2005. Milestones in ecological thought — A canon for plant ecology. Journal of Vegetation Science 16, 145–150. doi:10.1111/j.1654-1103.2005.tb02349.x

Kesanakurti, P.R., Fazekas, A.J., Burgess, K.S., Percy, D.M., Newmaster, S.G., Graham, S.W., Barrett, S.C.H., Hajibabaei, M., Husband, B.C., 2011. Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. Mol. Ecol. 20, 1289–1302. doi:10.1111/j.1365-294X.2010.04989.x

Kress, W.J., Erickson, D.L., Jones, F.A., Swenson, N.G., Perez, R., Sanjur, O., Bermingham, E., 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc Natl Acad Sci USA 106, 18621–18626. doi:10.1073/pnas.0909820106

Kress, W.J., 2017. Plant DNA barcodes: Applications today and in the future. J. Syst. Evol. 55, 291–307. doi:10.1111/jse.12254

Lughadha, E.N., Govaerts, R., Belyaeva, I., Black, N., Lindon, H., Allkin, R., Magill, R.E., Nicolson, N., 2016. Counting counts: revised estimates of numbers of accepted species of flowering plants, seed plants, vascular plants and land plants with a review of other recent estimates. Phytotaxa 272, 82. doi:10.11646/phytotaxa.272.1.5

Machado Nunes Romeiro, J., Eid, T., Antón-Fernández, C., Kangas, A., Trømborg, E., 2022. Natural disturbances risks in European Boreal and Temperate forests and their links to climate change – A review of modelling approaches. Forest Ecology and Management 509, 120071. doi:10.1016/j.foreco.2022.120071

Mariani, S., Baillie, C., Colosimo, G., Riesgo, A., 2019. Sponges as natural environmental DNA samplers. Curr. Biol. 29, R401–R402. doi:10.1016/j.cub.2019.04.031

McFrederick, Q.S., Rehan, S.M., 2016. Characterization of pollen and bacterial community composition in brood provisions of a small carpenter bee. Mol. Ecol. 25, 2302–2311. doi:10.1111/mec.13608

Meehan, M.C., Ban, N.C., Devillers, R., Singh, G.G., Claudet, J., 2020. How far have we come? A review of MPA network performance indicators in reaching qualitative elements of Aichi Target 11. Conserv. Lett. 13. doi:10.1111/conl.12746

Moen, A., 1999. Vegetasjon, Nasjonalatlas for Norge. Norges geografiske oppmåling, Hønefoss.

Molina-Venegas, R., Rodríguez, M.Á., Pardo-de-Santayana, M., Mabberley, D.J., 2021. A global database of plant services for humankind. PLoS ONE 16, e0253069. doi:10.1371/journal.pone.0253069

Moorhouse-Gann, R.J., Dunn, J.C., de Vere, N., Goder, M., Cole, N., Hipperson, H., Symondson, W.O.C., 2018. New universal ITS2 primers for

high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. Sci. Rep. 8, 8542. doi:10.1038/s41598-018-26648-2

Moura, M.R., Costa, H.C., Peixoto, M.A., Carvalho, A.L.G., Santana, D.J., Vasconcelos, H.L., 2018. Geographical and socioeconomic determinants of species discovery trends in a biodiversity hotspot. Biol. Conserv 220, 237–244. doi:10.1016/j.biocon.2018.01.024

Naas, A.E., Halvorsen, R., Horvath, P., Wollan, A.K., Bratli, H., Brynildsrud, K., Finne, E.A., Keetz, L.T., Lieungh, E., Olson, C., Simensen, T., Skarpaas, O., Tandstad, H.R., Torma, M., Værland, E.S., Bryn, A., 2023. What explains inconsistencies in field-based ecosystem mapping? Appl. Veg. Sci. doi:10.1111/avsc.12715

Nagler, M., Insam, H., Pietramellara, G., Ascher-Jenull, J., 2018. Extracellular DNA in natural environments: features, relevance and applications. Appl. Microbiol. Biotechnol. 102, 6343–6356. doi:10.1007/s00253-018-9120-4

Nocker, A., Fernández, P.S., Montijn, R., Schuren, F., 2012. Effect of air drying on bacterial viability: A multiparameter viability assessment. J. Microbiol. Methods 90, 86–95. doi:10.1016/j.mimet.2012.04.015

Økland, R., 1995. Boreal coniferous forest vegetation in the Solhomfjell area, S Norway: structure, dynamics and change, with particular reference to effects of long distance airborne pollution. Sommerfeltia 1995 (Monograph). Botanical garden and Museum, University of Oslo.

Økland, R.H., Eilertsen, O., 1993. Corticolous and lignicolous Pyrenomycetes s.lat. (Ascomycetes) on Salix along a mid-Scandinavian transect. Sommerfeltia. Botanical garden and museum, University of Oslo.

Økland, T., 1996. Vegetation-environment relationships of boreal spruce forests in ten monitoring reference areas in Norway. Sommerfeltia 22, 1–355. doi:10.2478/som-1996-0001

Petit, R.J., Vendramin, G.G., 2007. Plant phylogeography based on organelle genes: an introduction, in: Weiss, S., Ferrand, N. (Eds.), Phylogeography of Southern European Refugia. Springer Netherlands, Dordrecht, pp. 23–97. doi:10.1007/1-4020-4904-8_2

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M.T., Guerri, G., Nannipieri, P., 2009. Extracellular DNA in soil and sediment: fate and ecological relevance. Biol. Fertil. Soils 45, 219–235. doi:10.1007/s00374-008-0345-8

Polling, M., ter Schure, A.T.M., van Geel, B., van Bokhoven, T., Boessenkool, S., MacKay, G., Langeveld, B.W., Ariza, M., van der Plicht, H., Protopopov, A.V., Tikhonov, A., de Boer, H., Gravendeel, B., 2021. Multiproxy analysis of permafrost preserved faeces provides an unprecedented insight into the diets and habitats of extinct and extant megafauna. Quat. Sci. Rev. 267, 107084. doi:10.1016/j.quascirev.2021.107084

Porter, T.M., Hajibabaei, M., 2018. Scaling up: A guide to high-throughput

genomic approaches for biodiversity analysis. Mol. Ecol. 27, 313–338. doi:10.1111/mec.14478

Raclariu, A.C., Tebrencu, C.E., Ichim, M.C., Ciuperca, O.T., Brysting, A.K., de Boer, H.J., 2017. What's in the box? Authentication of *Echinacea* herbal products using DNA metabarcoding and HPTLC. bioRxiv 202721.

Rijal, D.P., Heintzman, P.D., Lammers, Y., Yoccoz, N.G., Lorberau, K.E., Pitelkova, I., Goslar, T., Murguzur, F.J.A., Salonen, J.S., Helmens, K.F., Bakke, J., Edwards, M.E., Alm, T., Bråthen, K.A., Brown, A.G., Alsos, I.G., 2021. Sedimentary ancient DNA shows terrestrial plant richness continuously increased over the Holocene in northern Fennoscandia. Sci. Adv. 7, eabf9557. doi:10.1126/sciadv.abf9557

Ritter, C.D., Zizka, A., Roger, F., Tuomisto, H., Barnes, C., Nilsson, R.H., Antonelli, A., 2018. High-throughput metabarcoding reveals the effect of physicochemical soil properties on soil and litter biodiversity and community turnover across Amazonia. doi:10.7287/peerj.preprints.27012

Robles, P., Quesada, V., 2021. Organelle genetics in plants. Int. J. Mol. Sci. 22. doi:10.3390/ijms22042104

Rodriguez-Ezpeleta, N., Morissette, O., Bean, C.W., Manu, S., Banerjee, P., Lacoursière-Roussel, A., Beng, K.C., Alter, S.E., Roger, F., Holman, L.E., Stewart, K.A., Monaghan, M.T., Mauvisseau, Q., Mirimin, L., Wangensteen, O.S., Antognazza, C.M., Helyar, S.J., de Boer, H., Monchamp, M.-E., Nijland, R., Deiner, K., 2021. Trade-offs between reducing complex terminology and producing accurate interpretations from environmental DNA: comment on "Environmental DNA: what's behind the term?" by Pawlowski et al., (2020). Mol. Ecol. 30, 4601–4605. doi:10.1111/mec.15942

Saah, D., Tenneson, K., Matin, M., Uddin, K., Cutter, P., Poortinga, A., Nguyen, Q.H., Patterson, M., Johnson, G., Markert, K., Flores, A., Anderson, E., Weigel, A., Ellenberg, W.L., Bhargava, R., Aekakkararungroj, A., Bhandari, B., Khanal, N., Housman, I.W., Potapov, P., Chishtie, F., 2019. Land cover mapping in data scarce environments: challenges and opportunities. Front. Environ. Sci. 7. doi:10.3389/fenvs.2019.00150

Shepherd, E., Milner-Gulland, E.J., Knight, A.T., Ling, M.A., Darrah, S., Soesbergen, A., Burgess, N.D., 2016. Status and trends in global ecosystem services and natural capital: assessing progress toward aichi biodiversity target 14. Conserv. Lett. 9, 429–437. doi:10.1111/conl.12320

Sigsgaard, E.E., Nielsen, I.B., Bach, S.S., Lorenzen, E.D., Robinson, D.P., Knudsen, S.W., Pedersen, M.W., Jaidah, M.A., Orlando, L., Willerslev, E., Møller, P.R., Thomsen, P.F., 2016. Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. Nat. Ecol. Evol. 1, 4. doi:10.1038/s41559-016-0004

Silvestro, D., Cascales-Miñana, B., Bacon, C.D., Antonelli, A., 2015. Revisiting the origin and diversification of vascular plants through a

comprehensive Bayesian analysis of the fossil record. New Phytol. 207, 425–436. doi:10.1111/nph.13247

Soininen, E.M., Gauthier, G., Bilodeau, F., Berteaux, D., Gielly, L., Taberlet, P., Gussarova, G., Bellemain, E., Hassel, K., Stenøien, H.K., Epp, L., Schrøder-Nielsen, A., Brochmann, C., Yoccoz, N.G., 2015. Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. PLoS ONE 10, e0115335. doi:10.1371/journal.pone.0115335

Sønstebø, J.H., Gielly, L., Brysting, A.K., Elven, R., Edwards, M., Haile, J., Willerslev, E., Coissac, E., Rioux, D., Sannier, J., Taberlet, P., Brochmann, C., 2010. Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. Mol. Ecol. Resour. 10, 1009–1018. doi:10.1111/j.1755-0998.2010.02855.x

Souza, A.C. de, Prevedello, J.A., 2020. The importance of protected areas for overexploited plants: Evidence from a biodiversity hotspot. Biol. Conserv 243, 108482. doi:10.1016/j.biocon.2020.108482

Starr, M., Vasander, H., 2021. Nature and characteristics of forest soils and peat soils as niches for microorganisms, in: Forest Microbiology. Elsevier, pp. 223–229. doi:10.1016/B978-0-12-822542-4.00021-8

Stefańska-Krzaczek, E., Swacha, G., Żarnowiec, J., Raduła, M.W., Kącki, Z., Staniaszek-Kik, M., 2022. Central European forest floor bryophytes: Richness, species composition, coexistence and diagnostic significance across environmental gradients of forest habitats. Ecological Indicators 139, 108954. doi:10.1016/j.ecolind.2022.108954

Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I., Bennett, E.M., Biggs, R., Carpenter, S.R., de Vries, W., de Wit, C.A., Folke, C., Gerten, D., Heinke, J., Mace, G.M., Persson, L.M., Ramanathan, V., Reyers, B., Sörlin, S., 2015. Sustainability. Planetary boundaries: guiding human development on a changing planet. Science 347, 1259855. doi:10.1126/science.1259855

Stein, A., Gerstner, K., Kreft, H., 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecol. Lett. 17, 866–880. doi:10.1111/ele.12277

Stewart, K.A., 2019. Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. Biodivers. Conserv. 28, 983–1001. doi:10.1007/s10531-019-01709-8

Taberlet, P., Bonin, A., Zinger, L., Coissac, E. (Eds.), 2018. Environmental DNA: For Biodiversity Research and Monitoring.Oxford University Press. Oxford.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Mol. Ecol. 21, 2045–2050. doi:10.1111/j.1365-294X.2012.05470.x

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., Willerslev, E., 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Res. 35, e14. doi:10.1093/nar/gkl938

Teixidor-Toneu, I., Jordan, F.M., Hawkins, J.A., 2018. Comparative phylogenetic methods and the cultural evolution of medicinal plant use. Nat. Plants 4, 754–761. doi:10.1038/s41477-018-0226-6

Thomsen, P.F., Willerslev, E., 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biol. Conserv 183, 4–18. doi:10.1016/j.biocon.2014.11.019

Torma, M., 2022. geco-nhm/natur-i-norge: 2.3-beta. Zenodo. doi:10.5281/zenodo.6687810

Twyford, A.D., Ness, R.W., 2017. Strategies for complete plastid genome sequencing. Mol. Ecol. Resour. 17, 858–868. doi:10.1111/1755-0998.12626

Valentini, A., 2007. Non-invasive diet analysis based on DNA barcoding: the Himalayan brown bears (*Ursus arctos isabellinus*) as a case study (Doctoral dissertation).University of Tucsia.

van Kleunen, M., Weber, E., Fischer, M., 2010. A meta-analysis of trait differences between invasive and non-invasive plant species. Ecol. Lett. 13, 235–245. doi:10.1111/j.1461-0248.2009.01418.x

Vellend, M., Baeten, L., Becker-Scarpitta, A., Boucher-Lalonde, V., McCune, J.L., Messier, J., Myers-Smith, I.H., Sax, D.F., 2017. Plant biodiversity change across scales during the anthropocene. Annu. Rev. Plant Biol. 68, 563–586. doi:10.1146/annurev-arplant-042916-040949

Watson, J.D., Crick, F.H., 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171, 737–738. doi:10.1038/171737a0

Watson, J.E.M., Shanahan, D.F., Di Marco, M., Allan, J., Laurance, W.F., Sanderson, E.W., Mackey, B., Venter, O., 2016. Catastrophic declines in wilderness areas undermine global environment targets. Curr. Biol. 26, 2929–2934. doi:10.1016/j.cub.2016.08.049

Wheeler, Q.D., Pennak, S., 2011. State of observed species. International Institute for Species Exploration.

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M.E., Lorenzen, E.D., Vestergård, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L.S., Pearman, P.B., Cheddadi, R., Murray, D., Bråthen, K.A., Yoccoz, N., Binney, H., Taberlet, P., 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. Nature 506, 47–51. doi:10.1038/nature12921

Williamson, J., Maurin, O., Shiba, S.N.S., van der Bank, H., Pfab, M., Pilusa, M., Kabongo, R.M., van der Bank, M., 2016. Exposing the illegal trade in cycad species (Cycadophyta: *Encephalartos*) at two traditional medicine markets in South Africa using DNA barcoding. Genome 59, 771–781. doi:10.1139/gen-2016-0032

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., Pang, X., Xu, H., Zhu, Y., Xiao, P., Chen, S., 2010. Use of ITS2 region as the universal DNA barcode for plants and animals. PLoS ONE 5. doi:10.1371/journal.pone.0013102

Yates, M.C., Derry, A.M., Cristescu, M.E., 2021. Environmental RNA: A revolution in ecological resolution? Trends Ecol. Evol. 36, 601–609. doi:10.1016/j.tree.2021.03.001

Yoccoz, N.G., Bråthen, K.A., Gielly, L., Haile, J., Edwards, M.E., Goslar, T., Von Stedingk, H., Brysting, A.K., Coissac, E., Pompanon, F., Sønstebø, J.H., Miquel, C., Valentini, A., De Bello, F., Chave, J., Thuiller, W., Wincker, P., Cruaud, C., Gavory, F., Rasmussen, M., Taberlet, P., 2012. DNA from soil mirrors plant taxonomic and growth form diversity. Mol. Ecol. 21, 3647–3655. doi:10.1111/j.1365-294X.2012.05545.x

Paper I

Chapter 24
Environmental and biodiversity assessments

Maria Ariza[1], Sandra Garcés-Pastor[2], Hugo J. de Boer[1]

1 Natural History Museum, University of Oslo, Norway
2 The Arctic University Museum of Norway, UiT - The Arctic University of Norway, Tromsø, Norway

Maria Ariza **mariadelosangelesariza@gmail.com**
Sandra Garcés-Pastor **sandra.garces-pastor@uit.no**
Hugo J. de Boer **h.de.boer@nhm.uio.no**

Citation: Ariza M, Garcés-Pastor S, de Boer HJ (2022) Chapter 24. Environmental and biodiversity assessments. In: de Boer H, Rydmark MO, Verstraete B, Gravendeel B (Eds) Molecular identification of plants: from sequence to species. Advanced Books. https://doi.org/10.3897/ab.e98875

Introduction

Being the world's most abundant life kingdom, plants are virtually everywhere: in terrestrial, freshwater, and marine ecosystems and even in the air in the form of pollen and spores (Bar-On et al. 2018). They can survive in extreme environments such as the arctic, deserts, and even concrete (Antonelli et al. 2020). Plants are crucial to nearly all ecosystems, and sustain primary production, nutrient cycling, food chains, and multi-scale networks (Corlett 2020). These characteristics make them good indicators of associated biodiversity, surrounding abiotic features, anthropogenic activities (Brunbjerg et al. 2018; Kier et al. 2005; Terwayet Bayouli et al. 2021; Uuemaa et al. 2013) and suitable organisms for environmental and total biodiversity assessments. Since many plants are sessile and perennial, their spatial distribution is not restricted to temporal fluctuations as with organisms, i.e., animals, and thus diversity can be easily quantified, leveraging the accuracy and efficiency of its assessment. Indeed, plant biodiversity assessments are often used to describe biome and landscape changes, to map habitats, and to monitor environmental quality, pollution, and responses to climate change (Halvorsen et al. 2020; Mucina 2019; Steinbauer et al. 2018; Terwayet Bayouli et al. 2021).

However, plant biodiversity assessments are impeded by problems associated with species detection, taxonomic assignment, abundance quantification, and sample bias given the unknown spatial and temporal distribution of target species (Beng and Corlett 2020). Traditional plant assessments have relied on plant morphological characters to identify and inventory diversity, these processes are also often limited to seasonal or life-history stages and require skilled botanists (Scott and Hallam 2003). Additionally, morphology-based assessments are labour intensive, invasive, and prone to observer-bias (Milberg et al. 2008). Although plant identification through organismal or extra-organismal DNA traces extracted from environmental samples (namely environmental DNA or eDNA) has enabled multiple and simultaneous detections at any season, including detection of rare taxa and those that are challenging to collect, complete and reliable plant biodiversity assessments remain challenging (Deiner et al. 2017; Hartvig et al. 2021; Taberlet et al. 2012). Hence, the complementary strength and knowledge of both traditional and eDNA-based assessments and from botanists and molecular ecologists is still required for better estimations of total plant diversity.

Improving plant biodiversity assessments is one of the century's greatest challenges as less than 10% of the world's plant diversity is currently known, and its loss outpaces the rate at which is discovered, inventoried, and protected (Corlett 2016). Furthermore, current global pressures on biodiversity, e.g., invasive species, climate change, environmental pollution and habitat loss, highlight the necessity of biodiversity data to mitigate these impacts (Corlett 2020). Molecular inventorying of plant diversity through eDNA-based assessments show great potential to meet these needs and offers novel opportunities to register the dynamics of species, populations, and communities over long time periods and across large spatial scales (Kersey et al. 2020; Rodríguez-Ezpeleta et al. 2021). This chapter focuses on plant biodiversity (green algae, liverworts, hornworts, mosses, and vascular plants) and environmental characteristics that can be assessed using both eDNA substrates and organismal DNA, and their applications to conservation, ecology, monitoring both diversity and invasive species.

Assessing plant DNA from the environment: power, precautions, and limitations

While many plants are sessile and their biomass is mainly located below or above anchoring surfaces, some vegetative and reproductive plant parts (i.e. flowers, leaf debris, pollen, seeds) detach and are transported on short or great distances from the main organismal body until they are finally deposited onto substrates (i.e., ground, water, and more). Hence, plant DNA can be found in environmental substrates as organismal and extra-organismal DNA at various proportions, with each substrate potentially tracking different spatial and temporal signatures of biodiversity (Rodríguez-Ezpeleta et al. 2021). The detection of these plant DNA sources can also be associated with both DNA status (intracellular or extracellular) and environmental conditions that may enhance or diminish DNA permanence and degradation, i.e., organic particles that bind DNA support its environmental persistence or UV light exposure that results in degradation (Nagler et al. 2018; Pietramellara et al. 2009). Nevertheless, DNA from environmental substrates degrades and decays over time and thus, its assessment can be facilitated by targeting short informative DNA fragments (Shogren et al. 2018). Indeed, eDNA-based plant assessments commonly employ metabarcoding analysis of the chloroplast *trnL* (UUA) intron p6 loop which has a short sequence ranging from 10–143 bp and primer binding sites that are well conserved in vascular and nonvascular plants (Taberlet et al. 2007). As DNA degrades over time, it is easier to target a short fragment for amplification for eDNA, sedaDNA and aDNA applications. Additionally, the p6 loop has a secondary structure that provides extra stability and resistance to degradation (Taberlet et al. 2007). However, low species resolution, particularly for bryophytes, and misidentification due to PCR bias hinders the use of this region to perform complete biodiversity assessments (Ariza et al. 2022).

As no single marker provides resolution for all taxa, eDNA-based assessments often employ metabarcoding of different nuclear and chloroplast regions such as ITS, *rbcL*, and *matK* to harvest their complementary resolution power (see Chapter 11 Amplicon metabarcoding for information about these regions and their suitable applicability; CBOL Plant Working Group 2009; Hollingsworth et al. 2011). Targeted capture of multiple informative genes and shotgun sequencing of environmental samples have recently gained attention as alternative approaches for assessment of plant diversity from eDNA samples as amplification-free methods (see Chapter 12 Metagenomics and Chapter 14 Target capture for more on these methods and the markers used; Chua et al. 2021a; Foster et al. 2021).

Despite the major recent advances in detection, eDNA-based assessments remain limited to reliably quantify abundance, which in turn makes it hard to assess population status and take management actions (Deiner et al. 2017). Although correlations between plant biomass and DNA concentration in the environmental samples are poorly understood, the use of sequence counts of identified taxa is becoming widely accepted in eDNA studies as a proxy of relative abundance (Deagle et al. 2019, 2013; Deiner et al. 2021). Particularly for plants, the assessment of eDNA from root communities has been shown to provide robust abundance estimations (Matesanz et al. 2019).

Furthermore, presence/absence estimations provided by eDNA-based assessments can be misleading as DNA may remain in the environment after the organism is no longer present (Harrison et al. 2019). Thus, plant eDNA-based assessments should be interpreted as merely detections of organismal DNA until evaluations of false occurrence estimations are investigated. Site occupancy-detection models have recently gained attention for this purpose, though false detections of plant DNA remain largely unexplored (Ficetola et al. 2016; Guillera-Arroita et al. 2017).

Substrates for eDNA in environmental and biodiversity assessments

About a decade after the term eDNA was introduced, the eDNA scientific community has adopted different terminology in reference to the state, source, or substrate from which eDNA is isolated (Pawlowski et al. 2020; Rodríguez-Ezpeleta et al. 2021). This chapter focuses on the substrates from which plant eDNA can be isolated, including air, faeces, pollen, soil, sediments, and water, as well as bulk samples such as flowers, leaves, or roots from which organismal DNA can be isolated. Each of these substrates harbours different plant eDNA sources and spatio-temporal signals from the environment. Careful consideration of the study questions and/or applications are required when selecting an eDNA substrate as this will impact the conclusions that can be derived from the assessments. More details on sampling and DNA extraction from eDNA substrates can be found in section 1 of this book.

Airborne samples

Pollen DNA is most commonly the main source of plant eDNA present in airborne samples, although single-cell algae, leaf and flower fragments may also be present (Eaton et al. 2018; Johnson et al. 2019; Núñez et al. 2019, 2017; Sherwood et al. 2017). Pollen from anemophilous terrestrial plants is especially abundant in airborne samples. Since airborne pollen can be transported over long distances it can provide information on regional vegetation (Eaton et al. 2018; Johnson et al. 2019; Núñez et al. 2019, 2017; Sherwood et al. 2017). Using dust traps, pollen from insect-pollinated plants can also be detected but its relation to local plant biomass and the effect of climatic conditions such as wind and temperature on detectability are poorly understood. Nevertheless, plant assessments through pollen metabarcoding from airborne samples have successfully characterised spatial and temporal heterogeneity (Leontidou et al. 2021; Polling et al. 2022), airborne communities (Craine et al. 2017; Núñez et al. 2017), and have been applied to pollen allergen monitoring (Kraaijeveld et al. 2015; Polling et al. 2022; Rowney et al. 2021). eDNA-based airborne monitoring in particular leverages the identification resolution of common plant-allergen families, i.e., Urticaceae, Taxaceae, Poaceae, and abundance estimations (Campbell et al. 2020; Polling et al. 2022; Rowney et al. 2021).

Faecal substrates

Faeces, mucus, and saliva contain DNA from the host and from the organisms that were ingested or that have been in contact with the host (Valentini 2007). Here, we follow Yoccoz (2012) and Pawlowski et al. (2020) and include faeces and other bodily substances as eDNA. Other authors have excluded these sources of DNA as host-associated and distinct from environmentally distributed DNA. It is important to consider that although such DNA transported in faeces and other materials associated with animals can become environmental DNA, it is not yet the case when faeces is collected for dietary assessments. Faecal samples are the most common excrement source of eDNA used for plant assessments and provide a snapshot of vegetation implicated in trophic interactions. Faeces from herbivorous animals are most commonly used, as droppings are easy to collect and represent a viable option to detect the diet of elusive animals (Holechek et al. 1982). Compared to

morphological assessments of plant remains in faeces, faecal DNA metabarcoding and metagenomics have leveraged the taxonomic resolution of plant dietary items from extinct megafauna, mammals, birds, reptiles, insects, and molluscs (Chua et al. 2021b; Koizumi et al. 2016; Polling et al. 2021; Valentini et al. 2009), revealing in turn more diverse diets than previously conceived (Chua et al. 2021b). Simultaneously, eDNA-inventorying of plant diet items have provided new ecological information to support habitat protection efforts (Chua et al. 2021b; Yamamoto and Uchida 2018), and the monitoring of invasive species (Mori et al. 2017), overgrazing (Craine et al. 2015; Fløjgaard et al. 2017), and dietary niche dynamics (Jorns et al. 2020; Kartzinel et al. 2015; Schure et al. 2021). Furthermore, parallel eDNA assessments of scats from communities of large herbivores has allowed the reconstruction of present and past landscape mosaics of the dominant vegetation (Polling et al. 2021; Schure et al. 2021). Moreover, the collection of residual saliva or mucus directly from plant organs can guide the evaluation of niche specialism and competition for plant resources (Harrer and Levi 2018).

Soil and sedimentary substrates

Soil and sediments, from both terrestrial and aquatic environments, are presumably the substrates where most plant DNA is present, as extra-organismal and organismal DNA from both active and dormant tissues including, roots, debris, fallen vegetative parts, seeds, and pollen are gathered or ultimately deposited in these substrates. Because of the major presence of plant eDNA and the ubiquity of these substrates in both aquatic and terrestrial ecosystems, soil and sedimentary eDNA samples are advantageously appropriate for plant assessments. Differences between soil and sediments can be ambiguous, as both are products of the earth's crusts weathering (Wood 1987). However, in soils the deposition of these products happens in situ and remains on the surface, while in sediments these products are transported and redeposited elsewhere in layers over time. As a consequence, these substrates have different spatio-temporal signals when it comes to the reconstruction of the environment (Deiner et al. 2017; Thomsen and Willerslev 2015). Plant eDNA from soil has been shown to signal local and contemporary vegetation (Ariza et al. 2022; Edwards et al. 2018; Yoccoz et al. 2012), whereas sedimentary samples from marine, lake, or terrestrial cores can combine local, regional, contemporary, and past vegetation signals (Alsos et al. 2018; Thomsen and Willerslev 2015; Willerslev et al. 2003).

Soil eDNA plant assessments have successfully characterised diversity in tropical (Osathanunkul et al. 2021; Yoccoz 2012; Zinger et al. 2019), temperate (Fahner et al. 2016; Yang et al. 2014; Yoccoz et al. 2012), and boreal ecosystems (Edwards et al. 2018; Yoccoz et al. 2012), including the hidden diversity from extreme environments such as deserts (Carrasco-Puga et al. 2021; Palacios Mejia et al. 2021), Antarctica (Carvalho-Silva et al. 2021), geothermal sites (Fraser et al. 2017), and permafrost (Willerslev et al. 2014). Soil eDNA plant inventories have been used to assess both natural and cultivated landscapes (Foucher et al. 2020; Yoccoz et al. 2012), woody encroachment in grasslands (Sepp et al. 2021), habitat from crime scenes (Fløjgaard et al. 2019), and rare terrestrial orchids (Hartvig et al. 2021).

As sediments are deposited throughout time and form distinguishable layers, the eDNA present in these layers (namely sedaDNA) can signal organisms that were likely locally present in ancient environments (Thomsen and Willerslev 2015). The assessment of plant eDNA present in terrestrial ancient sediments has been used to reconstruct the vegetation of the Pleistocene and Holocene in Siberia (Liu et al. 2021; Willerslev et al. 2003), and glacial and interglacial periods in the Arctic (Sønstebø et al. 2010). Further, plant eDNA from sedimentary samples collected in freshwater ecosystems, i.e., lake or riverine sediments, can signal contemporary

and surrounding terrestrial vegetation (Alsos et al. 2018; Giguet-Covex et al. 2019). However, ancient plant DNA present in these samples has been purposely targeted to infer past vegetations including paleo floras (Parducci et al. 2017; Thomsen and Willerslev 2015). Plant eDNA collected from lake sediments has revealed vegetation growing in the arctic during the last interglacial (Crump et al. 2021; Parducci et al. 2012) and post-glacial migration of trees from southern Europe (Epp et al. 2015), human-induced landscape changes and the biological invasions that followed (Ficetola et al. 2018; Giguet-Covex et al. 2014), and even a 5000 year timeline of tropical diversity (Bremond et al. 2017). eDNA metabarcoding of sediments from ancient water reservoirs at the centre of major Maya cities was used to reconstruct the forest types in these ancient cities (Lentz et al. 2021). Finally, eDNA present in coastal marine sediments has been used to monitor seagrasses, salt marshes, and mangrove communities (Foster et al. 2020; Ortega et al. 2020a).

Water samples

eDNA-based biodiversity assessments have proliferated in marine and freshwater environments in recent years, and our knowledge on the persistence, decay rates, and states of eDNA in water samples and its resolution compared to traditional assessments has in parallel increased (Ji et al. 2021; Mauvisseau et al. 2022). However, assessments of plant biodiversity from aquatic environments have been widely overlooked compared to assessments of other organisms across the tree of life. Presumably, plant eDNA present in water samples is mostly composed of extra-organismal DNA bound to suspended small particles derived either from terrestrial or aquatic habitats (Deiner et al. 2016; Drummond et al. 2021; Lacoursière-Roussel and Deiner 2021; Turner et al. 2014). In addition, DNA presence can be vertically stratified, influencing the signals that are retrieved with either shallow or deep water samples (Canals et al. 2021). However, comparisons of assessed diversity with water samples collected at different vertical and horizontal positions in small lakes revealed similar aquatic and terrestrial plant signals, suggesting that eDNA is evenly distributed in freshwater environments and that shore-based sampling can successfully capture beta diversity (Drummond et al. 2021). The latter study in addition showed that read abundances are heavily weighted toward aquatic macrophytes, while taxon richness was greatest in algae and other nonvascular plants. Similar detection patterns were registered in rivers (Ji et al. 2021). Furthermore, aquarium experiments suggest that eDNA concentration and submerged biomass are positively correlated (Matsuhashi et al. 2016).

The assessment of aquatic plant eDNA in freshwater ecosystems has simultaneously enabled the early detection of invasive species (Coghlan et al. 2021; Doi et al. 2021; Fujiwara et al. 2016; Gantz et al. 2018; Scriver et al. 2015), endangered species (Tsukamoto et al. 2021), and water quality indicator species (Gao et al. 2018; Kuzmina et al. 2018; Stoeck et al. 2018). Assessing plant diversity from eDNA in marine systems is harder due to salinity and the massive volumes of water in which plant DNA is diluted. Several studies have still shown the feasibility of using marine plant eDNA to study marine macrophytes (Foster et al. 2020) as well as blue carbon cycling (Ortega et al. 2019, 2020b).

Plant DNA can also be isolated from water samples in the form of snow, firn, and ice (Pedersen et al. 2015). In particular, glacier ice can be advantageous for plant assessments as it gathers plant remains from surrounding environments while maintaining freezing temperatures that preserve DNA naturally for long periods and thus allows the reconstruction of past environments (Varotto et al. 2021). Plant assessments from glacier ice cores have allowed the reconstruction of the conifer communities that once inhabited Greenland (Willerslev et al. 2007) and vegetation transitions during the Last Glacial Maximum throughout Beringia (Pedersen et al. 2016).

Bulk samples

Bulk samples from plants are distinctly different from pitfall or Malaise traps filled with insects. In bulk samples of plants, one can distinguish natural bulk samples such as pollen samples from pollen samplers, or those scraped or washed from pollinating vectors, and those that are artificially assembled such as collected roots, leaves, or flowers. Nevertheless, all bulk samples constitute organismal DNA from plant communities that can be used either to assess plant or other diversity (Deiner et al. 2017).

Flower bulk samples have been assembled to assess arthropod communities that leave DNA traces after either visitation or pollination (Thomsen and Sigsgaard 2019). Leaf bulk samples can be easily collected from the leaf litter. The latter has been particularly used to assess soil fauna and arthropod communities as it can reveal differences in habitat and beta diversity (Lopes et al. 2021; Ritter et al. 2018; Yang et al. 2014). However, the potential of leaf litter metabarcoding to assess vegetation remains unexplored. Artificially assembled leaf bulk samples have been used to assess plant diversity in tropical forests in the Brazilian canga (Vasconcelos et al. 2021). Natural pollen bulk samples are often collected from pollinators or flower visitors, particularly from their pollen baskets (Sookhan et al. 2021). Plant signals from these samples mainly correspond to vegetation involved in ecological interactions of pollination and parasitism and thus are valuable to reconstruct food webs (McFrederick and Rehan 2016; Sookhan et al. 2021). DNA metabarcoding of pollen bulk samples can be used to assess more diverse pollination networks from insects and bats as well as the seasonal availability of food resources (Koyama et al. 2018; Lim et al. 2018; Smart et al. 2017). Furthermore, pollen present in honey can be regarded as a bulk sample as it signals floral sources that bees have visited. Melissopalynology metabarcoding studies have focused either on identification of floral composition of honey, regional provenance, or identification of entomological sources of the honey (Chiara et al. 2021; Hawkins et al. 2015; Milla et al. 2021; Prosser and Hebert 2017; Richardson et al. 2015). Artificially assembled pollen samples such as pollen collected using Burkard samplers for allergenic pollen prognoses can be used to identify airborne pollen as well. Root bulk samples can be assembled to signal hidden belowground plant diversity and its abundance (Lamb et al. 2016; Matesanz et al. 2019). Metabarcoding root diversity in grasslands has revealed a larger fraction of diversity that cannot be detected with traditional assessments of aboveground diversity (Rucińska et al. 2022; Sepp et al. 2021). In addition, the assessment of root bulk samples has elucidated mycorrhizal and parasitic plant associations (Holá et al. 2017; Marčiulynienė et al. 2021).

Beyond eDNA samples: assessing biodiversity through eDNA biotic samplers

A recent development in eDNA metabarcoding is the use of organisms as natural samplers of DNA (coined nsDNA; Mariani et al. 2019). Siegenthaler et al. (2019) show how DNA assessment of gut contents from the European brown shrimp can recover the same number of taxa as using water or sediment eDNA samples from the same area where the shrimps were collected. Similarly, sponges have been shown to be robust natural samplers as they filter high volumes of water and simultaneously trap and concentrate DNA traces from faunal assemblages (Mariani et al. 2019; M. Turon et al. 2020). In terrestrial ecosystems, insectivorous bats have proven to be useful for assessing natural and invasive insect pests (Kemp et

al. 2019; Montauban et al. 2021). Most valuations of biotic samplers have focused on their potential to assess fauna whilst for flora this remains rather unexplored. Hence, we will highlight a few examples of potential biotic samplers that can characterise local floras and other environmental characteristics.

In aquatic ecosystems, macroinvertebrates (Chironomidae, Coleoptera, Hemiptera, Ephemeroptera) that feed both on aquatic vascular plants and plant fragments leached to the environment hold great potential to signal overall vegetation implicated in freshwater trophic relationships. Likewise, filtering organisms or animals that use specialised structures to filter fine particles from the water in lakes and rivers harbour the same potential, i.e., sponges (*Ephydatia*), Simuliidae, Ephemeroptera, Chironomidae, and Trichoptera.

For the assessment of terrestrial vegetation in tropical areas, bats hold great potential as biotic samplers of plant DNA since omnivorous and frugivorous communities are abundant and thus easy to collect (Kalko et al. 1996). For example, seed rains from leaf-nosed bats (Phyllostomidae) can signal understory vegetation that is presently abundant and part of secondary forest succession (Andrade et al. 2013; Charles-Dominique and Cockle 2001). Moreover, assessment of seed rain over time can help track phenological adaptations resulting from recent community turnover and reveal competition avoidance mechanisms of plant coexistence (Thies and Kalko 2004). The DNA assessment of seed rains may also overcome the low taxonomic resolution that traditional morphological identification of seeds yields. Furthermore, specialist organisms for pollination, nectarivore, and seed dispersal harbour the potential to detect elusive plants and reveal other plants that are visited or potentially pollinated. For example, 600 neotropical orchids are specifically pollinated by *Euglossa* bees, which in turn can visit other floral sources (Pemberton and Wheeler 2006; Ramírez et al. 2011).

Finally, amplifying hypervariable markers from biotic DNA samplers, i.e., COI for animals, has recently gained attention as it can assess diversity below the species level, and thus signals ecosystem population assemblages in space and time (metaphylogeography; X. Turon et al. 2020). Metaphylogeography datasets have the potential to provide novel insights that can be applied to conservation genetics, biodiversity management and assessment of protected areas.

While the exploration of eDNA samples and methods for plant assessments is still at its infancy, eDNA has already revolutionised the way and speed in which biodiversity can be inventoried. Plant detection via eDNA has enabled the discovery of plants living in extreme and/or ancient environments and yielded myriad applications with societal relevance. A decade after the rise of eDNA-based assessments, the limitations of this method across different eDNA samples are still being recognised while in parallel different strategies are being developed to overcome and mitigate these. In this rapidly developing field, it is essential to combine the basics of eDNA metabarcoding with the most recent insights and developments in the field to devise the most robust study design to answer your research questions.

Questions

1. You want to assess the floral resources available in summer for a butterfly species and identify potential food competitors. Describe your experimental design and the eDNA substrate(s) that you would use and why.
2. You are hired to conduct a vegetation assessment of a landscape mosaic composed of several small lakes and grasslands, however, you only have the time and budget to collect samples from a single eDNA substrate. Which eDNA substrate would you choose and why?

3. You use soil eDNA to detect the spread of an invasive alien gymnosperm tree species (Sitka spruce, *Picea sitchensis*). Though this species is conspicuously visible, you have not seen it nor has been reported around the sampling area.You detect OTUs in nearly every possible sample, and after a bout of cold sweat realise how this might be explained. What would explain this finding?

Glossary

Organismal DNA – The DNA that is isolated from bulk-extracted mixtures of organisms that are separated from the environmental sample. Also named community DNA.

Extra-organismal DNA – DNA originated (i) from biological material shed from an organism as part of tissue replacement or metabolic waste; (ii) as biologically active propagules such as gametes, pollen, seeds or spores; or (iii) as a result of cell lysis or cell extrusion (Rodriguez-Ezpeleta et al. 2021).

Environmental DNA – DNA captured from modern environments, i.e., seawater, freshwater, soil, or air; or ancient environments, i.e., cores from sediment, ice or permafrost (Thomsen and Willerslev 2015) that have originated from both organismal and extra organismal DNA (Rodriguez-Ezpeleta et al. 2021).

Intracellular DNA – DNA that is located within cell membranes.

Extracellular DNA – DNA that is located free in the environment after cell lysis or cell extrusion.

Anemophily – Plant pollination where pollen is distributed by wind, i.e. wind pollination.

Firn – Crystalline or granular snow, especially on the upper part of a glacier, where it has not yet been compressed into ice.

Melissopalynology – The study of pollen contained in honey and, in particular, the pollen's source.

References

Alsos IG, Lammers Y, Yoccoz NG, Jørgensen T, Sjögren P, Gielly L, Edwards ME (2018) Plant DNA metabarcoding of lake sediments: how does it represent the contemporary vegetation. PLoS ONE 13, e0195403. https://doi.org/10.1371/journal.pone.0195403

Andrade TY, Thies W, Rogeri PK, Kalko EKV, Mello MAR (2013) Hierarchical fruit selection by Neotropical leaf-nosed bats (Chiroptera: Phyllostomidae). J. Mammal. 94, 1094–1101. https://doi.org/10.1644/12-MAMM-A-244.1

Antonelli A, Fry C, Smith RJ, Simmonds MSJ, Kersey PJ, Pritchard HW, Abbo MS, Acedo C, Adams J, Ainsworth AM, Allkin B, Annecke W, Bachman SP, Bacon K, Bárrios S, Barstow C, Battison A, Bell E, Bensusan K, Bidartondo MI, et al. (2020) State of the World's Plants and Fungi 2020. Royal Botanic Gardens, Kew. https://doi.org/10.34885/172

Ariza M, Fouks B, Mauvisseau Q, Halvorsen R, Alsos IG, de Boer H (2022) Plant biodiversity assessment through soil eDNA reflects temporal and local diversity. Methods Ecol. Evol. https://doi.org/10.1111/2041-210X.13865

Bar-On YM, Phillips R, Milo R (2018) The biomass distribution on Earth. Proc Natl Acad Sci USA 115, 6506–6511. https://doi.org/10.1073/pnas.1711842115

Beng KC, Corlett RT (2020) Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. Biodivers. Conserv. 29, 2089–2121. https://doi.org/10.1007/s10531-020-01980-0

Bremond L, Favier C, Ficetola GF, Tossou MG, Akouégninou A, Gielly L, Giguet-Covex C, Oslisly R, Salzmann U (2017) Five thousand years of tropical lake sediment DNA records from Benin. Quat. Sci. Rev. 170, 203–211. https://doi.org/10.1016/j.quascirev.2017.06.025

Brunbjerg AK, Bruun HH, Dalby L, Fløjgaard C, Frøslev TG, Høye TT, Goldberg I, Læssøe T, Hansen MDD, Brøndum L, Skipper L, Fog K, Ejrnæs R (2018) Vascular plant species richness and bioindication predict multi-taxon species richness. Methods Ecol. Evol. 9, 2372–2382. https://doi.org/10.1111/2041-210X.13087

Campbell BC, Al Kouba J, Timbrell V, Noor MJ, Massel K, Gilding EK, Angel N, Kemish B, Hugenholtz P, Godwin ID, Davies JM (2020) Tracking seasonal changes in diversity of pollen allergen exposure: targeted metabarcoding of a subtropical aerobiome. Sci. Total Environ. 747, 141189. https://doi.org/10.1016/j.scitotenv.2020.141189

Canals O, Mendibil I, Santos M, Irigoien X, Rodríguez-Ezpeleta N (2021) Vertical stratification of environmentalDNA in the open ocean captures ecological patterns and behavior of deep-sea fishes. Limnol. Oceanogr. 6, 339–347. https://doi.org/10.1002/lol2.10213

Carrasco-Puga G, Díaz FP, Soto DC, Hernández-Castro C, Contreras-López O, Maldonado A, Latorre C, Gutiérrez RA (2021) Revealing hidden plant diversity in arid environments. Ecography 44, 98–111. https://doi.org/10.1111/ecog.05100

Carvalho-Silva M, Rosa LH, Pinto OHB, Da Silva TH, Henriques DK, Convey P, Câmara PEAS (2021) Exploring the plant environmental DNA diversity in soil from two sites on Deception Island (Antarctica, South Shetland Islands) using metabarcoding. Antarctic Science 33, 469–478. https://doi.org/10.1017/S0954102021000274

CBOL Plant Working Group (2009) A DNA barcode for land plants. Proc Natl Acad Sci USA 106, 12794–12797. https://doi.org/10.1073/pnas.0905845106

Charles-Dominique P, Cockle A (2001) Frugivory and seed dispersal by bats, in: Bongers, F., Charles-Dominique, P., Forget, P.-M., Théry, M. (Eds.), Nouragues, Monographiae Biologicae. Springer Netherlands, Dordrecht, pp. 207–216. https://doi.org/10.1007/978-94-015-9821-7_19

Chiara B, Francesco C, Fulvio B, Paola M, Annalisa G, Stefania S, Luigi AP, Simone P (2021) Exploring the botanical composition of polyfloral and monofloral honeys through DNA metabarcoding. Food Control 128, 108175. https://doi.org/10.1016/j.foodcont.2021.108175

Chua PYS, Crampton-Platt A, Lammers Y, Alsos IG, Boessenkool S, Bohmann K (2021a) Metagenomics: a viable tool for reconstructing herbivore diet. Mol. Ecol. Resour. 21, 2249–2263. https://doi.org/10.1111/1755-0998.13425

Chua PYS, Lammers Y, Menoni E, Ekrem T, Bohmann K, Boessenkool S, Alsos IG (2021b) Molecular dietary analyses of western capercaillies (*Tetrao urogallus*) reveal a diverse diet. Environmental DNA 3, 1156–1171. https://doi.org/10.1002/edn3.237

Coghlan SA, Shafer ABA, Freeland JR (2021) Development of an environmental DNA metabarcoding assay for aquatic vascular plant communities. Environmental DNA 3, 372–387. https://doi.org/10.1002/edn3.120

Corlett RT (2016) Plant diversity in a changing world: status, trends, and conservation needs. Plant Diversity 38, 10–16. https://doi.org/10.1016/j.pld.2016.01.001

Corlett RT (2020) Safeguarding our future by protecting biodiversity. Plant Diversity 42, 221–228. https://doi.org/10.1016/j.pld.2020.04.002

Craine JM, Barberán A, Lynch RC, Menninger HL, Dunn RR, Fierer N (2017) Molecular analysis of environmental plant DNA in house dust across the United States. Aerobiologia (Bologna) 33, 71–86. https://doi.org/10.1007/s10453-016-9451-5

Craine JM, Towne EG, Miller M, Fierer N (2015) Climatic warming and the future of bison as grazers. Sci. Rep. 5, 16738. https://doi.org/10.1038/srep16738

Crump SE, Fréchette B, Power M, Cutler S, de Wet G, Raynolds MK, Raberg JH, Briner JP, Thomas EK, Sepúlveda J, Shapiro B, Bunce M, Miller GH (2021) Ancient plant DNA reveals High Arctic greening during the Last Interglacial. Proc Natl Acad Sci USA 118, e2019069118. https://doi.org/10.1073/pnas.2019069118

Deagle BE, Thomas AC, McInnes JC, Clarke LJ, Vesterinen EJ, Clare EL, Kartzinel TR, Eveson JP (2019) Counting with DNA in metabarcoding studies: how should we convert sequence reads to dietary data? Mol. Ecol. 28, 391–406. https://doi.org/10.1111/mec.14734

Deagle BE, Thomas AC, Shaffer AK, Trites AW, Jarman SN (2013) Quantifying sequence proportions in a DNA-based diet study using Ion Torrent amplicon sequencing: which counts count? Mol. Ecol. Resour. 13, 620–633. https://doi.org/10.1111/1755-0998.12103

Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, Creer S, Bista I, Lodge DM, de Vere N, Pfrender ME, Bernatchez L (2017) Environmental DNA metabarcoding: transforming how we survey animal and plant communities. Mol. Ecol. 26, 5872–5895. https://doi.org/10.1111/mec.14350

Deiner K, Fronhofer EA, Mächler E, Walser J-C, Altermatt F (2016) Environmental DNA reveals that rivers are conveyer belts of biodiversity information. Nat. Commun. 7, 12544. https://doi.org/10.1038/ncomms12544

Deiner K, Yamanaka H, Bernatchez L (2021) The future of biodiversity monitoring and conservation utilizing environmental DNA. Environmental DNA 3, 3–7. https://doi.org/10.1002/edn3.178

Doi H, Akamatsu Y, Goto M, Inui R, Komuro T, Nagano M, Minamoto T (2021) Broad-scale detection of environmental DNA for an invasive macrophyte and the relationship between DNA concentration and coverage in rivers. Biol. Invasions 23, 507–520. https://doi.org/10.1007/s10530-020-02380-9

Drummond JA, Larson ER, Li Y, Lodge DM, Gantz CA, Pfrender ME, Renshaw MA, Correa AMS, Egan SP (2021) Diversity metrics are robust to differences in sampling location and depth for environmental DNA of plants in small temperate lakes. Front. Environ. Sci. 9, 617924. https://doi.org/10.3389/fenvs.2021.617924

Eaton S, Zúñiga C, Czyzewski J, Ellis C, Genney DR, Haydon D, Mirzai N, Yahr R (2018) A method for the direct detection of airborne dispersal in lichens. Mol. Ecol. Resour. 18, 240–250. https://doi.org/10.1111/1755-0998.12731

Edwards ME, Alsos IG, Yoccoz N, Coissac E, Goslar T, Gielly L, Haile J, Langdon CT, Tribsch A, Binney HA, von Stedingk H, Taberlet P (2018) Metabarcoding of modern soil DNA gives a highly local vegetation signal in Svalbard tundra. The Holocene 28, 2006–2016. https://doi.org/10.1177/0959683618798095

Epp LS, Gussarova G, Boessenkool S, Olsen J, Haile J, Schrøder-Nielsen A, Ludikova A, Hassel K, Stenøien HK, Funder S, Willerslev E, Kjær K, Brochmann C (2015) Lake sediment multi-taxon DNA from North Greenland records early post-glacial appearance of vascular plants and accurately tracks environmental changes. Quat. Sci. Rev. 117, 152–163. https://doi.org/10.1016/j.quascirev.2015.03.027

Fahner NA, Shokralla S, Baird DJ, Hajibabaei M (2016) Large-scale monitoring of plants through environmental DNA metabarcoding of soil: recovery, resolution, and annotation of four DNA markers. PLoS ONE 11, e0157505. https://doi.org/10.1371/journal.pone.0157505

Ficetola GF, Poulenard J, Sabatier P, Messager E, Gielly L, Leloup A, Etienne D, Bakke J, Malet E, Fanget B, Støren E, Reyss J-L, Taberlet P, Arnaud F (2018) DNA from lake sediments reveals long-term ecosystem changes after a biological invasion. Sci. Adv. 4, eaar4292. https://doi.org/10.1126/sciadv.aar4292

Ficetola GF, Taberlet P, Coissac E (2016) How to limit false positives in environmental DNA and metabarcoding? Mol. Ecol. Resour. 16, 604–607. https://doi.org/10.1111/1755-0998.12508

Fløjgaard C, De Barba M, Taberlet P, Ejrnæs R (2017) Body condition, diet and ecosystem function of red deer (*Cervus elaphus*) in a fenced nature reserve. Glob. Ecol. Conserv. 11, 312–323. https://doi.org/10.1016/j.gecco.2017.07.003

Fløjgaard C, Frøslev TG, Brunbjerg AK, Bruun HH, Moeslund J, Hansen AJ, Ejrnæs R (2019) Predicting provenance of forensic soil samples: linking soil to ecological habitats by metabarcoding and supervised classification. PLoS ONE 14, e0202844. https://doi.org/10.1371/journal.pone.0202844

Foster NR, Gillanders BM, Jones AR, Young JM, Waycott M (2020) A muddy time capsule: using sediment environmental DNA for the long-term monitoring of coastal vegetated ecosystems. Mar. Freshwater Res. 71, 869. https://doi.org/10.1071/MF19175

Foster NR, van Dijk K, Biffin E, Young JM, Thomson VA, Gillanders BM, Jones AR, Waycott M (2021) A multi-gene region targeted capture approach to detect plant DNA in environmental samples: a case study from coastal environments. Front. Ecol. Evol. 9, 735744. https://doi.org/10.3389/fevo.2021.735744

Foucher A, Evrard O, Ficetola GF, Gielly L, Poulain J, Giguet-Covex C, Laceby JP, Salvador-Blanes S, Cerdan O, Poulenard J (2020) Persistence of environmental DNA in cultivated soils: implication of this memory effect for reconstructing the dynamics of land use and cover changes. Sci. Rep. 10, 10502. https://doi.org/10.1038/s41598-020-67452-1

Fraser CI, Connell L, Lee CK, Cary SC (2017) Evidence of plant and animal communities at exposed and subglacial (cave) geothermal sites in Antarctica. Polar Biol. 41, 1–5. https://doi.org/10.1007/s00300-017-2198-9

Fujiwara A, Matsuhashi S, Doi H, Yamamoto S, Minamoto T (2016) Use of environmental DNA to survey the distribution of an invasive submerged plant in ponds. Freshwater Science 35, 748–754. https://doi.org/10.1086/685882

Gantz CA, Renshaw MA, Erickson D, Lodge DM, Egan SP (2018) Environmental DNA detection of aquatic invasive plants in lab mesocosm and natural field conditions. Biol. Invasions 20, 1–18. https://doi.org/10.1007/s10530-018-1718-z

Gao W, Chen Z, Li Y, Pan Y, Zhu J, Guo S, Hu L, Huang J (2018) Bioassessment of a drinking water reservoir using plankton: high throughput sequencing vs. traditional morphological method. Water (Basel) 10, 82. https://doi.org/10.3390/w10010082

Giguet-Covex C, Ficetola GF, Walsh K, Poulenard J, Bajard M, Fouinat L, Sabatier P, Gielly L, Messager E, Develle AL, David F, Taberlet P, Brisset E, Guiter F, Sinet R, Arnaud F (2019) New insights on lake sediment DNA from the catchment: importance of taphonomic and analytical issues on the record quality. Sci. Rep. 9, 14676. https://doi.org/10.1038/s41598-019-50339-1

Giguet-Covex C, Pansu J, Arnaud F, Rey P-J, Griggo C, Gielly L, Domaizon I, Coissac E, David F, Choler P, Poulenard J, Taberlet P (2014) Long livestock farming history and human landscape shaping revealed by lake sediment DNA. Nat. Commun. 5, 3211. https://doi.org/10.1038/ncomms4211

Guillera-Arroita G, Lahoz-Monfort JJ, van Rooyen AR, Weeks AR, Tingley R (2017) Dealing with false-positive and false-negative errors about species occurrence at multiple levels. Methods Ecol. Evol. 8, 1081–1091. https://doi.org/10.1111/2041-210X.12743

Halvorsen R, Skarpaas O, Bryn A, Bratli H, Erikstad L, Simensen T, Lieungh E (2020) Towards a systematics of ecodiversity: the EcoSyst framework. Global Ecol. Biogeogr. 29, 1887–1906. https://doi.org/10.1111/geb.13164

Harrer LEF, Levi T (2018) The primacy of bears as seed dispersers in salmon-bearing ecosystems. Ecosphere 9, e02076. https://doi.org/10.1002/ecs2.2076

Harrison JB, Sunday JM, Rogers SM (2019) Predicting the fate of eDNA in the environment and implications for studying biodiversity. Proc. Biol. Sci. 286, 20191409. https://doi.org/10.1098/rspb.2019.1409

Hartvig I, Kosawang C, Kjær ED, Nielsen LR (2021) Detecting rare terrestrial orchids and associated plant communities from soil samples with eDNA methods. Biodivers. Conserv. 30, 3879–3901. https://doi.org/10.1007/s10531-021-02279-4

Hawkins J, de Vere N, Griffith A, Ford CR, Allainguillaume J, Hegarty MJ, Baillie L, Adams-Groom B (2015) Using DNA metabarcoding to identify the floral composition of honey: A new tool for investigating honey bee foraging preferences. PLoS ONE 10, e0134735. https://doi.org/10.1371/journal.pone.0134735

Holá E, Kocková J, Těšitel J (2017) DNA barcoding as a tool for identification of host association of root-hemiparasitic plants. Folia Geobot. 52, 227–238. https://doi.org/10.1007/s12224-017-9286-z

Holechek JL, Vavra M, Pieper RD (1982) Botanical compostion determination of range herbivore diets. a review. Journal of Range Management 35, 309–315.

Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. PLoS ONE 6, e19254. https://doi.org/10.1371/journal.pone.0019254

Ji F, Yan L, Yan S, Qin T, Shen J, Zha J (2021) Estimating aquatic plant diversity and distribution in rivers from Jingjinji region, China, using environmental DNA metabarcoding and a traditional survey method. Environ. Res. 199, 111348. https://doi.org/10.1016/j.envres.2021.111348

Johnson MD, Cox RD, Barnes MA (2019) Analyzing airborne environmental DNA: A comparison of extraction methods, primer type, and trap type on the ability to detect airborne eDNA from terrestrial plant communities. Environmental DNA 1, 176–185. https://doi.org/10.1002/edn3.19

Jorns T, Craine J, Towne EG, Knox M (2020) Climate structures bison dietary quality and composition at the continental scale. Environmental DNA 2, 77–90. https://doi.org/10.1002/edn3.47

Kalko EKV, Handley CO, Handley D (1996) Organization, diversity, and long-term dynamics of a neotropical bat community, in: Cody, M.L., Smallwood, J.A. (Eds.), Long-Term Studies of Vertebrate Communities. Elsevier, pp. 503–553. https://doi.org/10.1016/B978-012178075-3/50017-9

Kartzinel TR, Chen PA, Coverdale TC, Erickson DL, Kress WJ, Kuzmina ML, Rubenstein DI, Wang W, Pringle RM (2015) DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. Proc Natl Acad Sci USA 112, 8019–8024. https://doi.org/10.1073/pnas.1503283112

Kemp J, López-Baucells A, Rocha R, Wangensteen OS, Andriatafika Z, Nair A, Cabeza M (2019) Bats as potential suppressors of multiple agricultural pests: a case study from Madagascar. Agriculture, Ecosystems & Environment 269, 88–96. https://doi.org/10.1016/j.agee.2018.09.027

Kersey PJ, Collemare J, Cockel C, Das D, Dulloo EM, Kelly LJ, Lettice E, Malécot V, Maxted N, Metheringham C, Thormann I, Leitch IJ (2020) Selecting for useful properties of plants and fungi – Novel approaches, opportunities, and challenges. Plants, People, Planet 2, 409–420. https://doi.org/10.1002/ppp3.10136

Kier G, Mutke J, Dinerstein E, Ricketts TH, Küper W, Kreft H, Barthlott W (2005) Global patterns of plant diversity and floristic knowledge. J. Biogeogr. 32, 1107–1116. https://doi.org/10.1111/j.1365-2699.2005.01272.x

Koizumi N, Mori A, Mineta T, Sawada E, Watabe K, Takemura T (2016) Exploratory environmental DNA analysis for investigating plant-feeding habit of the red-eared turtle using their feces samples. JT 78, 9–13. https://doi.org/10.11113/jt.v78.7253

Koyama A, Egawa C, Taki H, Yasuda M, Kanzaki N, Ide T, Okabe K (2018) Non-native plants are a seasonal pollen source for native honeybees in suburban ecosystems. Urban Ecosyst. 21, 1113–1122. https://doi.org/10.1007/s11252-018-0793-3

Kraaijeveld K, de Weger LA, Ventayol García M, Buermans H, Frank J, Hiemstra PS, den Dunnen JT (2015) Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. Mol. Ecol. Resour. 15, 8–16. https://doi.org/10.1111/1755-0998.12288

Kuzmina ML, Braukmann TWA, Zakharov EV (2018) Finding the pond through the weeds: eDNA reveals underestimated diversity of pondweeds. Appl. Plant Sci. 6, e01155. https://doi.org/10.1002/aps3.1155

Lacoursière-Roussel A, Deiner K (2021) Environmental DNA is not the tool by itself. J. Fish Biol. 98, 383–386. https://doi.org/10.1111/jfb.14177

Lamb EG, Winsley T, Piper CL, Freidrich SA, Siciliano SD (2016) A high-throughput belowground plant diversity assay using next-generation sequencing of the trnL intron. Plant Soil 404, 361–372. https://doi.org/10.1007/s11104-016-2852-y

Lentz DL, Hamilton TL, Dunning NP, Tepe EJ, Scarborough VL, Meyers SA, Grazioso L, Weiss AA (2021) Environmental DNA reveals arboreal cityscapes at the Ancient Maya Center of Tikal. Sci. Rep. 11, 12725. https://doi.org/10.1038/s41598-021-91620-6

Leontidou K, Vokou D, Sandionigi A, Bruno A, Lazarina M, De Groeve J, Li M, Varotto C, Girardi M, Casiraghi M, Cristofori A (2021) Plant biodiversity assessment through pollen DNA metabarcoding in Natura 2000 habitats (Italian Alps). Sci. Rep. 11, 18226. https://doi.org/10.1038/s41598-021-97619-3

Lim V-C, Ramli R, Bhassu S, Wilson J-J (2018) Pollination implications of the diverse diet of tropical nectar-feeding bats roosting in an urban cave. PeerJ 6, e4572. https://doi.org/10.7717/peerj.4572

Liu S, Li K, Jia W, Stoof-Leichsenring KR, Liu X, Cao X, Herzschuh U (2021) Vegetation reconstruction from Siberia and the Tibetan plateau using modern analogue technique–comparing sedimentary (ancient) DNA and pollen data. Front. Ecol. Evol. 9, 668611. https://doi.org/10.3389/fevo.2021.668611

Lopes CM, Baêta D, Sasso T, Vanzetti A, Raquel Zamudio K, Taberlet P, Haddad CFB (2021) Power and limitations of environmental DNA metabarcoding for surveying leaf litter eukaryotic communities. Environmental DNA 3, 528–540. https://doi.org/10.1002/edn3.142

Marčiulynienė D, Marčiulynas A, Lynikienė J, Vaičiukynė M, Gedminas A, Menkis A (2021) DNA-metabarcoding of belowground fungal communities in bare-root forest nurseries: focus on different tree species. Microorganisms 9, 150. https://doi.org/10.3390/microorganisms9010150

Mariani S, Baillie C, Colosimo G, Riesgo A (2019) Sponges as natural environmental DNA samplers. Curr. Biol. 29, R401–R402. https://doi.org/10.1016/j.cub.2019.04.031

Matesanz S, Pescador DS, Pías B, Sánchez AM, Chacón-Labella J, Illuminati A, de la Cruz M, López-Angulo J, Marí-Mena N, Vizcaíno A, Escudero A (2019) Estimating belowground plant abundance with DNA metabarcoding. Mol. Ecol. Resour. 19, 1265–1277. https://doi.org/10.1111/1755-0998.13049

Matsuhashi S, Doi H, Fujiwara A, Watanabe S, Minamoto T (2016) Evaluation of the environmental DNA method for estimating distribution and biomass of submerged aquatic plants. PLoS ONE 11, e0156217. https://doi.org/10.1371/journal.pone.0156217

Mauvisseau Q, Harper LR, Sander M, Hanner RH, Kleyer H, Deiner K (2022) The multiple states of environmental DNA and what is known about their persistence in aquatic environments. Environ. Sci. Technol. 56, 5322–5333. https://doi.org/10.1021/acs.est.1c07638

McFrederick QS, Rehan SM (2016) Characterization of pollen and bacterial community composition in brood provisions of a small carpenter bee. Mol. Ecol. 25, 2302–2311. https://doi.org/10.1111/mec.13608

Milberg P, Bergstedt J, Fridman J, Odell G, Westerberg L (2008) Observer bias and random variation in vegetation monitoring data. Journal of Vegetation Science 19, 633–644. https://doi.org/10.3170/2008-8-18423

Milla L, Sniderman K, Lines R, Mousavi-Derazmahalleh M, Encinas-Viso F (2021) Pollen DNA metabarcoding identifies regional provenance and high plant diversity in Australian honey. Ecol. Evol. 11, 8683–8698. https://doi.org/10.1002/ece3.7679

Montauban C, Mas M, Wangensteen OS, Sarto i Monteys V, Fornós DG, Mola XF, López-Baucells A (2021) Bats as natural samplers: first record of the invasive pest rice water weevil *Lissorhoptrus oryzophilus* in the Iberian Peninsula. Crop Prot. 141, 105427. https://doi.org/10.1016/j.cropro.2020.105427

Mori E, Mazza G, Galimberti A, Angiolini C, Bonari G (2017) The porcupine as "little thumbling": the role of *Hystrix cristata* in the spread of *Helianthus tuberosus*. Biologia 72, 1211–1216. https://doi.org/10.1515/biolog-2017-0136

Mucina L (2019) Biome: evolution of a crucial ecological and biogeographical concept. New Phytol. 222, 97–114. https://doi.org/10.1111/nph.15609

Nagler M, Insam H, Pietramellara G, Ascher-Jenull J (2018) Extracellular DNA in natural environments: features, relevance and applications. Appl. Microbiol. Biotechnol. 102, 6343–6356. https://doi.org/10.1007/s00253-018-9120-4

Núñez A, Amo de Paz G, Ferencova Z, Rastrojo A, Guantes R, García AM, Alcamí A, Gutiérrez-Bustillo AM, Moreno DA (2017) Validation of the Hirst-type spore trap for simultaneous monitoring of prokaryotic and eukaryotic biodiversities in urban air samples by next-generation sequencing. Appl. Environ. Microbiol. 83, e00472-17. https://doi.org/10.1128/AEM.00472-17

Núñez A, Amo de Paz G, Rastrojo A, Ferencova Z, Gutiérrez-Bustillo AM, Alcamí A, Moreno DA, Guantes R (2019) Temporal patterns of variability for prokaryotic and eukaryotic diversity in the urban air of Madrid (Spain). Atmos. Environ. 217, 116972. https://doi.org/10.1016/j.atmosenv.2019.116972

Ortega A, Geraldi NR, Alam I, Kamau AA, Acinas SG, Logares R, Gasol JM, Massana R, Krause-Jensen D, Duarte CM (2019) Important contribution of macroalgae to oceanic carbon sequestration. Nat. Geosci. 12, 748–754. https://doi.org/10.1038/s41561-019-0421-8

Ortega A, Geraldi NR, Díaz-Rúa R, Ørberg SB, Wesselmann M, Krause-Jensen D, Duarte CM (2020a) A DNA mini-barcode for marine macrophytes. Mol. Ecol. Resour. 20, 920–935. https://doi.org/10.1111/1755-0998.13164

Ortega A, Geraldi NR, Duarte CM (2020b) EnvironmentalDNA identifies marine macrophyte contributions to Blue Carbon sediments. Limnol. Oceanogr. 65, 3139–3149. https://doi.org/10.1002/lno.11579

Osathanunkul M, Sawongta N, Pheera W, Pechlivanis N, Psomopoulos F, Madesis P (2021) Exploring plant diversity through soil DNA in Thai national parks for influencing land reform and agriculture planning. PeerJ 9, e11753. https://doi.org/10.7717/peerj.11753

Palacios Mejia M, Curd E, Edalati K, Renshaw MA, Dunn R, Potter D, Fraga N, Moore J, Saiz J, Wayne R, Parker SS (2021) The utility of environmental DNA from sediment and water samples for recovery of observed plant and animal species from four Mojave Desert springs. Environmental DNA 3, 214–230. https://doi.org/10.1002/edn3.161

Parducci L, Bennett KD, Ficetola GF, Alsos IG, Suyama Y, Wood JR, Pedersen MW (2017) Ancient plant DNA in lake sediments. New Phytol. 214, 924–942. https://doi.org/10.1111/nph.14470

Parducci L, Jørgensen T, Tollefsrud MM, Elverland E, Alm T, Fontana SL, Bennett KD, Haile J, Matetovici I, Suyama Y, Edwards ME, Andersen K, Rasmussen M, Boessenkool S, Coissac E, Brochmann C, Taberlet P, Houmark-Nielsen M, Larsen NK, Orlando L, Willerslev E (2012) Glacial survival of boreal trees in northern Scandinavia. Science 335, 1083–1086. https://doi.org/10.1126/science.1216043

Pawlowski J, Apothéloz-Perret-Gentil L, Altermatt F (2020) Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. Mol. Ecol. 29, 4258–4264. https://doi.org/10.1111/mec.15643

Pedersen MW, Overballe-Petersen S, Ermini L, Sarkissian CD, Haile J, Hellstrom M, Spens J, Thomsen PF, Bohmann K, Cappellini E, Schnell IB, Wales NA, Carøe C, Campos PF, Schmidt AMZ, Gilbert MTP, Hansen AJ, Orlando L, Willerslev E (2015) Ancient and modern environmental DNA. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370, 20130383. https://doi.org/10.1098/rstb.2013.0383

Pedersen MW, Ruter A, Schweger C, Friebe H, Staff RA, Kjeldsen KK, Mendoza MLZ, Beaudoin AB, Zutter C, Larsen NK, Potter BA, Nielsen R, Rainville RA, Orlando L, Meltzer DJ, Kjær KH, Willerslev E (2016) Postglacial viability and colonization in North America's ice-free corridor. Nature 537, 45–49. https://doi.org/10.1038/nature19085

Pemberton RW, Wheeler GS (2006) Orchid bees don't need orchids: evidence from the naturalization of an orchid bee in Florida. Ecology 87, 1995–2001. https://doi.org/10.1890/0012-9658(2006)87[1995:obdnoe]2.0.co;2

Pietramellara G, Ascher J, Borgogni F, Ceccherini MT, Guerri G, Nannipieri P (2009) Extracellular DNA in soil and sediment: fate and ecological relevance. Biol. Fertil. Soils 45, 219–235. https://doi.org/10.1007/s00374-008-0345-8

Polling M, Sin M, de Weger LA, Speksnijder AGCL, Koenders MJF, de Boer H, Gravendeel B (2022) DNA metabarcoding using nrITS2 provides highly qualitative and quantitative results for airborne pollen monitoring. Sci. Total Environ. 806, 150468. https://doi.org/10.1016/j.scitotenv.2021.150468

Polling M, ter Schure ATM, van Geel B, van Bokhoven T, Boessenkool S, MacKay G, Langeveld BW, Ariza M, van der Plicht H, Protopopov AV, Tikhonov A, de Boer H, Gravendeel B (2021) Multiproxy analysis of permafrost preserved faeces provides an unprecedented insight into the diets and habitats of extinct and extant megafauna. Quat. Sci. Rev. 267, 107084. https://doi.org/10.1016/j.quascirev.2021.107084

Prosser SWJ, Hebert PDN (2017) Rapid identification of the botanical and entomological sources of honey using DNA metabarcoding. Food Chem. 214, 183–191. https://doi.org/10.1016/j.foodchem.2016.07.077

Ramírez SR, Eltz T, Fujiwara MK, Gerlach G, Goldman-Huertas B, Tsutsui ND, Pierce NE (2011) Asynchronous diversification in a specialized plant-pollinator mutualism. Science 333, 1742–1746. https://doi.org/10.1126/science.1209175

Richardson RT, Lin C-H, Sponsler DB, Quijia JO, Goodell K, Johnson RM (2015) Application of ITS2 metabarcoding to determine the provenance of pollen collected by honey bees in an agroecosystem. Appl. Plant Sci. 3, 1400066. https://doi.org/10.3732/apps.1400066

Ritter CD, Zizka A, Roger F, Tuomisto H, Barnes C, Nilsson RH, Antonelli A (2018) High-throughput metabarcoding reveals the effect of physicochemical soil properties on soil and litter biodiversity and community turnover across Amazonia. PeerJ 6, e5661. https://doi.org/10.7717/peerj.5661

Rodriguez-Ezpeleta N, Morissette O, Bean CW, Manu S, Banerjee P, Lacoursière-Roussel A, Beng KC, Alter SE, Roger F, Holman LE, Stewart KA, Monaghan MT, Mauvisseau Q, Mirimin L, Wangensteen OS, Antognazza CM, Helyar SJ, de Boer H, Monchamp M-E, Nijland R, Deiner K (2021) Trade-offs between reducing complex terminology and producing accurate interpretations from environmental DNA: comment on "Environmental DNA: what's behind the term?" by Pawlowski et al., (2020). Mol. Ecol. 30, 4601–4605. https://doi.org/10.1111/mec.15942

Rodríguez-Ezpeleta N, Zinger L, Kinziger A, Bik HM, Bonin A, Coissac E, Emerson BC, Lopes CM, Pelletier TA, Taberlet P, Narum S (2021) Biodiversity monitoring using environmental DNA. Mol. Ecol. Resour. 21, 1405–1409. https://doi.org/10.1111/1755-0998.13399

Rowney FM, Brennan GL, Skjøth CA, Griffith GW, McInnes RN, Clewlow Y, Adams-Groom B, Barber A, de Vere N, Economou T, Hegarty M, Hanlon HM, Jones L, Kurganskiy A, Petch GM, Potter C, Rafiq AM, Warner A, PollerGEN Consortium Wheeler B, Creer S (2021) Environmental DNA reveals links between abundance and composition of airborne grass pollen and respiratory health. Curr. Biol. 31, 1995-2003.e4. https://doi.org/10.1016/j.cub.2021.02.019

Rucińska A, Świerszcz S, Nobis M, Zubek S, Boczkowska M, Olszak M, Kosiński JG, Nowak S, Nowak A (2022) Is it possible to understand a book missing a quarter of the letters? Unveiling the belowground species richness of grasslands. Agriculture, Ecosystems & Environment 324, 107683. https://doi.org/10.1016/j.agee.2021.107683

Schure ATM, Pillai AAS, Thorbek L, Bhavani Shankar M, Puri R, Ravikanth G, Boer HJ, Boessenkool S (2021) eDNA metabarcoding reveals dietary niche overlap among herbivores in an Indian wildlife sanctuary. Environmental DNA 3, 681–696. https://doi.org/10.1002/edn3.168

Scott WA, Hallam CJ (2003) Assessing species misidentification rates through quality assurance of vegetation monitoring. Plant Ecology 165, 101–115.

Scriver M, Marinich A, Wilson C, Freeland J (2015) Development of species-specific environmental DNA (eDNA) markers for invasive aquatic plants. Aquatic Botany 122, 27–31. https://doi.org/10.1016/j.aquabot.2015.01.003

Sepp S, Davison J, Moora M, Neuenkamp L, Oja J, Roslin T, Vasar M, Öpik M, Zobel M (2021) Woody encroachment in grassland elicits complex changes in the functional structure of above- and belowground biota. Ecosphere 12, e03512. https://doi.org/10.1002/ecs2.3512

Sherwood AR, Dittbern MN, Johnston ET, Conklin KY (2017) A metabarcoding comparison of windward and leeward airborne algal diversity across the Ko'olau mountain range on the island of O'ahu, Hawai'i1. J. Phycol. 53, 437–445. https://doi.org/10.1111/jpy.12502

Shogren AJ, Tank JL, Egan SP, August O, Rosi EJ, Hanrahan BR, Renshaw MA, Gantz CA, Bolster D (2018) Water flow and biofilm cover influence environmental DNA detection in recirculating streams. Environ. Sci. Technol. 52, 8530–8537. https://doi.org/10.1021/acs.est.8b01822

Siegenthaler A, Wangensteen OS, Soto AZ, Benvenuto C, Corrigan L, Mariani S (2019) Metabarcoding of shrimp stomach content: Harnessing a natural sampler for fish biodiversity monitoring. Mol. Ecol. Resour. 19, 206–220. https://doi.org/10.1111/1755-0998.12956

Smart MD, Cornman RS, Iwanowicz DD, McDermott-Kubeczko M, Pettis JS, Spivak MS, Otto CRV (2017) A comparison of honey bee-collected pollen from working agricultural lands using light microscopy and ITS metabarcoding. Environ. Entomol. 46, 38–49. https://doi.org/10.1093/ee/nvw159

Sønstebø JH, Gielly L, Brysting AK, Elven R, Edwards M, Haile J, Willerslev E, Coissac E, Rioux D, Sannier J, Taberlet P, Brochmann C (2010) Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. Mol. Ecol. Resour. 10, 1009–1018. https://doi.org/10.1111/j.1755-0998.2010.02855.x

Sookhan N, Lorenzo A, Tatsumi S, Yuen M, MacIvor JS (2021) Linking bacterial diversity to floral identity in the bumble bee pollen basket. Environmental DNA 3, 669–680. https://doi.org/10.1002/edn3.165

Steinbauer MJ, Grytnes J-A, Jurasinski G, Kulonen A, Lenoir J, Pauli H, Rixen C, Winkler M, Bardy-Durchhalter M, Barni E, Bjorkman AD, Breiner FT, Burg S, Czortek P, Dawes MA, Delimat A, Dullinger S, Erschbamer B, Felde VA, Fernández-Arberas O, Wipf S (2018) Accelerated increase in plant species richness on mountain summits is linked to warming. Nature 556, 231–234. https://doi.org/10.1038/s41586-018-0005-6

Stoeck T, Pan H, Dully V, Forster D, Jung T (2018) Towards an eDNA metabarcode-based performance indicator for full-scale municipal wastewater treatment plants. Water Res. 144, 322–331. https://doi.org/10.1016/j.watres.2018.07.051

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. Mol. Ecol. 21, 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Res. 35, e14. https://doi.org/10.1093/nar/gkl938

Terwayet Bayouli I, Terwayet Bayouli H, Dell'Oca A, Meers E, Sun J (2021) Ecological indicators and bioindicator plant species for biomonitoring industrial pollution: eco-based environmental assessment. Ecological Indicators 125, 107508. https://doi.org/10.1016/j.ecolind.2021.107508

Thies W, Kalko EKV (2004) Phenology of neotropical pepper plants (Piperaceae) and their association with their main dispersers, two short-tailed fruit bats, *Carollia perspicillata* and *C. castanea* (Phyllostomidae). Oikos 104, 362–376. https://doi.org/10.1111/j.0030-1299.2004.12747.x

Thomsen PF, Sigsgaard EE (2019) Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods. Ecol. Evol. 9, 1665–1679. https://doi.org/10.1002/ece3.4809

Thomsen PF, Willerslev E (2015) Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

Tsukamoto Y, Yonezawa S, Katayama N, Isagi Y (2021) Detection of endangered aquatic plants in rapid streams using environmental DNA. Front. Ecol. Evol. 8. https://doi.org/10.3389/fevo.2020.622291

Turner CR, Barnes MA, Xu CCY, Jones SE, Jerde CL, Lodge DM (2014) Particle size distribution and optimal capture of aqueous macrobial eDNA. Methods Ecol. Evol. 5, 676–684. https://doi.org/10.1111/2041-210X.12206

Turon M, Angulo-Preckler C, Antich A, Præbel K, Wangensteen OS (2020) More than expected from old sponge samples: a natural sampler DNA metabarcoding assessment of marine fish diversity in Nha Trang bay (Vietnam). Front. Mar. Sci. 7, 605148. https://doi.org/10.3389/fmars.2020.605148

Turon X, Antich A, Palacín C, Praebel K, Wangensteen OS (2020) From metabarcoding to metaphylogeography: separating the wheat from the chaff. Ecol. Appl. 30, e02036. https://doi.org/10.1002/eap.2036

Uuemaa E, Mander Ü, Marja R (2013) Trends in the use of landscape spatial metrics as landscape indicators: a review. Ecological Indicators 28, 100–106. https://doi.org/10.1016/j.ecolind.2012.07.018

Valentini A, Miquel C, Nawaz MA, Bellemain E, Coissac E, Pompanon F, Gielly L, Cruaud C, Nascetti G, Wincker P, Swenson JE, Taberlet P (2009) New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. Mol. Ecol. Resour. 9, 51–60. https://doi.org/10.1111/j.1755-0998.2008.02352.x

Valentini A (2007) Non-invasive diet analysis based on DNA barcoding: the Himalayan brown bears (*Ursus arctos isabellinus*) as a case study (Doctoral dissertation).

Varotto C, Pindo M, Bertoni E, Casarotto C, Camin F, Girardi M, Maggi V, Cristofori A (2021) A pilot study of eDNA metabarcoding to estimate plant biodiversity by an alpine glacier core (Adamello glacier, North Italy). Sci. Rep. 11, 1208. https://doi.org/10.1038/s41598-020-79738-5

Vasconcelos S, Nunes GL, Dias MC, Lorena J, Oliveira RRM, Lima TGL, Pires ES, Valadares RBS, Alves R, Watanabe MTC, Zappi DC, Hiura AL, Pastore M, Vasconcelos LV, Mota NFO, Viana PL, Gil ASB, Simões AO, Imperatriz-Fonseca VL, Harley RM, Oliveira G (2021) Unraveling the plant diversity of the Amazonian canga through DNA barcoding. Ecol. Evol. 11, 13348–13362. https://doi.org/10.1002/ece3.8057

Willerslev E, Cappellini E, Boomsma W, Nielsen R, Hebsgaard MB, Brand TB, Hofreiter M, Bunce M, Poinar HN, Dahl-Jensen D, Johnsen S, Steffensen JP, Bennike O, Schwenninger J-L, Nathan R, Armitage S, de Hoog C-J, Alfimov V, Christl M, Beer J, Collins MJ (2007) Ancient biomolecules from deep ice cores reveal a forested southern Greenland. Science 317, 111–114. https://doi.org/10.1126/science.1141758

Willerslev E, Davison J, Moora M, Zobel M, Coissac E, Edwards ME, Lorenzen ED, Vestergård M, Gussarova G, Haile J, Craine J, Gielly L, Boessenkool S, Epp LS, Pearman PB, Cheddadi R, Murray D, Bråthen KA, Yoccoz N, Binney H, Taberlet P (2014) Fifty thousand years of Arctic vegetation and megafaunal diet. Nature 506, 47–51. https://doi.org/10.1038/nature12921

Willerslev E, Hansen AJ, Binladen J, Brand TB, Gilbert MTP, Shapiro B, Bunce M, Wiuf C, Gilichinsky DA, Cooper A (2003) Diverse plant and animal genetic records from Holocene and Pleistocene sediments. Science 300, 791–795. https://doi.org/10.1126/science.1084114

Wood JM (1987) Biological processes involved in the cycling of elements between soil or sediments and the aqueous environment. Hydrobiologia 149, 31–42. https://doi.org/10.1007/BF00048644

Yamamoto S, Uchida K (2018) A generalist herbivore requires a wide array of plant species to maintain its populations. Biological Conservation 228, 167–174. https://doi.org/10.1016/j.biocon.2018.10.018

Yang Chenxue Wang X, Miller JA, de Blécourt M, Ji Y, Yang Chunyan Harrison RD, Yu DW (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

Yoccoz NG, Bråthen KA, Gielly L, Haile J, Edwards ME, Goslar T, Von Stedingk H, Brysting AK, Coissac E, Pompanon F, Sønstebø JH, Miquel C, Valentini A, De Bello F, Chave J, Thuiller W, Wincker P, Cruaud C, Gavory F, Rasmussen M, Taberlet P (2012) DNA from soil mirrors plant taxonomic and growth form diversity. Mol. Ecol. 21, 3647–3655. https://doi.org/10.1111/j.1365-294X.2012.05545.x

Yoccoz NG (2012) The future of environmental DNA in ecology. Mol. Ecol. 21, 2031–2038. https://doi.org/10.1111/j.1365-294X.2012.05505.x

Zinger L, Taberlet P, Schimann H, Bonin A, Boyer F, De Barba M, Gaucher P, Gielly L, Giguet-Covex C, Iribar A, Réjou-Méchain M, Rayé G, Rioux D, Schilling V, Tymen B, Viers J, Zouiten C, Thuiller W, Coissac E, Chave J (2019) Body size determines soil community assembly in a tropical forest. Mol. Ecol. 28, 528–543. https://doi.org/10.1111/mec.14919

Answers

1. The analysis of eDNA from gut contents, faeces, or eDNA traces from the butterfly's body (vegetation fragments or pollen grains) would reveal the floral resources available and visited. To reveal other organisms that are using the same floral resources (other pollinators competitors), one could target insect eDNA present in flowers that have been visited.
2. eDNA water samples from near-shore sites would optimise the vegetation assessment as they are both easy to collect and signal terrestrial and aquatic diversity. Though airborne DNA could be also considered for this purpose, it may miss dormant DNA or non wind-dispersed plants. In addition, sedimentary eDNA may also signal nearby diversity.
3. Spruce and pine spores are tiny, light, and spread by wind. These have a tendency to show up anywhere, and are not a good indication for local presence. Invasive species monitoring needs approaches that provide a clear link between detected species and specific environments.

Paper II

Chapter 4
DNA from soil

Maria Ariza Salazar[1], Sandra Garcés-Pastor[2], Hugo de Boer[1]

1 Natural History Museum, University of Oslo, Norway
2 The Arctic University Museum of Norway, UiT - The Arctic University of Norway, Tromsø, Norway

Maria Ariza Salazar **mariadelosangelesariza@gmail.com**
Sandra Garcés-Pastor **sandra.garces-pastor@uit.no**
Hugo de Boer **h.de.boer@nhm.uio.no**

Citation: Salazar MA, Garcés-Pastor S, de Boer H (2022) Chapter 4. DNA from soil. In: de Boer H, Rydmark MO, Verstraete B, Gravendeel B (Eds) Molecular identification of plants: from sequence to species. Advanced Books. https://doi.org/10.3897/ab.e98875

Introduction

The natural presence of any plant entails the existence of a substrate where it can anchor itself and absorb nutrients for its development and survival (Wardle et al. 2004). This is most commonly the ground and specifically, soil. Nevertheless, the link between soil and plants goes beyond soil supporting plants as plants are one of the main soil-forming forces of pedogenesis through the accumulation of organic matter as well as modification of the soil biochemistry surrounding the roots (Corti et al. 2005). This process over time leads to the formation of soil layers, termed horizons, that can commonly be visibly identified (Schulz et al. 2013; Shlemon 1985; Vogt et al. 1995). Near the ground surface, the first soil horizon is an organic layer composed of growing roots and decomposing vegetative and reproductive plant material from local or regional origins, i.e., fallen debris, pollen particles, seeds (Vogt et al. 1995). Hence, this soil horizon is particularly rich in plant DNA from the environment (soil eDNA in short; Taberlet et al. 2018) and can be used as a proxy for plant identification and other biodiversity assessments (Fahner et al. 2016; Taberlet et al. 2018; Yoccoz et al. 2012).

Since the first isolation of DNA from soil bacteria, soil eDNA has gained attention for the assessment of terrestrial environments for several reasons: soil is virtually everywhere, it is easy to collect and transport, harbors signals from above and below biota including both active and dormant cells, and is a non-invasive sample collection technique (Torsvik et al. 1990; Yoccoz 2012); for more on soil eDNA applications see Chapter 24 Environment and biodiversity assessments). Soil eDNA assessments targeting modern plant diversity commonly employ samples that are collected near the surface (organic horizont). However, some studies may refer to sediments which can lead to confusing eDNA samples coexisting in underground environments (Kristensen and Rabenhorst 2015). Although both soil and sediments are products of mineral weathering (Wood 1987), in soils the deposition of these products happens in situ and remains on the surface, while in sediments these products are transported and redeposited elsewhere in layers over time, e.g., the ground or the bottom of a lake or stream (Burdige 2020). Moreover, sediments in general have very different organic content, particle size and mineralogy, and lesser organismal activity than soil, although the transition from soil to sediment can be gradual and depends on the eco-physiological characteristics of the regional environment (e.g., tropical vs. boreal forest; Shackley 1975; Smol et al. 2001). Yet, during flooding events sediments can be transported very rapidly from one place to another while sedimenting in new layers mixed with soil (Baldwin and Mitchell 2000). In these contexts, soil and sedimentary eDNA samples may have a mix of different spatio-temporal signals when it comes to the reconstruction of terrestrial or aquatic environments (Deiner et al. 2017; Thomsen and Willerslev 2015). Ancient sedimentary DNA (sedaDNA) is commonly sampled from bottom sediment layers in either aquatic or terrestrial environments (Parducci et al. 2018), and its temporal signal is usually correlated with sampling depth (Willerslev and Cooper 2005). For more on sedaDNA and its applications see Chapter 8 aDNA from sediments. Sedimentary DNA (sedDNA) usually refers to modern sediments that were either recently deposited or signal contemporary environments. Plant biodiversity assessments of modern environments often employ surface lake sediments (Andersen et al. 2012; Pedersen et al. 2015; Willerslev et al. 2014) as it captures current biodiversity from the entire watershed catchment area (Alsos et al. 2018). This chapter focuses on modern DNA isolated from soil eDNA.

Further, studies may also refer to bulk soil DNA when using soil samples to identify unknown communities, especially in forensic contexts (Boggs et al. 2019; Gothwal et al. 2007; Meiklejohn et al. 2018). Bulk DNA is commonly used in contexts where known taxa are mixed, molecularly identified (usually by metabarcoding), and then studied. There is no consensus

on the precise use of these different terms, and the terminology often reflects disciplinary backgrounds and study approaches (Kristensen and Rabenhorst 2015). Yet, it is worth noting that all terms mentioned so far are not mutually exclusive nor encapsulate a particular environment. For example, soil may also be used in aquatic contexts when pedogenic processes lead to horizon differentiation, e.g., estuarine substrata (Wardle et al. 2004). Thus, careful interpretation of the context in which the term is employed is recommended to ensure correct interpretation of data and studies.

Soil DNA: degradation, persistence, and decay

Molecular (plant) identification using soil or sediment eDNA relies on isolating DNA traces from roots, debris, seeds, and pollen (Levy-Booth et al. 2007), which signal diverse spatial and temporal origins, i.e., local or regional, ancient or contemporary. When these plant parts settle into the ground, DNA can be present either in intact cells (intracellular DNA or iDNA) or free in the environment following cell lysis or rupture (extracellular DNA or exDNA; Nagler et al. 2018). The largest fraction of eDNA in underground environments is exDNA that originates from bacteria and fungal soil communities (Levy-Booth et al. 2007; Nagler et al. 2018; Pietramellara et al. 2009; Poté et al. 2009).

The state of DNA in the soil is subject to intrinsic and extrinsic DNA properties related to the origins of the DNA as well as factors influencing its decay (Barnes et al. 2014; Lacoursière-Roussel and Deiner 2021; Sirois and Buckley 2019). For more on leaf DNA decay together with organic horizon formation, see the infographic. Soil eDNA is therefore a combination of iDNA and exDNA, that can degrade rapidly or persist over time. Intrinsic DNA properties that can affect its persistence in the ground include characteristics such as DNA GC content, purity, and weight (Nielsen et al. 2000; Pietramellara et al. 2009; Sirois and Buckley 2019; Taberlet et al. 2018; Vuillemin et al. 2017). Intrinsic DNA properties are those of the organism that affect the magnitude of DNA deposition such as life history traits like biomass, feeding, social, nesting, burrowing, hibernation, etc. Extrinsic DNA properties are more related to abiotic and biotic processes operating in the ground, e.g., soil mineralogy, organic components, pH, electrostatic properties, moisture, the presence/absence of UV radiation, bioturbation**,** enzymatic activity by microbial communities, and decomposition (Cozzolino et al. 2007; Gardner and Gunsch 2017; Gulden et al. 2005; Levy-Booth et al. 2007; Prosser and Hedgpeth 2018; Saeki et al. 2011). Examples of biotic processes operating in natural environments can be found in the infographic.

iDNA persists due to protection from the cell wall and membranes against abiotic processes. Cells are more likely to remain intact in the ground if there is decreased enzymatic activity as a result of rapid soil desiccation, low temperatures, or extreme pH values (Pietramellara et al. 2009; Taberlet et al. 2018). exDNA is more likely to persist when it binds to surface-reactive particles and hydrophobic soil components such as clay, sand, silt, and humic acids (Levy-Booth et al. 2007; Pietramellara et al. 2009). DNA may also indirectly persist via bacterial integration of DNA fragments (Levy-Booth et al. 2007). Bacterial enzymatic activity plays a central role in DNA degradation in soil (Blum et al. 1997). DNase is secreted copiously to access the phosphorus and nitrogen from the DNA and acts more rapidly on DNA at higher temperatures (Levy-Booth et al. 2007). Since both the temperature and underground biota activity levels are higher in tropical climates, there are generally increased degradation rates in tropical vs. boreal soils. Soil types may also affect degradation rates (Sirois and Buckley 2019) using a controlled microcosm reported that synthetic DNA degraded slower in forest than in agricultural soils where tillage and other disruptive processes can affect persistence. Predicting the origins and persistence of

eDNA remains a thorny issue, mainly because of the complex nature of the properties involved (Barnes and Turner 2015; Deiner et al. 2017).

Soil memory

Plant eDNA bound to soil particles can originate from multiple taxa and multiple vegetative parts, each one with particular mechanisms to bind, persist and degrade in soil substrates. Plant DNA persistence within soil allows us to harvest its botanical memory for identifying vegetation through time. Indeed, comparisons of plant identifications through both visual vegetation surveys and soil eDNA assessments have shed light on the temporal signals stored in top soils. In boreal areas, plant identification through soil eDNA signal mostly registered contemporary vegetation (Ariza et al. 2022; Edwards et al. 2018; Yoccoz et al. 2012), however, taxa surveyed up to 30 years ago was also reported, suggesting that soil eDNA harbors more of a contemporary memory (Ariza et al. 2022). The extent of this memory effect across soil types and environments is poorly understood while its implications are relevant for society (e.g., biodiversity assessments and monitoring, forensics, biosafety). For more on applications of soil eDNA see Chapter 24 Environment and biodiversity assessments.

Designing a soil eDNA study

The flora and study area are key in any study to ensure sound conclusions. Below you will find considerations that can help you to answer common questions when designing field and wet lab experiments.

How to sample and how much?

Soil sampling can be done either by scooping out the soil, drilling down a tube, i.e., a 50 ml falcon tube, or with a soil core sampler. We recommend to use sampling protocols specifically validated in an environment similar to your study site, e.g., woodlands, grasslands, meadows, boreal temperate, and tropical forest (Bienert et al. 2012; Dopheide et al. 2019; Fahner et al. 2016; Taberlet et al. 2012; Yoccoz et al. 2012). It is also recommended to sample in flat areas as slopes can cause erosion and colluvium that can interfere with soil stratification. Soil and sedimentary particles are deposited in sequence, thus we can expect the bottom soil horizons to harbor older eDNA signals than those at the top. However, mixing across vertical layers can be expected as a result of bioturbation, and it is thus very important to assess the stratigraphy of the soil/sediment that is being investigated. If bioturbation is absent, sampling specific soil horizons can thus be used to capture vegetation with particular time signals (Dickie et al. 2018). Similarly, the amount of soil collected, as well as the number of samples and replicates, can affect the spatial and time signal captured (Calderón-Sanou et al. 2020; Dopheide et al. 2019; Taberlet et al. 2012; Zinger et al. 2019a). We recommend sampling at least 10 g of soil, but power analysis and rarefaction curves can aid to determine and optimize this parameter (Dickie et al. 2018; Dopheide et al. 2019). If one prefers to reduce the effect of local heterogeneity in the sampling strategy, several dozens of subsamples (between 20 and 50 g) can be mixed (Dickie et al. 2018; Taberlet et al. 2012). This strategy is however not suitable for studies dealing with patterns at small spatial scales (< 1 m^2; Edwards et al. 2018).

How to process the soil samples?

Obtaining clean DNA samples as well as avoiding cross contamination is challenging when sampling soil eDNA. Collection instruments should therefore be decontaminated between each sample (e.g., flaming, chlorine cleaning), gloves and masks should be worn and changed regularly to avoid introduction of DNA, and samples should be stored in separate plastic bags. In order to stop (or greatly reduce) enzymatic activity, samples should be stored cold or frozen, preferably at -20 °C, if immediate sample processing is not possible (Taberlet et al. 2012). Post-collection treatment of soil samples can also include air drying or freeze-drying to stop enzymatic activity and preserve DNA integrity in the sample (Nocker et al. 2012; Ritter et al. 2018). Soil samples are usually a mix of both above and below ground fragments of fauna and flora, i.e., debris, manure, roots, seeds, pollen, insects, and worms. DNA from organisms that are present in large total biomass may complicate detection of DNA signals from rare organisms. Thus, particularly for plant identification studies, it is worth considering whether root and leaf fragments should be sieved out from the soil samples. This will also contribute towards amplifying the signal from those low abundant taxa and normalize amplifications for all organisms present in a sample.

Extraction of iDNA or exDNA?

DNA extraction is a key bottleneck when capturing molecular data, and protocols need to be tailored to both the study area and the question(s). At a minimum, you need to decide which fraction of the total soil eDNA (iDNA or exDNA) you want to isolate to answer your research question. In general, isolating exDNA is preferred when targeting non-microorganisms and avoiding diversity patterns across short temporal scales (Taberlet et al. 2012; Zinger et al. 2009). While both extraction protocols are generally similar, iDNA extraction requires a cell lysis step. Breaking the cell wall or pollen exine can be achieved with soil grinding, sonication, thermal shocks, or chemical treatments (Frostegård et al. 1999; Zhou et al. 2007). For DNA extraction protocols specifically for pollen DNA, see Chapter 5 DNA from pollen. Commercial kits for DNA extraction are readily available for joint or separate extraction of iDNA and exDNA from soil, and these are commonly used in soil eDNA studies (Alsos et al. 2018; Edwards et al. 2018; Fahner et al. 2016; Foucher et al. 2020; Yoccoz et al. 2012; Zinger et al. 2019b). Taberlet et al. (2012) proposed an extraction protocol targeting exDNA that is suitable for tropical and nontropical areas, and can be performed with material that is commonly found in molecular laboratories. Depending on the soil properties in your study area, you can adapt commercial kits to increase the quality and quantity of DNA. For example, adding chloroform can increase the separation of the organic phase and aqueous phase, which in turn optimizes DNA quality (Fatima et al. 2014). However, chloroform is highly abrasive and can induce cell lysis. Alternatively, slightly alkaline solutions of phosphate buffers can remove soil particles to which exDNA might be bound while simultaneously preventing lysis of the cells (Nagler et al. 2018).

Which DNA marker(s) to use?

If (meta)barcoding is used for identification, there are three desired features for a barcode in any study: sufficient polymorphism for identification at the desired taxonomic resolution, conserved primer binding sites for universal amplification, and available reference sequences for

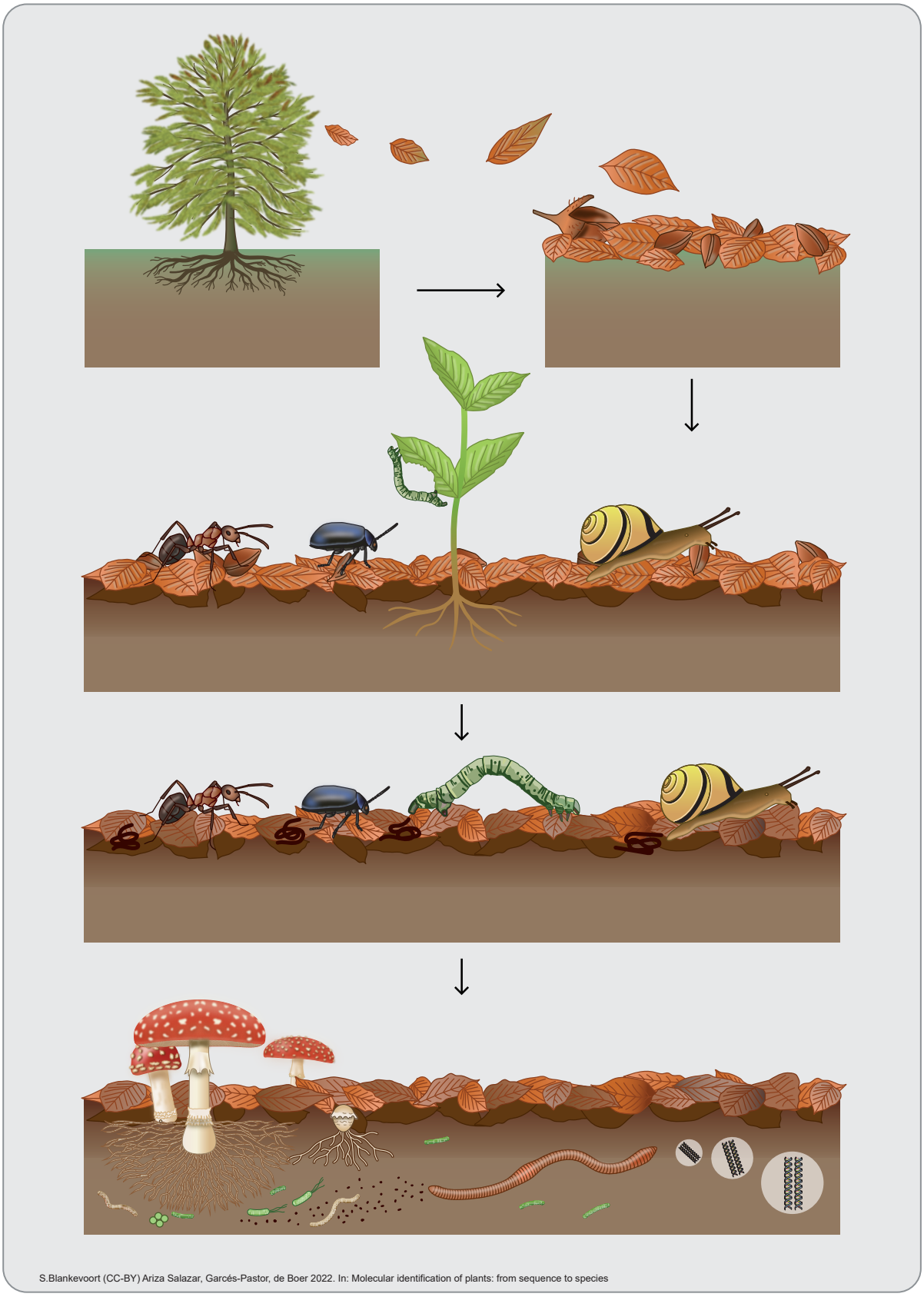

Figure 1. Chapter 4 Infographic: From leaf DNA to soil environmental DNA. One of the ways in which plant DNA is deposited in soil surfaces is through the accumulation of fallen leaves from trees.

the target organism. In many cases, not all features can be met. You may therefore need to decide on which features are most important for your research question. For more general information about choosing suitable markers and available reference databases, see Chapter 10 DNA barcoding and Chapter 11 Amplicon metabarcoding. Soil eDNA studies targeting plants have used markers found in chloroplast DNA (*trnL* P6 loop, *matK*, *rbcL*) and in ribosomal DNA (ITS2; Epp et al. 2018; Fahner et al. 2016; Yoccoz et al. 2012). However, metagenomic and target enrichment approaches are also starting to gain popularity as these avoid bias by PCR amplification and reduce the noise from non-target organisms (Johnson et al. 2019; Murchie et al. 2021). Fahner et al. (2016) compared the performance of plant barcodes (long vs. short barcodes) and recommended ITS2 and *rbcL* when identifying plants through soil eDNA metabarcoding, because these outperformed other markers in terms of recovery, reference completeness and identification resolution. Since the nuclear region, ITS2, is shared across plants and fungi, and the latter are abundantly present in soil, increased amplification of fungi can be expected. To avoid this, plant-specific primers targeting these regions can be used (Cheng et al. 2016). Furthermore, to avoid biased assessments towards particular plant groups when using ITS2, i.e., flowering plants or mosses, a combination of both TS2F/ITSp4 and ITSp3/ITSu4 primers pairs, is recommended to yield most of the land plant communities (Cheng et al. 2016; Timpano et al. 2020). In addition, the *trnL* P6 loop is the most commonly used marker in plant eDNA studies for a number of reasons: it has sufficient variability across both angiosperms and gymnosperms, there are a number of available reference databases as well as taxa-specific primers, and its small size works well for degraded eDNA (Alsos et al. 2020; Epp et al. 2018; Foucher et al. 2020).

Questions

1. The laboratory technician hands you an extraction protocol that has been used previously to extract DNA from soil and sediments. How do you know if this protocol will extract both iDNA and exDNA? Motivate your answer.
2. You are designing your soil eDNA study for a plant taxon that is distributed heterogeneously across plots. Describe the soil sampling strategy that will take into account the target taxon distribution.
3. You want to reconstruct vegetation types based on soil eDNA targeting the *trnL* P6 loop. This marker will not allow you to identify all taxa to species level. Will this affect your ability to determine the vegetation types? Motivate why or why not?

Glossary

Bioturbation – Biological processes involved in the dissemination of genetic media through terrestrial media.
DNA degradation – Refers to the physical changes of the DNA molecule.
DNA decay – Refers to the reduction in detectable quantity of eDNA.
DNA persistence – Refers to the amount of DNA that remains detectable across time.
DNA polymorphism – Presence of two or more variants of a particular DNA sequence.
Horizon – A layer parallel to the soil surface whose physical, chemical and biological characteristics differ from the layers above and beneath.

Power analysis – Probability of detecting an effect, given that the effect is really there. Can also be seen as rejecting the null hypothesis when it is in fact false.

Pedogenesis – The process of soil formation as regulated by the effects of place, environment, and history.

Rarefaction curves (in ecology) – A technique to assess species richness given the number of samples collected.

References

Alsos IG, Lammers Y, Yoccoz NG, Jørgensen T, Sjögren P, Gielly L, Edwards ME (2018) Plant DNA metabarcoding of lake sediments: how does it represent the contemporary vegetation. PLoS ONE 13, e0195403. https://doi.org/10.1371/journal.pone.0195403

Alsos IG, Lavergne S, Merkel MKF, Boleda M, Lammers Y, Alberti A, Pouchon C, Denoeud F, Pitelkova I, Pușcaș M, Roquet C, Hurdu B-I, Thuiller W, Zimmermann NE, Hollingsworth PM, Coissac E (2020) The treasure vault can be opened: large-scale genome skimming works well using herbarium and silica gel dried material. Plants 9, 432. https://doi.org/10.3390/plants9040432

Andersen K, Bird KL, Rasmussen M, Haile J, Breuning-Madsen H, Kjaer KH, Orlando L, Gilbert MTP, Willerslev E (2012) Meta-barcoding of "dirt" DNA from soil reflects vertebrate biodiversity. Mol. Ecol. 21, 1966–1979. https://doi.org/10.1111/j.1365-294X.2011.05261.x

Ariza M, Fouks B, Mauvisseau Q, Halvorsen R, Alsos IG, de Boer H (2022) Plant biodiversity assessment through soil eDNA reflects temporal and local diversity. Methods Ecol. Evol. https://doi.org/10.1111/2041-210X.13865

Baldwin DS, Mitchell AM (2000) The effects of drying and re-flooding on the sediment and soil nutrient dynamics of lowland river-floodplain systems: a synthesis. Regul. Rivers: Res. Mgmt. 16, 457–467. https://doi.org/10.1002/1099-1646(200009/10)16:5<457::AID-RRR597>3.0.CO;2-B

Barnes MA, Turner CR, Jerde CL, Renshaw MA, Chadderton WL, Lodge DM (2014) Environmental conditions influence eDNA persistence in aquatic systems. Environ. Sci. Technol. 48, 1819–1827. https://doi.org/10.1021/es404734p

Barnes MA, Turner CR (2015) The ecology of environmental DNA and implications for conservation genetics. Conserv. Genet. 17, 1–17. https://doi.org/10.1007/s10592-015-0775-4

Bienert F, De Danieli S, Miquel C, Coissac E, Poillot C, Brun J-J, Taberlet P (2012) Tracking earthworm communities from soil DNA. Mol. Ecol. 21, 2017–2030. https://doi.org/10.1111/j.1365-294X.2011.05407.x

Blum SAE, Lorenz MG, Wackernagel W (1997) Mechanism of retarded DNA degradation and prokaryotic origin of dnases in nonsterile soils. Syst. Appl. Microbiol. 20, 513–521. https://doi.org/10.1016/S0723-2020(97)80021-5

Boggs LM, Scheible MKR, Machado G, Meiklejohn KA (2019) Single fragment or bulk soil DNA metabarcoding: which is better for characterizing biological taxa found in surface soils for sample separation? Genes (Basel) 10, 431. https://doi.org/10.3390/genes10060431

Burdige DJ (2007) Geochemistry of marine sediments. Princeton University Press, Princeton. 624 pp.

Calderón-Sanou I, Münkemüller T, Boyer F, Zinger L, Thuiller W (2020) From environmental DNA sequences to ecological conclusions: how strong is the influence of methodological choices? J. Biogeogr. 47, 193–206. https://doi.org/10.1111/jbi.13681

Cheng T, Xu C, Lei L, Li C, Zhang Y, Zhou S (2016) Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Mol. Ecol. Resour. 16, 138–149. https://doi.org/10.1111/1755-0998.12438

Corti G, Agnelli A, Cuniglio R, Sanjurjo MF, Cocco S (2005) Characteristics of rhizosphere soil from natural and agricultural environments, in: Huang, P.M., Gobran, G.R. (Eds.), Biogeochemistry of Trace Elements in the Rhizosphere. Elsevier, pp. 57–128. https://doi.org/10.1016/B978-044451997-9/50005-2

Cozzolino S, Cafasso D, Pellegrino G, Musacchio A, Widmer A (2007) Genetic variation in time and space: the use of herbarium specimens to reconstruct patterns of genetic variation in the endangered orchid *Anacamptis palustris*. Conserv. Genet. 8, 629–639. https://doi.org/10.1007/s10592-006-9209-7

Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, Creer S, Bista I, Lodge DM, de Vere N, Pfrender ME, Bernatchez L (2017) Environmental DNA metabarcoding: transforming how we survey animal and plant communities. Mol. Ecol. 26, 5872–5895. https://doi.org/10.1111/mec.14350

Dickie IA, Boyer S, Buckley HL, Duncan RP, Gardner PP, Hogg ID, Holdaway RJ, Lear G, Makiola A, Morales SE, Powell JR, Weaver L (2018) Towards robust and repeatable sampling methods in eDNA-based studies. Mol. Ecol. Resour. 18, 940–952. https://doi.org/10.1111/1755-0998.12907

Dopheide A, Xie D, Buckley TR, Drummond AJ, Newcomb RD (2019) Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity. Methods Ecol. Evol. 10, 120–133. https://doi.org/10.1111/2041-210X.13086

Edwards ME, Alsos IG, Yoccoz N, Coissac E, Goslar T, Gielly L, Haile J, Langdon CT, Tribsch A, Binney HA, von Stedingk H, Taberlet P (2018) Metabarcoding of modern soil DNA gives a highly local vegetation signal in Svalbard tundra. The Holocene 28, 2006–2016. https://doi.org/10.1177/0959683618798095

Epp LS, Kruse S, Kath NJ, Stoof-Leichsenring KR, Tiedemann R, Pestryakova LA, Herzschuh U (2018) Temporal and spatial patterns of mitochondrial haplotype and species distributions in Siberian larches inferred from ancient environmental DNA and modeling. Sci. Rep. 8, 17436. https://doi.org/10.1038/s41598-018-35550-w

Fahner NA, Shokralla S, Baird DJ, Hajibabaei M (2016) Large-scale monitoring of plants through environmental DNA metabarcoding of soil: recovery, resolution, and annotation of four DNA markers. PLoS ONE 11, e0157505. https://doi.org/10.1371/journal.pone.0157505

Fatima F, Pathak N, Rastogi Verma S (2014) An improved method for soil DNA extraction to study the microbial assortment within rhizospheric region. Mol. Biol. Int. 2014, 518960. https://doi.org/10.1155/2014/518960

Foucher A, Evrard O, Ficetola GF, Gielly L, Poulain J, Giguet-Covex C, Laceby JP, Salvador-Blanes S, Cerdan O, Poulenard J (2020) Persistence of environmental DNA in cultivated soils: implication of this memory effect for reconstructing the dynamics of land use and cover changes. Sci. Rep. 10, 10502. https://doi.org/10.1038/s41598-020-67452-1

Frostegård A, Courtois S, Ramisse V, Clerc S, Bernillon D, Le Gall F, Jeannin P, Nesme X, Simonet P (1999) Quantification of bias related to the extraction of DNA directly from soils. Appl. Environ. Microbiol. 65, 5409–5420. https://doi.org/10.1128/AEM.65.12.5409-5420.1999

Gardner CM, Gunsch CK (2017) Adsorption capacity of multiple DNA sources to clay minerals and environmental soil matrices less than previously estimated. Chemosphere 175, 45–51. https://doi.org/10.1016/j.chemosphere.2017.02.030

Gothwal RK, Nigam VK, Mohan MK, Sasmal D, Ghosh P (2007) Extraction of bulk DNA from Thar Desert soils for optimization of PCR-DGGE based microbial community analysis. Electron. J. Biotechnol. 10, 400–408. https://doi.org/10.2225/vol10-issue3-fulltext-6

Gulden RH, Lerat S, Hart MM, Powell JR, Trevors JT, Pauls KP, Klironomos JN, Swanton CJ (2005) Quantitation of transgenic plant DNA in leachate water: real-time polymerase chain reaction analysis. J. Agric. Food Chem. 53, 5858–5865. https://doi.org/10.1021/jf0504667

Johnson MG, Pokorny L, Dodsworth S, Botigué LR, Cowan RS, Devault A, Eiserhardt WL, Epitawalage N, Forest F, Kim JT, Leebens-Mack JH, Leitch IJ, Maurin O, Soltis DE, Soltis PS, Wong GK-S, Baker WJ, Wickett NJ (2019) A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. Syst. Biol. 68, 594–606. https://doi.org/10.1093/sysbio/syy086

Kristensen E, Rabenhorst MC (2015) Do marine rooted plants grow in sediment or soil? A critical appraisal on definitions, methodology and communication. Earth-Science Reviews 145, 1–8. https://doi.org/10.1016/j.earscirev.2015.02.005

Lacoursière-Roussel A, Deiner K (2021) Environmental DNA is not the tool by itself. J. Fish Biol. 98, 383–386. https://doi.org/10.1111/jfb.14177

Levy-Booth DJ, Campbell RG, Gulden RH, Hart MM, Powell JR, Klironomos JN, Pauls KP, Swanton CJ, Trevors JT, Dunfield KE (2007) Cycling of extracellular DNA in the soil environment. Soil Biology and Biochemistry 39, 2977–2991. https://doi.org/10.1016/j.soilbio.2007.06.020

Meiklejohn KA, Jackson ML, Stern LA, Robertson JM (2018) A protocol for obtaining DNA barcodes from plant and insect fragments isolated from forensic-type soils. Int. J. Legal Med. 132, 1515–1526. https://doi.org/10.1007/s00414-018-1772-1

Murchie TJ, Kuch M, Duggan AT, Ledger ML, Roche K, Klunk J, Karpinski E, Hackenberger D, Sadoway T, MacPhee R, Froese D, Poinar H (2021) Optimizing extraction and targeted capture of ancient environmental DNA for reconstructing past environments using the PalaeoChip Arctic-1.0 bait-set. Quaternary Research 99, 305–328. https://doi.org/10.1017/qua.2020.59

Nagler M, Insam H, Pietramellara G, Ascher-Jenull J (2018) Extracellular DNA in natural environments: features, relevance and applications. Appl. Microbiol. Biotechnol. 102, 6343–6356. https://doi.org/10.1007/s00253-018-9120-4

Nielsen KM, Smalla K, van Elsas JD (2000) Natural transformation of Acinetobacter sp. strain BD413 with cell lysates of Acinetobacter sp., Pseudomonas fluorescens, and Burkholderia cepacia in soil microcosms. Appl. Environ. Microbiol. 66, 206–212. https://doi.org/10.1128/aem.66.1.206-212.2000

Nocker A, Fernández PS, Montijn R, Schuren F (2012) Effect of air drying on bacterial viability: A multiparameter viability assessment. J. Microbiol. Methods 90, 86–95. https://doi.org/10.1016/j.mimet.2012.04.015

Parducci L, Nota K, Wood J (2018) Reconstructing past vegetation communities using ancient DNA from lake sediments, in: Lindqvist, C., Rajora, O.P. (Eds.), Paleogenomics: Genome-Scale Analysis of Ancient DNA. Springer International Publishing, Cham, pp. 163–187. https://doi.org/10.1007/13836_2018_38

Pedersen MW, Overballe-Petersen S, Ermini L, Sarkissian CD, Haile J, Hellstrom M, Spens J, Thomsen PF, Bohmann K, Cappellini E, Schnell IB, Wales NA, Carøe C, Campos PF, Schmidt AMZ, Gilbert MTP, Hansen AJ, Orlando L, Willerslev E (2015) Ancient and modern environmental DNA. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370, 20130383. https://doi.org/10.1098/rstb.2013.0383

Pietramellara G, Ascher J, Borgogni F, Ceccherini MT, Guerri G, Nannipieri P (2009) Extracellular DNA in soil and sediment: fate and ecological relevance. Biol. Fertil. Soils 45, 219–235. https://doi.org/10.1007/s00374-008-0345-8

Poté J, Ackermann R, Wildi W (2009) Plant leaf mass loss and DNA release in freshwater sediments. Ecotoxicol. Environ. Saf. 72, 1378–1383. https://doi.org/10.1016/j.ecoenv.2009.04.010

Prosser CM, Hedgpeth BM (2018) Effects of bioturbation on environmental DNA migration through soil media. PLoS ONE 13, e0196430. https://doi.org/10.1371/journal.pone.0196430

Ritter CD, Zizka A, Roger F, Tuomisto H, Barnes C, Nilsson RH, Antonelli A (2018) High-throughput metabarcoding reveals the effect of physicochemical soil properties on soil and litter biodiversity and community turnover across Amazonia. PeerJ 6, e5661. https://doi.org/10.7717/peerj.5661

Saeki K, Ihyo Y, Sakai M, Kunito T (2011) Strong adsorption of DNA molecules on humic acids. Environ. Chem. Lett. 9, 505–509. https://doi.org/10.1007/s10311-011-0310-x

Schulz S, Brankatschk R, Dümig A, Kögel-Knabner I, Schloter M, Zeyer J (2013) The role of microorganisms at different stages of ecosystem development for soil formation. Biogeosciences 10, 3983–3996. https://doi.org/10.5194/bg-10-3983-2013

Shackley M (1975) Archaeological sediments: a survey of analytical methods. Butterworth, London and Boston.

Shlemon RJ (1985) Application of soil-stratigraphic techniques to engineering geology. Environmental & Engineering Geoscience xxii, 129–142. https://doi.org/10.2113/gseegeosci.xxii.2.129

Sirois SH, Buckley DH (2019) Factors governing extracellular DNA degradation dynamics in soil. Environ. Microbiol. Rep. 11, 173–184. https://doi.org/10.1111/1758-2229.12725

Smol JP, Birks HJB, Last WM, Bradley RS, Alverson K (Eds) (2001) Tracking environmental change using lake sediments: terrestrial, algal, and siliceous indicators, Developments in paleoenvironmental research. Springer Netherlands, Dordrecht. https://doi.org/10.1007/0-306-47668-1

Taberlet P, Bonin A, Zinger L, Coissac E (Eds) (2018) Environmental DNA: for biodiversity research and monitoring. Oxford University Press. https://doi.org/10.1093/oso/9780198767220.001.0001

Taberlet P, Prud'Homme SM, Campione E, Roy J, Miquel C, Shehzad W, Gielly L, Rioux D, Choler P, Clément J-C, Melodelima C, Pompanon F, Coissac E (2012) Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. Mol. Ecol. 21, 1816–1820. https://doi.org/10.1111/j.1365-294X.2011.05317.x

Thomsen PF, Willerslev E (2015) Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

Timpano EK, Scheible MKR, Meiklejohn KA (2020) Optimization of the second internal transcribed spacer (ITS2) for characterizing land plants from soil. PLoS ONE 15, e0231436. https://doi.org/10.1371/journal.pone.0231436

Torsvik V, Goksøyr J, Daae FL (1990) High diversity in DNA of soil bacteria. Appl. Environ. Microbiol. 56, 782–787. https://doi.org/10.1128/aem.56.3.782-787.1990

Vogt KA, Vogt DJ, Palmiotto PA, Boon P, O'Hara J, Asbjornsen H (1995) Review of root dynamics in forest ecosystems grouped by climate, climatic forest type and species. Plant Soil 187, 159–219. https://doi.org/10.1007/BF00017088

Vuillemin A, Horn F, Alawi M, Henny C, Wagner D, Crowe SA, Kallmeyer J (2017) Preservation and significance of extracellular DNA in ferruginous sediments from Lake Towuti, Indonesia. Front. Microbiol. 8, 1440. https://doi.org/10.3389/fmicb.2017.01440

Wardle DA, Bardgett RD, Klironomos JN, Setälä H, van der Putten WH, Wall DH (2004) Ecological linkages between aboveground and belowground biota. Science 304, 1629–1633. https://doi.org/10.1126/science.1094875

Willerslev E, Cooper A (2005) Ancient DNA. Proc. Biol. Sci. 272, 3–16. https://doi.org/10.1098/rspb.2004.2813

Willerslev E, Davison J, Moora M, Zobel M, Coissac E, Edwards ME, Lorenzen ED, Vestergård M, Gussarova G, Haile J, Craine J, Gielly L, Boessenkool S, Epp LS, Pearman PB, Cheddadi R, Murray D, Bråthen KA, Yoccoz N, Binney H, Taberlet P (2014) Fifty thousand years of Arctic vegetation and megafaunal diet. Nature 506, 47–51. https://doi.org/10.1038/nature12921

Wood JM (1987) Biological processes involved in the cycling of elements between soil or sediments and the aqueous environment. Hydrobiologia 149, 31–42. https://doi.org/10.1007/BF00048644

Yoccoz NG, Bråthen KA, Gielly L, Haile J, Edwards ME, Goslar T, Von Stedingk H, Brysting AK, Coissac E, Pompanon F, Sønstebø JH, Miquel C, Valentini A, De Bello F, Chave J, Thuiller W, Wincker P, Cruaud C, Gavory F, Rasmussen M, Taberlet P (2012) DNA from soil mirrors plant taxonomic and growth form diversity. Mol. Ecol. 21, 3647–3655. https://doi.org/10.1111/j.1365-294X.2012.05545.x

Yoccoz NG (2012) The future of environmental DNA in ecology. Mol. Ecol. 21, 2031–2038. https://doi.org/10.1111/j.1365-294X.2012.05505.x

Zhou L-J, Pei K-Q, Zhou B, Ma K-P (2007) A molecular approach to species identification of Chenopodiaceae pollen grains in surface soil. Am. J. Bot. 94, 477–481. https://doi.org/10.3732/ajb.94.3.477

Zinger L, Bonin A, Alsos IG, Bálint M, Bik H, Boyer F, Chariton AA, Creer S, Coissac E, Deagle BE, De Barba M, Dickie IA, Dumbrell AJ, Ficetola GF, Fierer N, Fumagalli L, Gilbert MTP, Jarman S, Jumpponen A, Kauserud H, Taberlet P (2019a) DNA metabarcoding-Need for robust experimental designs to draw sound ecological conclusions. Mol. Ecol. 28, 1857–1862. https://doi.org/10.1111/mec.15060

Zinger L, Shahnavaz B, Baptist F, Geremia RA, Choler P (2009) Microbial diversity in alpine tundra soils correlates with snow cover dynamics. ISME J. 3, 850–859. https://doi.org/10.1038/ismej.2009.20

Zinger L, Taberlet P, Schimann H, Bonin A, Boyer F, De Barba M, Gaucher P, Gielly L, Giguet-Covex C, Iribar A, Réjou-Méchain M, Rayé G, Rioux D, Schilling V, Tymen B, Viers J, Zouiten C, Thuiller W, Coissac E, Chave J (2019b) Body size determines soil community assembly in a tropical forest. Mol. Ecol. 28, 528–543. https://doi.org/10.1111/mec.14919

Answers

1. By checking if there is a step that can lyse the cells to extract iDNA. This step can be grinding, sonication, thermal shocks, or chemical treatments such as with chloroform.
2. To take into account heterogeneity the strategy is to take many subsamples and mix them.
3. Soil eDNA using *trnL* P6 loop will not give you accurate species lists in most floras, but rather lists of genera with occasional low-level or higher-level identifications. Most vegetation types are characterized by a few key species only, so having limited taxonomic resolution of your identifications is unlikely to affect the overall vegetation type calling. However in some floras or vegetation types this approach will be insufficient, e.g., for those characterized by specific taxa in locally speciose genera.

Paper III

Received: 19 November 2021 | Accepted: 10 February 2022

DOI: 10.1111/2041-210X.13865

Methods in Ecology and Evolution

RESEARCH ARTICLE

Leveraging Natural History Collections to Understand the Impacts of Global Change

Plant biodiversity assessment through soil eDNA reflects temporal and local diversity

María Ariza[1] | Bertrand Fouks[2] | Quentin Mauvisseau[1] | Rune Halvorsen[1] | Inger Greve Alsos[3] | Hugo J. de Boer[1]

[1]Universitetet i Oslo, Naturhistorisk Museum, Oslo, Norway

[2]Westfälische Wilhelms-Universität, Institute for Evolution and Biodiversity, Molecular Evolution and Bioinformatics, Münster, Germany

[3]The Arctic University Museum of Norway, UiT - The Arctic University of Norway, Tromsø, Norway

Correspondence
María Ariza
Email: mariadelosangelesariza@gmail.com

Funding information
H2020 MSCA-ITN-ETN Plant.ID network, Grant/Award Number: No 765000

Handling Editor: Lynsey Rebecca Harper

Abstract

1. Several studies have shown the potential of eDNA-based proxies for plant identification, but little is known about their spatial and temporal resolution. This limits its use for plant biodiversity assessments and monitoring of vegetation responses to environmental changes. Here we calibrate the temporal and spatial plant signals detected with soil eDNA surveys by comparing with a standard visual above-ground vegetation survey.
2. Our approach compares vegetation in an old-growth boreal forest in southern Norway, surveyed in 100 permanent 1-m^2 plots seven times over a 30-year period, with a single soil eDNA metabarcoding-based survey from soil samples collected at the same 100 plots in the year of the last vegetation survey.
3. On average, 60% and 10% of the vascular plants and bryophytes recorded across all vegetation surveys were detected by soil eDNA. Taxa detected by soil eDNA were more representative for the local taxa pool than for the specific plot, and corresponded to those surveyed over the 30-year period although most closely matched the current taxa composition. Soil eDNA detected abundant taxa better than rare ones although both rare taxa and taxa unrecorded by the visual survey were detected.
4. Our study highlights the potential of soil eDNA assessments for monitoring of vegetation responses over broad spatial and temporal scales. The method's ability to detect abundant taxa makes it suitable for assessment of vegetation composition in a specific area and for broad-scale plant diversity assessments.

KEYWORDS
metabarcoding, plant identification, soil eDNA, spatial scale, temporal change, vegetation assessments

1 | INTRODUCTION

The current global warming crisis and the fast pace of global biodiversity losses relative to its appraisal require innovative and rapid operational approaches to biodiversity assessment like never before. Plants are central to most biodiversity assessments, as they are predominant and ubiquitous (Kier et al., 2005), as well as valuable indicators of associated diversity (Brunbjerg et al., 2018), surrounding

abiotic features (Terwayet Bayouli et al., 2021) and human impact (Uuemaa et al., 2013). In addition, plants are suitable organisms for climate changes monitoring as community reshuffling and spatial redistribution towards summits and higher latitudes are well-established biotic responses to increased temperatures (Bertrand et al., 2011; Chen et al., 2011; Steinbauer et al., 2018; Wiens, 2016). Moreover, plant assessments are often required to map habitats, monitor environmental quality and assess habitat changes in space and time (Halvorsen et al., 2020).

Recording the taxonomic composition is the cornerstone of any biodiversity assessment, and this task requires the use of morphological and/or molecular proxies for detection and identification of taxa (Ruppert et al., 2019). Morphological proxies require inspection of plant characters that are diagnostic for the specimens' identity. However, this is a rather lengthy, and thus expensive, process that usually requires participation from trained botanists. In addition, a plant must be noticeably present and/or have the characteristics necessary to enable observation and identification, thus limiting the seasons in which a majority of different taxa present may be monitored. Historically, censuses of plant diversity, forest inventories and monitoring programmes have relied on morphological proxies (Corona et al., 2011). These have contributed greatly to the discovery and current knowledge of the known diversity.

More recently, molecular proxies for taxon identification based on environmental DNA (eDNA) have been taken into use for biodiversity assessment purposes (Beng & Corlett, 2020). The use of eDNA extracted from soil, water, faeces or bulk samples (Taberlet et al., 2012) grants the possibility of collecting organismal or extra-organismal DNA from multiple individuals and taxa simultaneously, saving lengthy and costly hours in the field collecting specimens. In addition, this non-invasive and non-destructive method may be useful for the detection of rare, elusive and/or challenging-to-collect taxa (Alsos et al., 2018; Carrasco-Puga et al., 2021; Hartvig et al., 2021). eDNA-based surveys thus may open for rapid assessment and monitoring of biodiversity within a particular region, which is a critical aspect to understand effects of the current climate change crisis and biodiversity losses. Indeed, soil eDNA samples have gained attention as a potentially valuable tool for the assessment of plant diversity, as it may harbour DNA from both above-ground and below-ground signals (i.e. pollen, debris, roots), from active as well as dormant plant tissues (Hiiesalu et al., 2012). Accordingly, eDNA may provide a series of past and present plant signals that can assist the documentation of local extinctions and long-term ecosystem changes. Soil eDNA has most often been used to assess plant diversity in Arctic and boreal regions where the low temperature facilitates DNA preservation (Edwards et al., 2018; Wang et al., 2021; Willerslev et al., 2014; Yoccoz et al., 2012), although it may also be successful in tropical (Yoccoz et al., 2012; Zinger et al., 2019), and extreme environments such as deserts (Carrasco-Puga et al., 2021) and geothermal sites (Fraser et al., 2018). Furthermore, soil eDNA assessments have been successfully applied to identify present and past diversity from natural or cultivated areas (Yoccoz et al., 2012), assess woody encroachment in grasslands (Sepp et al., 2021) and predict habitats from crime scenes (Fløjgaard et al., 2019).

While eDNA-based detection for assessing diversity is already in widespread use (Deiner et al., 2021), knowledge about the sources, fate, persistence and transport of eDNA in the environment is scarce and mainly explored in aquatic environments (Mauvisseau et al., 2021). Knowledge of these properties of eDNA is indispensable to establish the temporal and spatial resolution expected of an eDNA assessment and to assess the utility of soil eDNA-based methods for ecological monitoring, for example, of biotic responses to climate changes (Deiner et al., 2021). Soil eDNA has been shown to reflect plant diversity at local (Beng & Corlett, 2020; Edwards et al., 2018; Kumpula, 2020; Yoccoz et al., 2012) as well as regional scales (Carrasco-Puga et al., 2021). Furthermore, crop signals from 10 to 50 years into the past have been detected in cultivated soils (Foucher et al., 2020; Yoccoz, 2012). These insights, obtained from diverse environments with different anthropogenic pressures, may indicate that soil eDNA assessments are adequate for monitoring of vegetation. However, exploration of soil eDNA-based methods is still in its infancy and important knowledge gaps still exist as exemplified by the combined effects of spatial and temporal resolution on plant eDNA signals in soils, which to our knowledge has not yet been studied. This and other knowledge gaps have to be filled to establish a calibration field protocol for monitoring vegetation areas in near-natural state, for example, for assessment of vegetation responses to climate changes.

Here, we use time-series data from intermittent vegetation surveys originally designed to track vegetation responses to climate changes in an old-growth forest in South Norway, to investigate the temporal and spatial resolution of a soil eDNA assessment made in 2018. Our research questions are as follows: (a) Do soil eDNA assessments capture the current diversity or integrate diversity over a longer time period?; (b) Do soil eDNA assessments reflect the plot-specific diversity or the local or regional pool of taxa?; (c) Can soil eDNA assessments provide taxa abundance estimates, that is, are common and abundant taxa detected with higher probability than rare, subordinate taxa? We used metabarcoding analysis of the trnL (UUA) intron P6 loop and identified plant sequences with customed reference libraries previously built by sequencing herbarium collections (Alsos et al., 2020; Soininen et al., 2015; Sønstebø et al., 2010; Willerslev et al., 2014). While the marker region is chosen based on vascular plants, bryophytes are common by-catch, and we also report results for bryophytes.

2 | MATERIALS AND METHODS

2.1 | Vegetation plot surveys

In 1988, one hundred 1-m^2 plots were placed across eight transects with a total length of 1,320 m, subjectively selected to cover the broad-scale variation in forest vegetation in response to natural edaphic gradients in the Solhomfjell Forest Reserve, southern

Norway, 58°58′N, 8°58′E, at 350–480 m a.s.l. (Figure 1; Økland & Eilertsen, 1993). Fieldwork permission was granted by the Environmental protection authorities at the County Governor's office in Aust-Agder before fieldwork started in 1988. The studied area is situated within the southern boreal zone and harbours protected old-growth forests with overstories dominated by either Norway spruce (*Picea abies* [L.] H.Karst.) or Scots pine (*Pinus sylvestris* L.), alternating with mires and rock outcrops. The one hundred 1-m^2 plots were distributed semi-systematically along the eight transects. Every 10th m along each transect was a candidate plot position. From the 132 candidate positions, 100 were selected randomly. Of these 100 plots, 61 were dominated by Norway spruce (hereafter referred to as 'spruce subset') and 39 plots were dominated by Scots pine ('pine subset'), respectively. All 1-m^2 plots (hereafter called 'plots') were censused for vegetation composition every fifth year from 1988 to 2018 (t_1 = 1988, t_2 = 1993, t_3 = 1998, t_4 = 2003, t_5 = 2008, t_6 = 2013, t_7 = 2018). At census, all vascular plants (including lignified taxa <80 cm high) and bryophytes were carefully searched for and their presence/absence recorded in each of 16 equal-sized subplots of 625 cm^2 in each plot. Subplot frequency (0–16) was used as a taxon abundance measure. A total of 157 taxa were registered, including 69 vascular plants and 88 bryophytes (Table S1; Figure 1). This time series of vegetation data, hereafter referred to as 'vegetation survey', has been documented and subjected to analyses in a series of publications from the Natural History Museum, University of Oslo (Halvorsen et al., 2019; Økland et al., 2004; Økland & Eilertsen, 1994; Økland & Eilertsen, 1996).

The turnover of the vegetation plot composition was calculated for vascular plants and bryophytes as the sum of proportions of taxa gained and lost from survey time t−1 to survey time t as fraction of the all taxa recorded at the two time points, using the R library 'codyn' (Hallett et al., 2016). Turnover values with 95% confidence intervals for each period were obtained separately for spruce and pine subsets, by averaging across plots.

For all taxa recorded in the vegetation survey, we created a local Solhomfjell sequence reference library for the chloroplast trnL intron (UUA) P6 loop with retrieved sequences from the regional arctic and boreal reference library (Arcborbryo; Soininen et al., 2015; Sønstebø et al., 2010; Willerslev et al., 2014) when available there, otherwise GenBank and for a few taxa unpublished P6 loop sequences were retrieved from PhyloNorway (Alsos et al., 2020, Table S1). Different taxa with identical sequences for this marker were merged at the

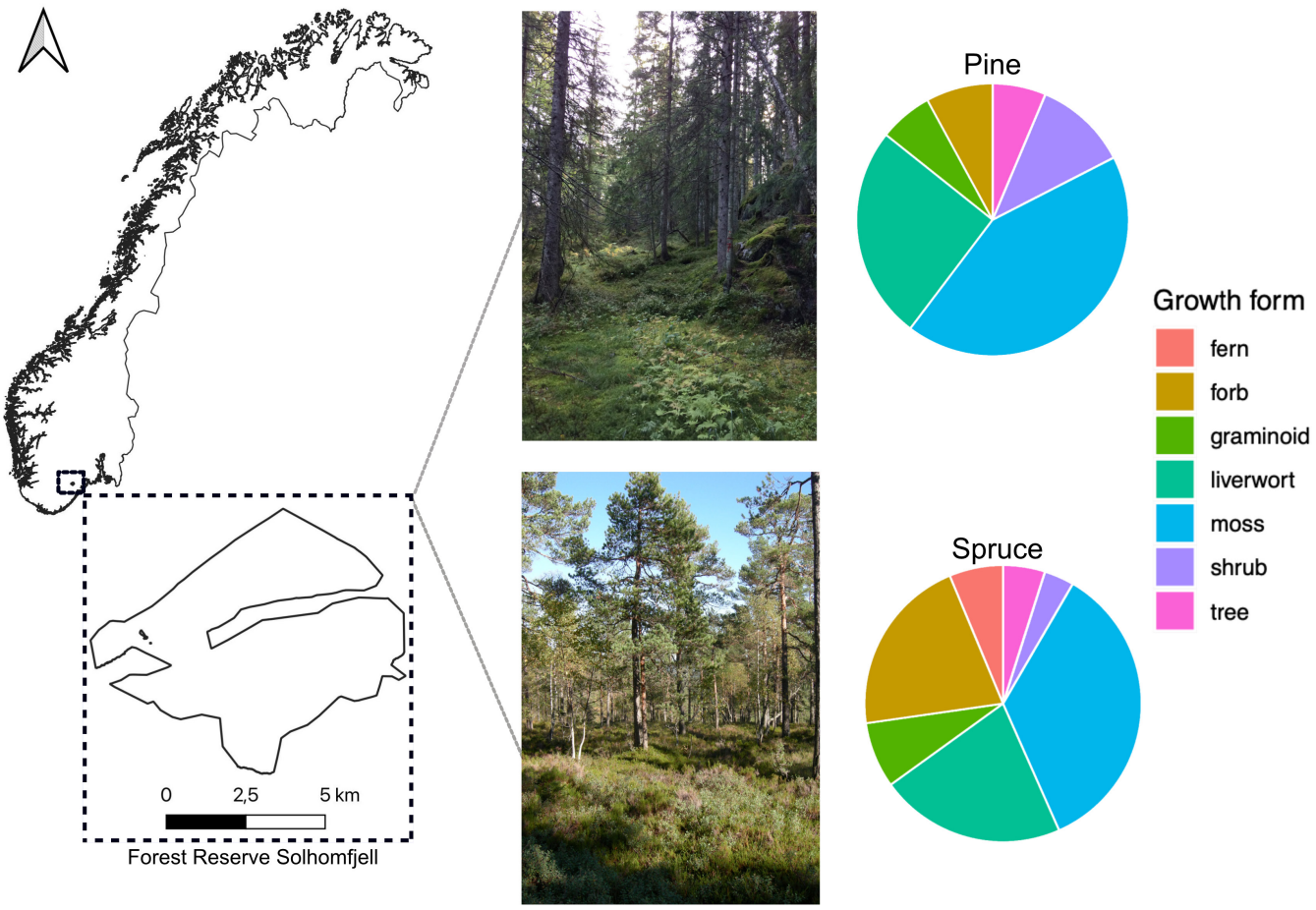

FIGURE 1 Location of the study area, Solhomfjell Forest Reserve, within Norway (left), typical interior of forests dominated by Norway spruce (above) and scots pine (below), and the relative number of taxa of different growth forms registered in the vegetation surveys. A map of the 100 survey plots with specific locations is given in Økland and Eilertsen (1993)

lowest possible taxonomic level and named accordingly. Presence/absence values (1 or 0) for the merged taxon were obtained from the original data by summation of presence values for all parent taxa. A sum ≥1 was scored as presence (1), otherwise absence (0). If each parent taxa had subplot frequency ≥ 16, biomass of the merged taxon was obtained by averaging subplot frequency values for the parent taxa, otherwise the sum of subplot frequency values for parent taxa was used (Table S1).

2.2 | Soil eDNA sampling, amplification and sequencing

A single soil eDNA sample was collected from the centre of each vegetation plot surveyed in the Solhomfjell Forest Reserve in August 2018. Debris and living plant parts were removed to expose the topsoil for sampling, and 50 ml Falcon tubes (11 cm) were pushed into the organic soil. The soil-filled falcon tubes were pulled up and immediately capped after retrieval. The soil eDNA samples were stored in individual plastic bags for transportation to the laboratory and stored at −20 °C prior to freeze-drying under vacuum. Each soil eDNA sample was separately homogenised with ceramic beads and 1 g was used for eDNA extraction. The latter was done in five rounds of two steps: (a) CTAB/chloroform pre-treatment to increase the separation of the organic phase and (b) aqueous phase and using the E.Z.N.A. soil DNA kit following the manufacturer's protocol (Omega Bio-tek; see Data S1 for a detailed protocol). The chloroplast marker trnL (UAA) intron P6 loop was chosen as its short sequence can yield amplification of old DNA material degraded in eDNA samples. This marker was amplified for each sample with the *g* and *h* primers by PCR, using three technical replicates (Taberlet et al., 2007; 5′-GGGCAATCCTGAGCCAA-3′, 5′-CCATTGAGTCTCTGCACCTATC-3′). Forward and reverse primers were tagged with a unique 12 bp oligonucleotide on the 5′ end (Fadrosh et al., 2014). Unique combinations of tagged primers were set up in panels for each PCR reaction for a total of 309 samples (100 samples with 3 PCR replicates each, 5 extractions blanks and 4 PCR negatives). The PCR negatives had no DNA template and were placed on the 96th well position in each panel. Composition of PCRs, final volumes and number of cycles can be found in Data S1. The PCR products were run on a 2% agarose gel, and the amplicon concentrations were measured via band intensity using ImageLab software (Bio-Rad). The lowest concentration (μM) available for all PCR products and its relative volume was identified and the relative concentrations of the PCR products were adjusted to this same concentration. Amplicons were pooled in one library using a Biomek 4000 automated liquid handler (Beckman Coulter Life Sciences). The library was cleaned using AMPure XP reagent beads (Beckman Coulter Life Sciences). The length for all amplicons in the library was determined using a Fragment Analyser (Agilent Technologies). The library was sequenced on an Illumina MiSeq platform with 150 bp paired-end reads (Illumina Inc.).

2.3 | Sequence analysis and taxonomic identification

Sequence data were analysed and curated using OBITools 2 (Boyer et al., 2016; sequences and a detailed script is available at Ariza et al., 2022) following the wolf tutorial with adaptations for demultiplexing dual indexes from QIIME2 (Caporaso et al., 2010). Sequences were retained with both indexes for dereplication for further analysis. Similar sequences were clustered with *obiclean* (Boyer et al., 2016) only when the read count of the less abundant sequence was below 5% of the most abundant sequence. To reduce multiple identifications of the same sequence, taxonomic assignment of dereplicated and denoised sequences was done by matching to three reference sequences databases containing: (a) only taxa registered in the local Solholmfjell reference library (see above); (b) the complete arctic boreal database for vascular plants and bryophytes (Soininen et al., 2015; Sønstebø et al., 2010; Willerslev et al., 2014); and (c) taxa available in the EMBL database (downloaded on 7/02/2020) filtered to sequences with trnL (UUA) intron *g-h* primers using *ecoPCR* tool from OBITools (Boyer et al., 2016). Resulting identifications from the three databases were merged by sequence and duplicates were eliminated giving priority to reference databases (a), (b) and (c) in that order. To minimise erroneous taxonomic assignments, only taxa with a 100% match to a reference sequence were retained. We observed that below this threshold, sequences remained without a taxonomic rank assigned. Furthermore, assigned taxa names were changed to the lowest taxonomic rank possible with trnL (UUA) intron and thus are identical to those registered in vegetation surveys. When different sequences were identified with identical taxa names, a unique entry was retained and the read counts within plots and replicates were summed. Read counts were averaged across all samples and negative controls (extraction + PCR).

2.4 | Comparison between vegetation surveys and eDNA survey

The vegetation survey composition served as a baseline to assess the overlap with the composition of the soil eDNA survey, and from this, the spatial and temporal resolution was derived. Thus, our comparison framework comprises (a) a 7-point time-series of vegetation surveys in one hundred 1-m^2 plots (spruce and pine data subsets) from the Solhomfjell Forest Reserve, carried out in the years 1988, 1993, 1998, 2003, 2008, 2013 and 2018; and (b) a single soil eDNA survey made in 2018 by sampling soil eDNA at the centre of these plots. All the following analyses are plot based, and coded using R v 1.4.17 (R Core Team, 2019) and with packages listed in the code (available at Ariza et al., 2022). Separate analyses are made for vascular plants and bryophytes, and/or for spruce and pine data subsets, or combinations thereof, when relevant.

For comparison between vegetation and soil eDNA survey(s), we quantified the taxonomic overlap for each plot by the *number*

of matching taxa variable, a count of identical taxon names registered in both surveys. Furthermore, for each plot, we calculated the total number of taxa recorded across the seven vegetation surveys by *number of taxa in total vegetation surveys* variable, and the fraction with available reference sequences for the trnL (UUA) p6 loop by *number of detectable taxa in total vegetation surveys* variable. Thus, this variable corresponds to the maximum number of taxa recorded in the vegetation survey that could possibly be identified by soil eDNA. Spearman's nonparametric correlation coefficients (e.g. Sokal & Rohlf, 1995) were calculated between the *number of matching taxa* and *number of taxa in total and detectable vegetation surveys*.

Proportions of *number of matching taxa* out of *number of taxa in total vegetation surveys* and *number of detectable taxa in total vegetation survey* are referred to as *number of matching taxa in total vegetation survey* and *number of matching taxa in total and detectable vegetation survey*, respectively. The proportion of *number of matching taxa in total and detectable vegetation survey* was compared between spruce and pine subsets by use of a two-sample unpaired Wilcoxon–Mann–Whitney rank-sum test (e.g. Sokal & Rohlf, 1995).

2.5 | Temporal resolution of eDNA survey data

The temporal resolution of soil eDNA assessments, that is, at which temporal scale (survey year) the two inventories were most comparable, was evaluated by recalculating per plot the *number of matching taxa* and *number of detectable taxa in vegetation survey t* for taxa compositions from each year of vegetation survey ($t_1 = 1988 \ldots t_7 = 2018$). Then, the proportion of *number of matching taxa* out of the *number of detectable taxa in vegetation survey t* was calculated for each survey, and the seven variables thus obtained per plot are referred to as the *number of matching taxa in detectable vegetation survey t* ($t_1 = 1988 \ldots t_7 = 2018$), etc. An overall assessment was based upon calculation of mean values across all plots with 95% confidence intervals. The 'best detected vegetation survey' (t_{bdvs}) was determined as the vegetation survey *t* with the highest mean plot *number of matching taxa in detectable vegetation survey*. To determine if the mean plot *number of matching taxa in detectable vegetation survey t* was similar across years, multiple pairwise comparisons were tested across all years (groups) using a Friedman Test. A *post-hoc* Tukey test was used to identify significantly different groups.

To determine whether the similarity of the plot *number of matching taxa in detectable vegetation survey t* across years was due to soil eDNA survey recording the same dominant taxa over vegetation survey years, we investigated the relation of this variable to the unchanged composition between survey t_{bdvs} and $t_1, \ldots, t_6$ with a Spearman correlation test. First, we calculated the number of identical taxa between vegetation survey t_{bdvs} and $t_1, \ldots, t_6$ by *unchanged composition* variable. Second, for each plot, we summed the number of taxa between vegetation survey t_{bdvs} and $t_1, \ldots, t_6$ by *taxon count* variable. Thus, six iterations were calculated for both of these variables in each plot. Finally, we calculated the proportion of *unchanged composition* out of the *taxon count* $t_{bdvs} + t$ for each iteration.

To investigate plant DNA permanence and the past plant signals stored in a soil eDNA sample, we annotated each taxon from the plot *number* of *matching taxa in total and detectable vegetation survey* variable with the vegetation survey year *t* in which it was recorded. Since taxa permanence across vegetation survey years will lead to many annotations, we focused on those taxa that were recorded only in 1 year of vegetation survey. We call this subset a *temporal number of matching taxa* and describe how many years ago ($t_1 = 30$, $t_2 = 25 \ldots t_7 = 0$) each taxon was recorded.

2.6 | Spatial resolution of eDNA survey data

The spatial resolution of taxa registered in the soil eDNA survey in a given plot was assessed by categorising each taxon name according to the spatial scale on which was recorded: *matching taxa* in the best detected vegetation survey if recorded within the plot (<1 m^2), *vegetation survey match* if also registered in the best detected vegetation survey but in another plots (>1 m^2), or *regional flora match* if registered in artsdatabanken.no for the 41.77 km^2 Solhomfjell Forest Reserve area (Figure 1; >1 m^2). In addition, taxa with match to non-native plants or with higher mean number of reads in the negative controls than across samples were categorised as *false positives*. Sequence reads assigned to taxa within all categories were counted, and taxon count and proportions were calculated. Moreover, the spatial scales of both soil eDNA and vegetation surveys were compared plot-wise by correlating the number of taxa from above categories to the number of taxa registered in the best detected vegetation survey. Correlations were fitted to a linear model and predicted with a 95% confidence level interval.

2.7 | Representation of vegetation in soil DNA survey

Each taxon name in the best detected vegetation survey was categorised as a *matching taxa* if also present in the soil eDNA survey, *no trnL reference* if the reference sequence for the *trnL* (UUA) intron with *g-h* primers was not available or *undetected by eDNA* if not present in the soil eDNA survey but a reference sequence was available. Taxon counts and proportions within all categories were calculated.

Finally, we investigated if soil eDNA detections of taxa occurring within 1 m^2 plots (*matching taxa*) reflected their abundance. For this, we summarised the subplot frequency (used as proxy for biomass; Porté et al., 2009; Wilson, 2011) separately for *matching taxa* and taxa *undetected by eDNA*. The distribution of biomass data for each taxon in each plot from these two categories was compared with a Spearman's rank correlation test. We also investigated the correlation between the number of reads assigned to each matching taxa within major growth forms and the biomass registered in the best detected survey year.

3 | RESULTS

3.1 | Vegetation surveys

Counts of taxa registered in plots, both in each survey and across all seven vegetation survey years can be found in Data S3. Across all vegetation surveys, the spruce subset registered more unique vascular plants and bryophytes than the pine subset. The number of vascular plant taxa recorded per plot in each vegetation survey was on average 9.23 ± 5.12 *SD* and 5.23 ± 1.40 *SD* in the spruce and pine subsets, respectively, while the corresponding numbers for bryophytes were 9.29 ± 3.33 *SD* and 5.48 ± 2.04 *SD*.

Furthermore, a compositional turnover of 15%–20% and 20%–25% of the vascular and bryophyte composition, respectively, was found between consecutive vegetation survey years (Figure 2). Similarly, higher turnover was found for pine than for spruce subsets for most 5-year periods (Figure 2). Across consecutive vegetation survey years, 7%–15% and 4%–14% of the total bryophytes and vascular plants were gained, and 10%–20% and 10%–15% lost, respectively (Figure S1). Accordingly, the number of taxa of both groups declined over the 30-year survey period. Differences between bryophytes and vascular plants were more pronounced for gains than those for losses (Figure S1).

Reference sequences for the trnL (UUA) intron were available for 84% (133/157) plant taxa from the total vegetation survey. Of the remaining 13% (24/157), 20 were bryophytes and four were vascular plants (Table S1), and these taxa were filtered out from most analyses. Several taxa, 21% (28/133), had identical sequences (mainly bryophytes), and these were lumped into 12 unique taxa. Our final vegetation dataset thus used for comparison with the soil eDNA survey consisted of 65 vascular plants and 68 bryophytes (117 in total; Table S1).

3.2 | Soil eDNA survey

The Illumina MiSeq PE150 run yielded approximately 10 million reads assigned either to plots or blanks (mean reads/plot for all replicates: >100,000; mean reads/blank: 384). Of these, more than 4 million reads matched 100% to 130 taxa in the reference libraries. A summary of the soil eDNA survey is shown in Table S2. The large majority of these reads (4,792,356) were assigned to 116 vascular plants, while only 5,295 reads were assigned to 11 bryophytes. Taxon identifications for vascular plants included 75 species, 31 genera, 8 families and 2 subtribes, while for bryophytes included 6 species and 5 species complexes. Thus, 65% and 55% of the taxa were identified at species level for vascular plants and bryophytes, respectively. The ericaceous genus *Vaccinium*, represented by *V. myrtillus*, *V. uliginosum* and *V. vitis-idaea*, conifers (*Pinus sylvestris* and *Picea abies*), and deciduous trees (*Betula* spp, *Populus tremula*) were the most common vascular plants found in the eDNA records. The most common bryophytes found in eDNA records were the feather mosses *Hylocomium splendens* and

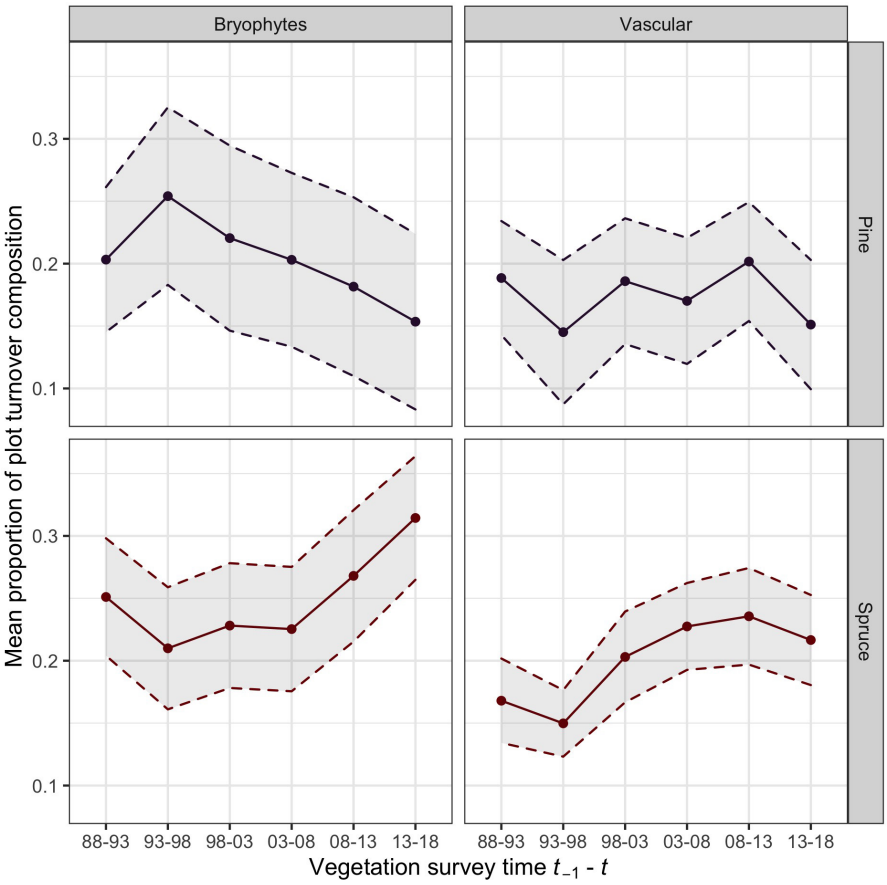

FIGURE 2 Compositional turnover over 5-year periods based upon the vegetation surveys, expressed as averages of plot values for total turnover and calculated separately for pine and spruce subsets. Confidence intervals (95%) are indicated by grey shaded areas delimited by dotted lines in the respective colours

Pleurozium schreberi. The mean number of vascular plants registered in the soil eDNA survey per plot differed among subsets but was similar for bryophytes (spruce forest subset: vascular plants: 37.46 ± 11.36 *SD*; bryophytes: 3.27 ± 2.21 *SD*; pine forest subset: vascular plants: 29.66 ± 9.08 *SD*; bryophytes: 4.41 ± 2.51 *SD*; Table S3).

3.3 | Matching taxa and temporal resolution of eDNA survey data

A total of 53 vascular plants and 11 bryophytes were registered both in the vegetation and the eDNA survey (*matching taxa*), accounting for 81% (53/65) and 13% (8/68) of the total vegetation survey, respectively (a summary is found in Table S3). The mean proportion values from the *number of matching taxa in the total and detectable vegetation survey* per plot were 0.60 ± 0.18 *SD* for the vascular and 0.10 ± 0.12 *SD* for bryophytes, respectively. For both taxa groups, the *number of matching taxa* and the *number of detectable taxa in total vegetation survey* were positively correlated (vascular plants: Spearman's ρ = 0.891, p = < 2.2e-16, n = 100; bryophytes: Spearman's ρ = 0.219, p = 0.028, n = 39; Figure 3). The proportion of *number of matching taxa in total & detectable vegetation survey* did not differ between spruce and pine subsets (Wilcoxon–Mann–Whitney rank-sum tests: pine subset: W = 1,752, p = 0.632; spruce subset: W = 1,963, p = 0.274; Figure S2).

Plot-wise, the *number of matching taxa* calculated for the soil eDNA survey with respect to each of the seven detectable vegetation surveys (t_1, ...,t_7) varied between 0.55 and 0.7 across years for vascular plants, and between 0.10 and 0.20 for bryophytes (Figure 4). For both taxonomic groups, the highest mean *number of matching taxa in detectable vegetation survey t* was observed for the last vegetation survey (t_7 = 2018), the year soil eDNA was sampled (Figure 4). Thus, subsequent analyses were made only comprising taxa from plot compositions registered on the 2018 vegetation survey and referred to as the 'best detected vegetation survey' (t_{bdvs} = t_7 = 2018). However, the mean *number of matching taxa in detectable vegetation survey t* for vascular plants only differed significantly between four last vegetation surveys and the first survey made (Figure 4; Friedman test: χ^2 = 24.005, df = 6, p = 0.0005211; *Post-hoc* Tukey test: t_7–t_1: z = 4.137, p < 0.001, t_6–t_1: z = 3.577, p = 0.006, t_5–t_1: z = 3.074, p = 0.034; t_4 = –t_1: z = 3.076, p = 0.341) and, for bryophytes, significant differences were found only between the last and the three first surveys (Figure 4; Friedman test: χ^2 = 41.35, df = 6, p < 0.001; *Post-hoc* Tukey test: t_7–t_1: z = 3.409, p = 0.0117, t_7–t_2: z = 4.499, p < 0.001, t_7–t_3: z = 3.790, p = 0.003). The low *number of matching taxa in total and detectable vegetation survey* resulted from 132 and 142 unique appearances of vascular plants and bryophytes across the seven individual surveys, respectively (Figure 4).

Moreover, in each plot, the proportions of unchanged composition between survey t_{bdvs} and t_1, ...,t_6 varied from 0.4 to 1 for vascular plants. However, these proportions were not correlated to the

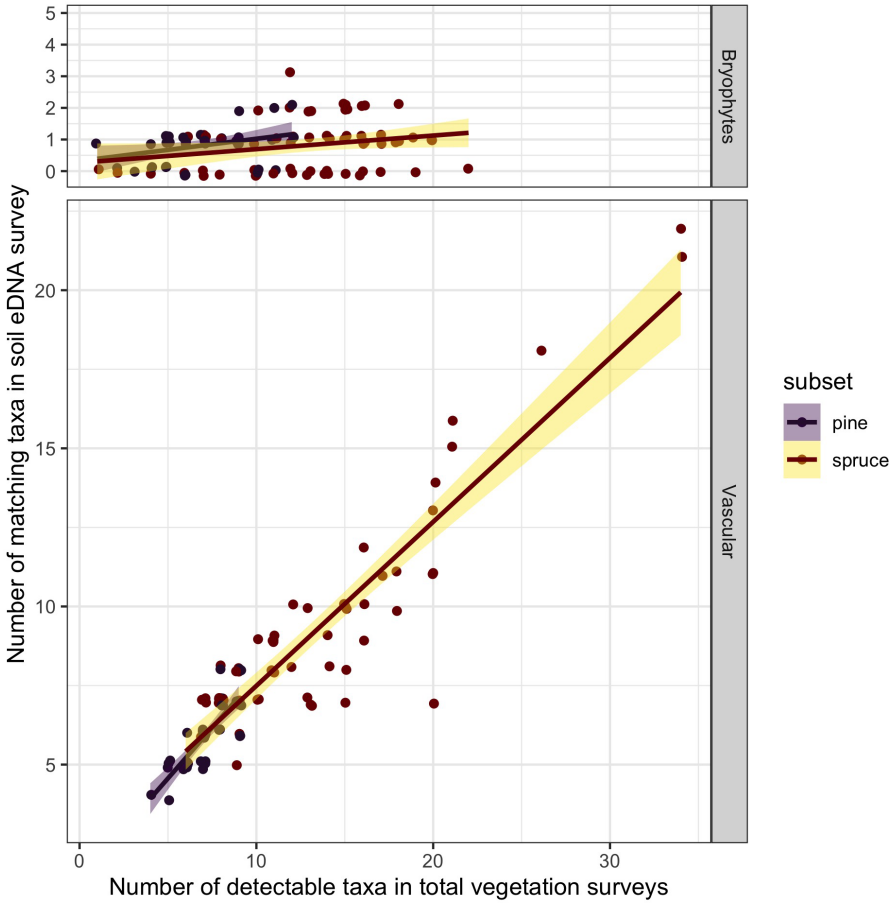

FIGURE 3 Matching taxa between the total vegetation survey and the soil eDNA survey. The total number of detectable taxa in each 1-m^2 plot of the spruce or pine subset registered across seven vegetation surveys is related to the number of identical taxon names recorded in the soil eDNA survey. Lines represent linear models for *number of matching taxa in soil eDNA survey* regressed on *the number of detectable taxa in total vegetation survey*; shaded areas are 95% confidence level intervals for model predictions. To avoid spatial overlap points are jittered by up to 0.15 units along both axes when necessary

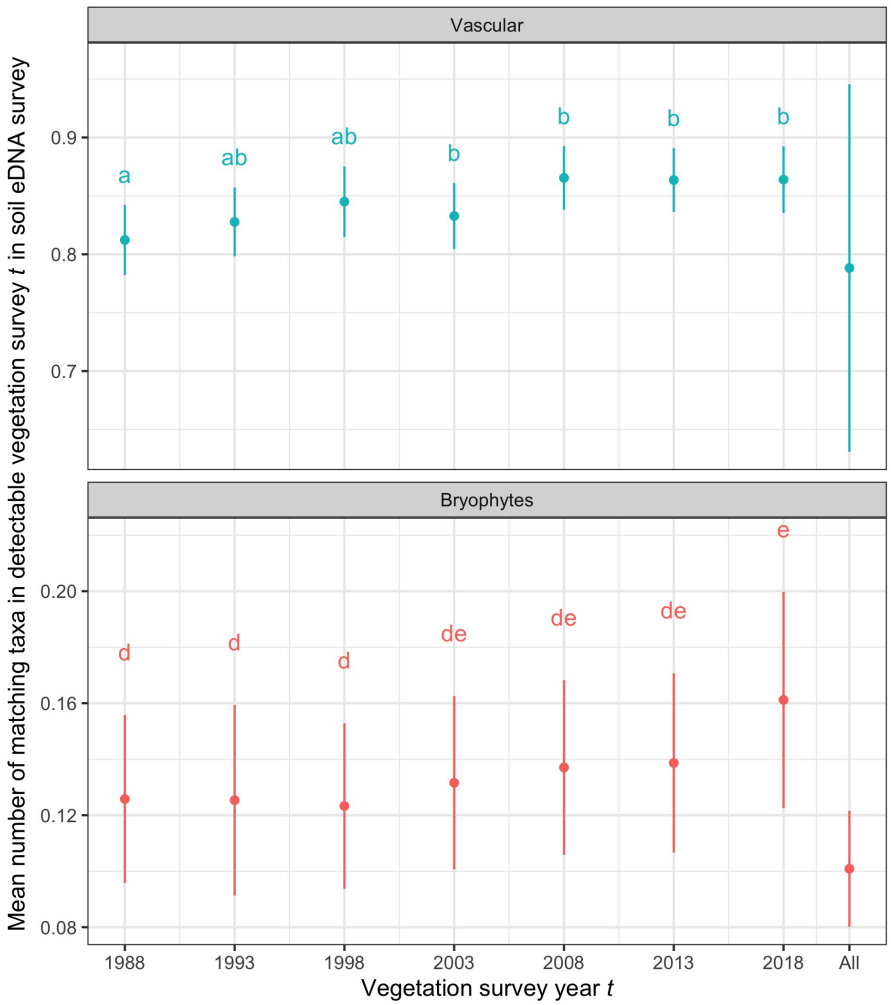

FIGURE 4 Matching taxa in each detectable vegetation survey and temporal resolution of soil eDNA survey. The number of matching taxa (the number of identical taxon names registered in both detectable vegetation and soil eDNA surveys) are calculated per plot for each plant group, and for compositions from each vegetation survey (t_1, ...,t_6) and for the total vegetation survey ('all'). The proportion of matching taxa in a vegetation survey *t* is the fraction of number of matching taxa out of the detectable taxon count in a vegetation survey t or in the total vegetation survey. Points and bars indicate means with 95% confidence intervals for the means. Identical small letters indicate non-significant differences ($p > 0.05$) in pairwise Friedman multiple comparison tests

proportion of matching taxa calculated for the same plot (Figures S3; Spearman's $\rho = -0.041$, $p = 0.316$). Conversely, for bryophytes, the plot proportions of unchanged composition between survey t_{bdvs} and t_1, ...,t_6 varied from 0 to 1 and these were slightly correlated to the proportion of matching taxa (Figures S3; Spearman's $\rho = 0.216$, $p < 0.05$).

Furthermore, the mean number of vascular plants and bryophytes that were registered only in a vegetation survey year were 18.57 ± 14.70 *SD* and 20.28 ± 7.27 *SD*, respectively (a summary for each vegetation survey *t* is found in Table S4). Consequently, the subset *temporal number of matching taxa* contained 40% of vascular plants from each vegetation survey year *t* and 6% of bryophytes only for years 1988, 1998 and 2003. Trees such as pine *Pinus sylvestris*, birch (*Betula* spp.) and European aspen *Populus tremula* were most common across years in this *temporal number of matching taxa* subset, but taxa from all plant forms were also detected.

3.4 | Spatial resolution of eDNA survey data

When comparing the soil eDNA survey to the best detected vegetation survey ($t_7 = 2018$), the majority of records of both vascular plants and bryophytes matched the vegetation survey at the 1-m^2 plot scale (*matching taxa*) or, at broader scale, other taxa from the total vegetation survey (*vegetation survey match*; Figure 5a). The large majority of sequence reads were assigned to taxa from these two categories (Figures S4). In addition, on average, 5.94 ± 2.70 *SD* vascular plants per plot from soil eDNA survey were not registered in any vegetation survey but were known from the Solhomfjell Forest Reserve (*regional flora match*; Table S3). The total number of taxa in this group was 51. These are mainly perennial shrubs, herbs and graminoids typical of boreal forests (Table S3 for a taxonomic overview and descriptive statistics). False positives, all vascular plants, were registered in 61 plots accounting for 8% of the soil eDNA survey (10/129) and less than 3% of the total reads (Figure S4).

For vascular plants, number of taxa registered in the best detected vegetation survey ($t_7 = 2018$) was positively correlated with *matching taxa* (Spearman's $\rho = 0.762$, $p < 0.001$, $n = 100$) and *regional flora match* variables (Spearman's $\rho = 0.375$, $p < 0.001$, $n = 100$), while the *vegetation survey match* (Spearman's $\rho = 0.088$, $p = 0.381$, $n = 100$) and false positives (Spearman's $\rho = 0.154$, $p = 0.125$, $n = 100$) were not significantly correlated. For bryophytes, positive correlations were found for *matching taxa* (Spearman's $\rho = 0.157$, $p = 0.117$, $n = 100$), whereas a negative correlation was found with *vegetation survey match* (Spearman's $\rho = 0.088$, $p = 0.380$, $n = 100$). A summary figure is found in Figure S5.

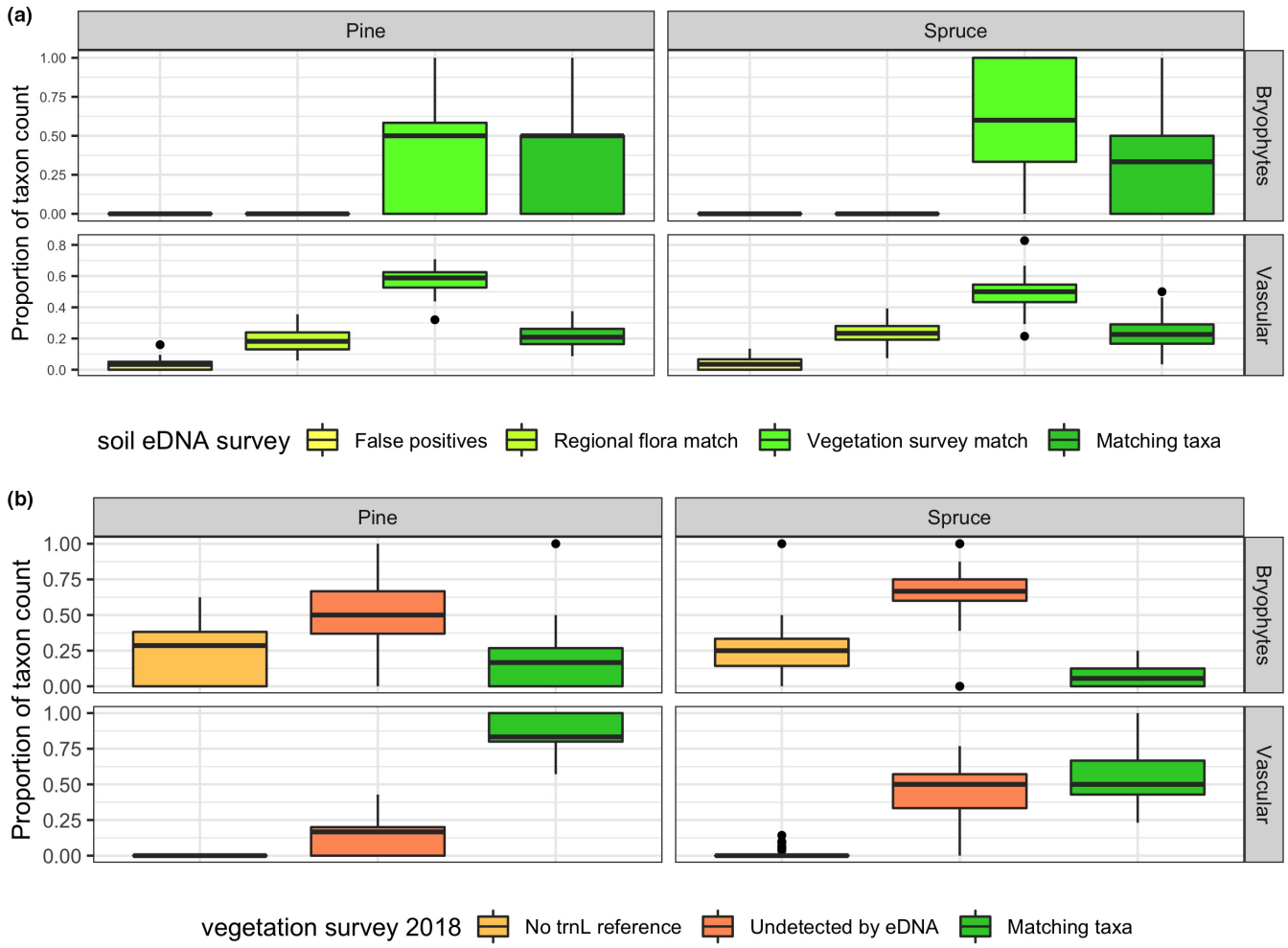

FIGURE 5 Spatial resolution of the soil eDNA survey and best detected vegetation survey. (a) Plant taxa detected with a single soil eDNA sample from each plot is assessed according to the spatial scale on which was recorded by categorising into: Matching taxa (taxa registered in the plot in the best detected vegetation survey $t_7 = 2018$); vegetation survey match (taxa recorded in any other plot in the best detected vegetation survey $t_7 = 2018$); regional flora match (taxa recorded in the Solhomfjell area outside plots); and false positives (taxa not recorded in the Solhomfjell area or with higher mean read count in PCR negatives than across samples and replicates). (b) Plant taxa recorded in each plot at the last vegetation survey $t_7 = 2018$ are categorised into: Matching taxa (taxa also registered in the soil eDNA survey); undetected by eDNA (trnL reference sequence available, but taxa were not registered in the soil eDNA survey); and no trnL reference available (no reference sequence was available for identification with eDNA data). For both panels, results are shown for taxon count proportions from each category and subsets with boxplots: horizontal bold lines depict median, the boxes show interquartile range and vertical lines indicate non-outlier range; the dots represent the outliers

3.5 | Representation of vegetation in soil DNA

On average, 6% ± 21 *SD* and 16% ± 19.00 *SD* of the vascular plants and bryophytes proportion registered in the best detected vegetation survey ($t_7 = 2018$), respectively, were also present in the soil eDNA survey (*matching taxa*; Figure 5b). The full vascular plant composition of this vegetation survey was recovered by the soil eDNA survey for 13 and 3 plots in the pine and spruce forest subsets, respectively (Figures S6). The majority of bryophytes recorded in the vegetation survey belonged to the *undetected by eDNA* category while for vascular plants the proportion of this category accounted for 0.32 of all taxa (Figure 5b). The proportion of vascular plants with *no trnL reference available* accounted for less than 0.03 of the total composition from only in six plots of the spruce subset, whereas for bryophytes this category accounted for more than 0.30 across plots of both pine and spruce subsets (Figure 5b).

The abundance of taxa registered in the best detected vegetation survey ($t_7 = 2018$) was significantly higher for taxa detected with soil eDNA than for taxa undetected by soil eDNA, for both vascular plants and bryophytes (Wilcoxon–Mann–Whitney rank-sum tests: vascular plants: $W = 19{,}386$, $p < 0.001$; bryophytes: $W = 16{,}986$, $p = 0.0087$; Figure 6). Sequence read counts for each taxa of ferns, forbs, mosses and trees taxa were slightly positively correlated to the respective abundances registered in the best detected vegetation survey, while more strongly positive correlations were found for graminoids, herbs and shrubs taxa (summary figure is found in Figure S7).

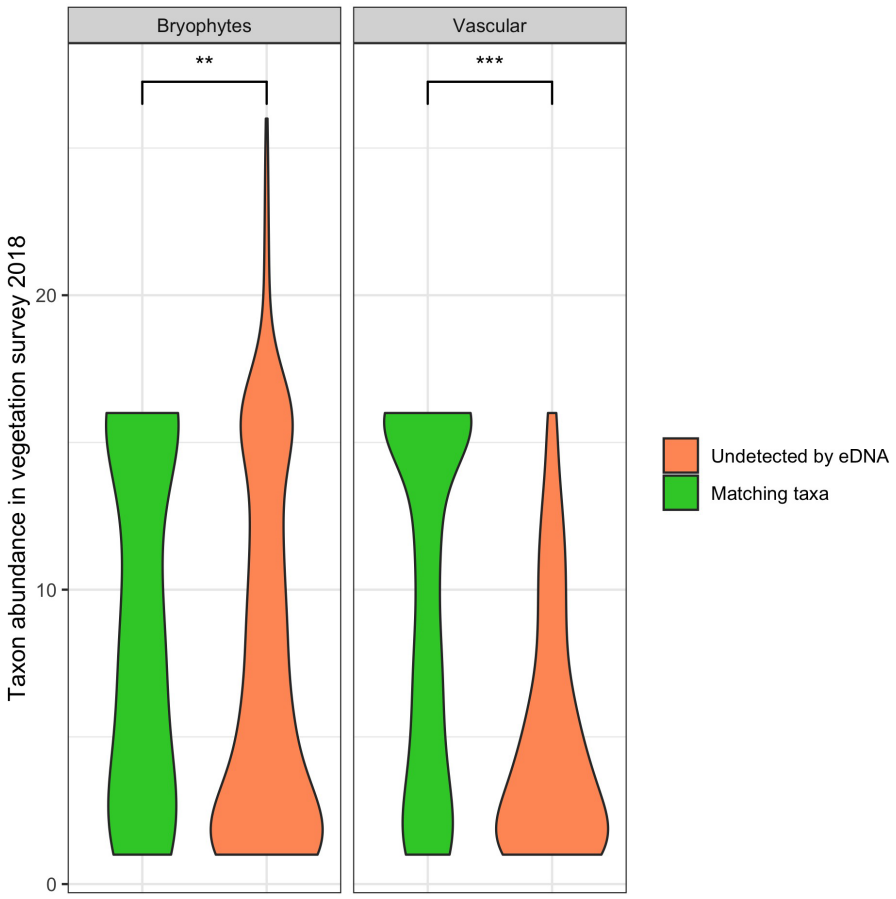

FIGURE 6 Correspondence between taxon detectability and abundance in the year of the best detected vegetation survey ($t_7 = 2018$). The composition of each plot from the best detected vegetation survey is categorised as matching taxa and undetected by eDNA. The distribution of summed subplot frequency (proxy for biomass) from each taxon in these two categories is shown by violin plots: Shape width is proportional with the frequency of observations. Subplot frequency corresponds to taxon presence in 625 cm^2 grids, and values >16 correspond to taxa with identical trnL (UUA) intron sequences that were merged into one single taxon with their subplot frequencies summed. Significance of differences between biomass from taxa within both categories was tested using a Wilcoxon–Mann–Whitney rank-sum test and are shown with ** and *** for p values <0.05 and <0.01, respectively

4 | DISCUSSION

4.1 | Vegetation survey inventories and its change across time

More bryophytes than vascular plants were registered in plots from the spruce subset than pine subset (Figure 1), and its average taxon counts in the total vegetation survey follow the same plot trends across all survey years (Data S3). Overstory dominance by pine or spruce influences the understory vegetation by different effects on ground-level light conditions (Felton et al., 2020), with lower light levels in the spruce forest (Esseen et al., 1997). Reduced richness in pine forest from the Solhomfjell area has been attributed to severe effects of drought from its shallow soils (Økland & Eilertsen, 1996). Furthermore, nutrient-demanding taxa are restricted to the richer and deeper soils that in the Solhomfjell area are found in spruce forests only (Økland & Eilertsen, 1993). Dwarf shrubs such as bilberry and lingonberry, and mosses such as feather moss and shaded wood-moss, registered in the majority of plots, are common and dominant species in boreal understory forest vegetation (Nilsson & Wardle, 2005; Økland et al., 2004).

An average of one to two vascular plants and bryophytes were either lost or gained between consecutive vegetation surveys, respectively, and fluctuations observed in plots located below overstories dominated by pine were generally lower than spruce (Figure 2). Both trends were also observed in Solhomfjell plots analysed every year from 1988 to 1993 (Økland & Eilertsen, 1996), and in other boreal forests in the south of Norway between 1988 and 2003 (Økland et al., 2004). This suggests these turnover rates are representative for Norwegian boreal forests over the last 30 years. The latter studies also detected a decrease in vascular plant richness, notably in spruce forests, and increase in large bryophytes in both forest types, attributed to a combination of past soil acidification due to higher deposition of air pollutants and longer and warmer growth seasons. Though we did not detect steady trends of gain or loss of vascular plants and bryophytes between consecutive years of vegetation survey, we detected peaks that may correspond to exacerbation of these climatic conditions (Figures S1).

4.2 | Soil eDNA survey and detection of taxa

The soil eDNA survey consisted of 127 taxa assigned mostly to species level (63.77%), from which 53 were vascular plants not registered in the total vegetation survey (Table S2). In all, 57 taxon detections accounted for 81% and 13% of vascular and bryophyte taxa registered in the total vegetation survey, respectively (Table S2). Soil eDNA surveys in temperate and tropical forests, tundra and deserts have routinely found 'hidden taxa' that were not observed when surveying above-ground diversity using vegetation surveys (Carrasco-Puga et al., 2021; Edwards et al., 2018; Osathanunkul et al., 2021; Palacios et al., 2021; Yoccoz et al., 2012). Small herbaceous plants

and seedlings can be easily overlooked while surveying vegetation, and eDNA-based assessments detect DNA rather than organisms (Deiner et al., 2017). The pool of plant DNA in top soils is composed of locally deposited debris, roots, rhizomes and seeds. In addition, it may contain local and/or exotic pollen, but this does not seem to contribute to the local eDNA signal (Edwards et al., 2018). Metabarcoding analysis of root diversity has highlighted the DNA contribution of many perennial plants that persist below-ground even in the temporary absence of above-ground parts, which, in turn, increases below-ground richness estimates compared to above-ground (Pärtel et al., 2012; Rucińska et al., 2022; Träger et al., 2019). In our study, soil eDNA detections of taxa not registered in the vegetation surveys but present in the Solhomfjell area are indeed mostly perennial plants (Figure 5a; Table S2), suggesting that these are most likely local signals from plants growing in between vegetation survey plots or seedlings not recruited in the plot.

Reference sequences for the trnL (UUA) intron P6 loop were available for about 95% and 75% of the vascular and bryophyte taxa registered across all vegetation surveys, respectively, and the taxonomic resolution was 75% and 65% identified to species level. While longer markers such as ITS, matK and rbcL in general may provide higher taxonomic resolution, the actual taxonomic resolution obtained depends on the marker region used, the representation in the reference library and the size and nature of the local flora (Hollingsworth et al., 2016). The P6 loop of the chloroplast trnL (UUA) intron (Taberlet et al., 2007) is the most commonly used marker for soil eDNA studies targeting vascular plants (Capo et al., 2021; Parducci et al., 2017), as its short sequence may be found in the degraded DNA that is typically present in underground decomposed material and sediments (Taberlet et al., 2007). This primer is designed to target vascular plants, and our results on vascular plants show high detection and high taxonomic resolution similar to other studies that are based on this primer (e.g. Alsos et al., 2018; Edwards et al., 2018). Other primer pairs for the P6 loop such as *c-d* (Taberlet et al., 2007) and Bryo_P6 (Epp et al., 2012) are conserved from Bryophytes to Angiosperms, but these markers have not been widely used and very few reference sequences exist in public repositories (Boukhdoud et al., 2021; Soininen et al., 2017). Though other nuclear ribosomal (ITS) and chloroplast (rbcL) markers may yield higher specificity in bryophytes (Lang et al., 2014; Liu et al., 2010), detection of bryophytes with eDNA-based assessments is limited by low intraspecific variation in marker regions (Hassel et al., 2013) and particularly for boreal forests when relying on non-exhaustive reference libraries. However, the fact that we did detect bryophytes in almost every soil eDNA sample suggests that there is a great potential for eDNA also for this group, but we recommend further development of primer design and build up of reference library

The detection of taxa was related to abundance, which, in turn, is an expression of biomass. The soil eDNA survey failed to detect some taxa registered in vegetation surveys even when reference sequences were available (Figure 6), but most of these had significantly lower plot abundance than those that were detected (Figure 6). A positive relation between detectability and plant biomass has been also observed in tundra and temperate sites (Alsos et al., 2018; Edwards et al., 2018; Yoccoz et al., 2012) and in aquatic environments (Alsos et al., 2018; Anglès d'Auriac et al., 2019; Matsuhashi et al., 2016). The relation is often attributed to the greater chance of deposited or suspended plant DNA that can be detected with higher organismal biomass. However, our study also reports detections of taxa present in only one out of the 16,625 cm^2 subplots of a vegetation plot (Figure 6), demonstrating that soil eDNA metabarcoding also detects some rare taxa. In addition, the apparent stochasticity of rare taxa raises questions on how biomass differences between root and shoot at different life-history stages (Qi et al., 2019) can potentially underlie the detection of less abundant taxa. Furthermore, in metabarcoding studies, sequence read counts are often interpreted as a proxy for abundance since DNA template availability for PCR amplification covariates with biomass (Amend et al., 2010; Beng & Corlett, 2020; Deagle et al., 2019). However, the signal of biomass can be diluted by technical and biological biases in marker recovery rates across different taxa (Deiner et al., 2017). In our study, sequence read counts assigned to taxa from all considered growth forms registered in the last year of vegetation survey correlated positively to their plot abundance, and these correlations were significant for forb, graminoid, moss and shrub taxa as well as for all life-forms combined (See Figure S7 for figures and *p* values). The correlations for ferns and trees were not significant, and this might be due to the abundance being skewed by the larger aerial vegetation cover in relation to the smaller underground cover of these taxa, as may decrease the DNA contribution of roots and rhizomes to the soil eDNA pool in relation to other growth forms (Qi et al., 2019).

4.3 | Matching taxa and temporal calibration

Each plot has a measure of the total number of unique taxa registered during the seven survey years as well as the number of taxa detected in the single eDNA survey. On average, 60% and 10% of the vascular and bryophyte taxa from the total vegetation survey matched with a single soil eDNA survey, respectively (Figure 4; Table S3). Conversely, 18% and 30% of the vascular and bryophyte taxa registered by a single soil eDNA survey matched with the total vegetation survey, respectively (Table S3). Similar rates of undersampled vegetation, that is, taxa that were detected in only one of the two surveys (Edwards et al., 2018), and matching taxa values, have been found when comparing surveys at similar and even larger plot scales (1-m^2 plots in alpine subarctic vegetation Kumpula, 2020; 1–4 m radius from circular plots in Svalbard tundra Edwards et al., 2018; and 15-m^2 plots in Varanger boreal forest Yoccoz et al., 2012), highlighting the ability of both proxies to assess the total vegetation of a site. In our study, we show that the number of matching taxa increased with richness registered in the total vegetation survey in both pine and spruce subsets (Figure 3), and this suggests that differences in soil and vegetation properties in these two environments probably have no effect on detectability. Furthermore, surveys built by both

proxies detected the same dominant taxa. This highlights the ability of both proxies to recover the main components of the vegetation.

Our soil eDNA survey made in 2018 on average matched best with the plot composition surveyed the same year. This holds true for vascular plants as well as for bryophytes. Roots and shoots from live plants and derived litter are probably the biggest contributors to plant DNA in the soil, and also the least degraded and therefore more likely to be detected. Although the match between vascular plant compositions registered each vegetation survey year and the soil eDNA survey were similar, these similarities were not correlated to soil eDNA detections of composition persisting (or unchanged) across surveys (Figure S3; Figure 2). Soil eDNA detections of vascular plants registered uniquely at a vegetation survey *t* also support this (Table S4). Altogether, our study indicates that a single soil eDNA survey can detect taxa from multiple vegetation surveys and its turnover fraction across time with the same power, and pinpoint how soil eDNA samples can encapsulate the vascular composition going back at least 30 years. Detections of past signals from 30 up to 50 years ago have also been found in crop soils (Foucher et al., 2020; Yoccoz et al., 2012), though these are more likely to be detected since plant biomass exponentially increases in monocultures. However, unique past signals detected in our study correspond to taxa with median to low abundance, registered in 50% or fewer subplots. This suggests that biomass may not play a role in the detectability of past signals in natural environments (Data S4). Nevertheless, detections of past signals can be also attributed to the resurfacing of deep soil particles through bioturbation by biotic underground DNA transporters such as insects, moles, worms, etc. (Prosser & Hedgpeth, 2018). The match with the composition of bryophytes registered in the best detected vegetation survey ($t_7 = 2018$) was significantly higher than with the rest of the vegetation surveys and significantly correlated with the proportion of unchanged taxa (Figures S3). This indicates that the soil eDNA survey mainly detected a similar fraction of bryophyte composition from each vegetation survey. Mosses and liverworts are poorly detected, and this is probably due to a combination of factors including mismatch of the trnL (UUA) intron *g-h* binding site for these taxa, but also that their DNA is probably underrepresented in the soil pool in comparison to that of vascular plants as most of their biomass is allocated in the forest floor making them less detectable over years (Bergamini et al., 2001).

4.4 | Spatial patterns of detection

On average, about 55% of the soil eDNA survey composition in each plot matched the composition registered in any other (near) plots or in the Solhomfjell area, whereas about 22% matched the vegetation plot composition (Figure 5a), suggesting that soil eDNA samples reflect mainly local vegetation rather than plot specific signals. Our results contrast those from Edwards et al. (2018) in Svalbard tundra where soil eDNA signals were highly specific to those recorded within a circular plot of 1 m radius and no taxa existing beyond a 4 m radius were found. Instead, our results are more in concordance with <1 km signals speculated by Yoccoz et al. (2012), as these authors did not find signals in uncultivated meadows from crops located a kilometre away but found signals that are likely part of the regional species pool. Differences in taxa richness and vegetation distribution between tundra (low, homogeneous) and temperate forests (high, heterogeneous) such as that of Solhomfjell area led us to consider how distribution patterns of vegetation may hinder the spatial recovery of soil eDNA signals. Furthermore, the hilly Solhomfjell landscape in which the plots are located may contribute to DNA transport from one plot to another via snow-melt, rainfall run-off and through-flow, thus enabling detections from other plots. Although we attempted to calibrate each eDNA detection with a match to an area, that is, a match to the vegetation survey is a match to an area >1 m^2, our categories may disguise a temporal match within plots (a match to taxa detected in the same plot back in time). Our study is the first to assess simultaneously the temporal and spatial resolution of soil eDNA samples in natural environments, yet our approach cannot disentangle the contribution of each signal separately. This limitation highlights the need for studies in controlled microcosms where plant signals can be followed with biomarkers, thus enabling the possibility of tracing both the spatial and temporal signals. Nevertheless, our taxa comparison in space allowed us to identify plant richness detected by both methods with similar sampling efforts (surveying 1-m^2 plot vs taking a soil sample at the centre of 1-m^2 plot; Figure S5) and so provide a baseline for decision-making when designing sampling for soil eDNA assessments.

4.5 | Limitations and considerations for soil eDNA for plant diversity assessments

Soil and sediments are suitable substrates for eDNA-based plant assessments in terrestrial environments, as most extra-organismal and organismal plant DNA from both active and dormant tissues are gathered or ultimately deposited in these substrates (Rodriguez-Ezpeleta et al., 2021). Accordingly, soil eDNA assessment is a valuable tool for identification of plant diversity at any season, especially when non-destructive and easy sampling is needed. Our study has shown how a single eDNA sample can signal local dominant flora and thus might be useful for general plant diversity assessments. However, when detection of less abundant and/or rare taxa is desired, collecting multiple eDNA soil samples is recommended. Moreover, our study shows how a single soil eDNA survey can provide a series of local, regional, past and present plant signals that can help track long-term responses to climate and ecosystem changes. However, as with any method, there are some limitations to consider before embarking on a plant soil eDNA study (see Figures S8 for a summary of methodological steps). Since eDNA-based organismal detections are dependent on both DNA presence (intracellular or extracellular) and environmental conditions that may enhance or diminish DNA permanence, degradation and/or decay (Nagler et al., 2018; Pietramellara et al., 2009; Rodriguez-Ezpeleta et al., 2021), an evaluation of the potential state of DNA given the study environment is

essential before choosing an appropriate approach. Plant eDNA from soil substrates is particularly subject to degradation or decay from decomposition processes of organic matter by both underground and above-ground biota (Pietramellara et al., 2009). Thus, long DNA fragments are expected to account for the lowest fraction of target soil eDNA that can be isolated. In tropical areas, warmer environments and richer decomposing communities may exacerbate DNA degradation rates in soils (Pietramellara et al., 2009). Furthermore, DNA decay increases with time and past plant signals may be only in the form of short DNA fragments (Kistler et al., 2017). For these reasons, it is generally recommended to employ markers targeting short DNA sequences. Indeed, most plant eDNA-based studies employ the chloroplast trnL (UAA) intron which amplifies on average a 50 bp region that has been robustly catalogued for the flora of Fennoscandia (Alsos et al., 2020). Noteworthy, in temperate areas, studies have successfully amplified matK, rbcL, ITS2 markers with target regions of more than 490 bp (Fahner et al., 2016). Furthermore, the use of short markers may be hindered by their low variability across species resolutions, that is, that identification of related taxa is supported by a few base pairs only (Taberlet et al., 2007). Employing short markers may thus require stricter thresholds of OTU (head sequence) matching to a reference sequence as the probability of identification mismatching resulting from polymerase errors is amplified. Although the prospecting of new plant DNA markers with targeted capture of multiple informative genes is promising, eDNA-based assessments can only identify taxa present in a reference sequence library. Thus, an eDNA assessment is only as good as its reference library. Although correlations between plant biomass and DNA concentration in the environmental samples are poorly understood, several studies—including this one—show that read counts may be used as a proxy for biomass of some plant life-forms (Deagle et al., 2013; Deiner et al., 2021). Finally, as DNA may remain in the environment after the organism is no longer present (Harrison et al., 2019), taxon detections provided by eDNA-based assessments should be interpreted merely as detections of organismal DNA. If one is interested in detections of live organisms, RNA approaches should be considered. These limitations highlight the importance of considering the current literature carefully to ensure that the study design is suited to address feasible and measurable questions.

5 | CONCLUSIONS

In this study, we investigated the temporal and spatial resolution of soil eDNA surveys of plant biodiversity to interpret the utility of this approach to effectively assess biodiversity and monitor vegetation changes through time and space. Our results show that a combination of aboveground vegetation surveys and soil eDNA surveys yields the most comprehensive inventory of plant diversity for a site. In particular, a single soil eDNA sample mainly detects local plant diversity rather than site specific diversity. In addition, a soil eDNA sample captures plant diversity going back at least 30 years in time while matching most closely with current diversity. Similarly, we find that soil eDNA samples can be useful to detect both rare and unrecorded taxa, but are best at detecting abundant taxa. Our results highlight the potential of soil eDNA surveys to monitor vegetation responses over broader spatial and temporal scales, and encourage a rethinking of the optimal strategies for assessment of vegetation if soil eDNA is used as a method.

ACKNOWLEDGEMENTS

Morphological surveys were conducted as a part of a project from the Natural History Museum, University of Oslo. We recognise the great effort put into these surveys by their participants over 30 years. We like to thank Andreas Wollan for his support during soil sampling and plot recognition in the Solhomfjell Forest Reserve, Jarl Andreas Anmarkrud, Lisbeth Torbek and Audun Schrøder-Nielsen for their support in the laboratory for DNA extraction, library building and sequencing, Youri Lammers for his support in metabarcoding bioinformatics, Marcella Orwick Rydmark for her support and valuable feedback in manuscript writing, Rebecca Angelica Blakeney for proof-reading the manuscript. This research is part of the H2020 MSCA-ITN-ETN Plant.ID network, and has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 765000.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHORS' CONTRIBUTIONS

M.A., I.G.A., R.H. and H.J.d.B. conceived the study and methodology; M.A. collected the data and built the metabarcoding libraries in the laboratory; M.A. analysed the data with scientific and coding advice from I.G.A., R.H., and B.F., Q.M., respectively; M.A. and H.J.d.B. drafted the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13865.

DATA AVAILABILITY STATEMENT

Raw sequences from both MiSeq runs (F and R) and bash and R codes are available in Ariza et al. (2022).

ORCID

María Ariza https://orcid.org/0000-0001-8396-4902
Bertrand Fouks https://orcid.org/0000-0002-3675-3499
Quentin Mauvisseau https://orcid.org/0000-0003-1215-2987
Rune Halvorsen https://orcid.org/0000-0002-6859-7726
Inger Greve Alsos https://orcid.org/0000-0002-8610-1085
Hugo J. de Boer https://orcid.org/0000-0003-1985-7859

REFERENCES

Alsos, I. G., Lammers, Y., Yoccoz, N. G., Joergensen, T., Sjoegren, P., Gielly, L., Edwards, M. E., Alsos, I. G., Lammers, Y., Yoccoz, N. G.,

Jørgensen, T., Sjögren, P., Gielly, L., & Edwards, M. E. (2018). Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. *PLOS ONE*, *13*(4), e0195403. https://doi.org/10.1371/journal.pone.0195403

Alsos, I. G., Lavergne, S., Merkel, M. K. F., Boleda, M., Lammers, Y., Alberti, A., Pouchon, C., Denoeud, F., Pitelkova, I., Pușcaș, M., Roquet, C., Hurdu, B.-I., Thuiller, W., Zimmermann, N. E., Hollingsworth, P. M., & Coissac, E. (2020). The treasure vault can be opened: Large-scale genome skimming works well using herbarium and silica gel dried material. *Plants*, *9*(4), 432. https://doi.org/10.3390/plants9040432

Amend, A. S., Seifert, K. A., & Bruns, T. D. (2010). Quantifying microbial communities with 454 pyrosequencing: Does read abundance count? *Molecular Ecology*, *19*(24), 5555–5565. https://doi.org/10.1111/j.1365-294X.2010.04898.x

Anglès d'Auriac, M. B., Strand, D. A., Mjelde, M., Demars, B. O. L., & Thaulow, J. (2019). Detection of an invasive aquatic plant in natural water bodies using environmental DNA. *PLoS ONE*, *14*(7), e0219700. https://doi.org/10.1371/journal.pone.0219700

Ariza, M., Fouks, B., Mauvisseau, Q., Halvorsen, R., Alsos, I.G. & de Boer, J. H. (2022). Using soil eDNA for plant diversity assessments. Zenodo, https://doi.org/10.5281/zenodo.6400498

Beng, K. C., & Corlett, R. T. (2020). Applications of environmental DNA (eDNA) in ecology and conservation: Opportunities, challenges and prospects. *Biodiversity and Conservation*, *29*(7), 2089–2121. https://doi.org/10.1007/s10531-020-01980-0

Bergamini, A., Pauli, D., Peintinger, M., & Schmid, B. (2001). Relationships between productivity, number of shoots and number of species in bryophytes and vascular plants. *Journal of Ecology*, *89*(6), 920–929. https://doi.org/10.1111/j.1365-2745.2001.00613.x

Bertrand, R., Lenoir, J., Piedallu, C., Riofrío-Dillon, G., de Ruffray, P., Vidal, C., Pierrat, J.-C., & Gégout, J.-C. (2011). Changes in plant community composition lag behind climate warming in lowland forests. *Nature*, *479*(7374), 517–520. https://doi.org/10.1038/nature10548

Boukhdoud, L., Saliba, C., Parker, L. D., Rotzel McInerney, N., Ishak Mouawad, G., Kharrat, M., Kahale, R., Chahine, T., Maldonado, J. E., & Bou Dagher-Kharrat, M. (2021). First DNA sequence reference library for mammals and plants of the Eastern Mediterranean Region. *Genome*, *64*(1), 39–49. https://doi.org/10.1139/gen-2019-0194

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: Aunix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*(1), 176–182. https://doi.org/10.1111/1755-0998.12428

Brunbjerg, A. K., Bruun, H. H., Dalby, L., Fløjgaard, C., Frøslev, T. G., Høye, T. T., Goldberg, I., Læssøe, T., Hansen, M. D. D., Brøndum, L., Skipper, L., Fog, K., & Ejrnæs, R. (2018). Vascular plant species richness and bioindication predict multi-taxon species richness. *Methods in Ecology and Evolution*, *9*(12), 2372–2382. https://doi.org/10.1111/2041-210X.13087

Capo, E., Giguet-Covex, C., Rouillard, A., Nota, K., Heintzman, P. D., Vuillemin, A., Ariztegui, D., Arnaud, F., Belle, S., Bertilsson, S., Bigler, C., Bindler, R., Brown, A. G., Clarke, C. L., Crump, S. E., Debroas, D., Englund, G., Ficetola, G. F., Garner, R. E., … Parducci, L. (2021). Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: Overview and recommendations. *Quaternary*, *4*(1), 6. https://doi.org/10.3390/quat4010006

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*(5), 335–336. https://doi.org/10.1038/nmeth.f.303

Carrasco-Puga, G., Díaz, F. P., Soto, D. C., Hernández-Castro, C., Contreras-López, O., Maldonado, A., Latorre, C., & Gutiérrez, R. A. (2021). Revealing hidden plant diversity in arid environments. *Ecography*, *44*(1), 98–111. https://doi.org/10.1111/ecog.05100

Chen, I.-C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science*, *333*(6045), 1024–1026. https://doi.org/10.1126/science.1206432

Corona, P., Chirici, G., McRoberts, R. E., Winter, S., & Barbati, A. (2011). Contribution of large-scale forest inventories to biodiversity assessment and monitoring. *Forest Ecology and Management*, *262*(11), 2061–2069. https://doi.org/10.1016/j.foreco.2011.08.044

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., Kartzinel, T. R., & Eveson, J. P. (2019). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology*, *28*(2), 391–406. https://doi.org/10.1111/mec.14734

Deagle, B. E., Thomas, A. C., Shaffer, A. K., Trites, A. W., & Jarman, S. N. (2013). Quantifying sequence proportions in a DNA-based diet study using Ion Torrent amplicon sequencing: Which counts count? *Molecular Ecology Resources*, *13*(4), 620–633. https://doi.org/10.1111/1755-0998.12103

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Deiner, K., Yamanaka, H., & Bernatchez, L. (2021). The future of biodiversity monitoring and conservation utilizing environmental DNA. *Environmental DNA*, *3*(1), 3–7. https://doi.org/10.1002/edn3.178

Edwards, M. E., Alsos, I. G., Yoccoz, N., Coissac, E., Goslar, T., Gielly, L., Haile, J., Langdon, C. T., Tribsch, A., Binney, H. A., von Stedingk, H., & Taberlet, P. (2018). Metabarcoding of modern soil DNA gives a highly local vegetation signal in Svalbard tundra. *The Holocene*, *28*(12), 2006–2016. https://doi.org/10.1177/0959683618798095

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., Erséus, C., Gusarov, V. I., Edwards, M. E., Johnsen, A., Stenøien, H. K., Hassel, K., Kauserud, H., Yoccoz, N. G., Bråthen, K. A., Willerslev, E., Taberlet, P., Coissac, E., & Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Molecular Ecology*, *21*(8), 1821–1833. https://doi.org/10.1111/j.1365-294X.2012.05537.x

Esseen, P.-A., Ehnström, B., Ericson, L., & Sjöberg, K. (1997). Boreal forests. *Ecological Bulletins*, *46*, 16. Retrieved from http://www.jstor.org/stable/20113207

Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M., & Ravel, J. (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*, *2*(1), 6. https://doi.org/10.1186/2049-2618-2-6

Fahner, N. A., Shokralla, S., Baird, D. J., & Hajibabaei, M. (2016). Large-scale monitoring of plants through environmental DNA metabarcoding of soil: Recovery, resolution, and annotation of four DNA markers. *PLoS ONE*, *11*(6), e0157505. https://doi.org/10.1371/journal.pone.0157505

Felton, A., Petersson, L., Nilsson, O., Witzell, J., Cleary, M., Felton, A. M., Björkman, C., Sang, Å. O., Jonsell, M., Holmström, E., Nilsson, U., Rönnberg, J., Kalén, C., & Lindbladh, M. (2020). The tree species matters: Biodiversity and ecosystem service implications of replacing Scots pine production stands with Norway spruce. *Ambio*, *49*(5), 1035–1049. https://doi.org/10.1007/s13280-019-01259-x

Fløjgaard, C., Frøslev, T. G., Brunbjerg, A. K., Bruun, H. H., Moeslund, J., Hansen, A. J., & Ejrnæs, R. (2019). Predicting provenance of forensic soil samples: Linking soil to ecological habitats by metabarcoding

and supervised classification. *PLoS ONE*, *14*(7), e0202844. https://doi.org/10.1371/journal.pone.0202844

Foucher, A., Evrard, O., Ficetola, G. F., Gielly, L., Poulain, J., Giguet-Covex, C., Laceby, J. P., Salvador-Blanes, S., Cerdan, O., & Poulenard, J. (2020). Persistence of environmental DNA in cultivated soils: Implication of this memory effect for reconstructing the dynamics of land use and cover changes. *Scientific Reports*, *10*(1), 10502. https://doi.org/10.1038/s41598-020-67452-1

Fraser, C. I., Connell, L., Lee, C. K., & Cary, S. C. (2018). Evidence of plant and animal communities at exposed and subglacial (cave) geothermal sites in Antarctica. *Polar Biology*, *41*(3), 417–421. https://doi.org/10.1007/s00300-017-2198-9

Hallett, L. M., Jones, S. K., MacDonald, A. A. M., Jones, M. B., Flynn, D. F. B., Ripplinger, J., Slaughter, P., Gries, C., & Collins, S. L. (2016). codyn: An R package of community dynamics metrics. *Methods in Ecology and Evolution*, *7*(10), 1146–1151. https://doi.org/10.1111/2041-210X.12569

Halvorsen, R., Bratli, H., & Wollan, A. K. (2019). Vegetasjonsundersøkelser av boreal barskog i Solhomfjell 2018. In Norsk institutt for naturforskning (NINA) (Ed.), *Terrestrisk naturovervåking i 2018: Markvegetasjon, epifytter, smågnagere og fugl. Sammenfatning av resultater*. Norsk institutt for naturforskning (NINA).

Halvorsen, R., Skarpaas, O., Bryn, A., Bratli, H., Erikstad, L., Simensen, T., & Lieungh, E. (2020). Towards a systematics of ecodiversity: The EcoSyst framework. *Global Ecology and Biogeography*, *29*(11), 1887–1906. https://doi.org/10.1111/geb.13164

Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proceedings of the Royal Society B. Biological Sciences*, *286*(1915), 20191409. https://doi.org/10.1098/rspb.2019.1409

Hartvig, I., Kosawang, C., Kjær, E. D., & Nielsen, L. R. (2021). Detecting rare terrestrial orchids and associated plant communities from soil samples with eDNA methods. *Biodiversity and Conservation*, *30*(13), 3879–3901. https://doi.org/10.1007/s10531-021-02279-4

Hassel, K., Segreto, R., & Ekrem, T. (2013). Restricted variation in plant barcoding markers limits identification in closely related bryophyte species. *Molecular Ecology Resources*, *13*(6), 1047–1057. https://doi.org/10.1111/1755-0998.12074

Hiiesalu, I., Opik, M., Metsis, M., Lilje, L., Davison, J., Vasar, M., Moora, M., Zobel, M., Wilson, S. D., & Pärtel, M. (2012). Plant species richness belowground: Higher richness and new patterns revealed by next-generation sequencing. *Molecular Ecology*, *21*(8), 2004–2016. https://doi.org/10.1111/j.1365-294X.2011.05390.x

Hollingsworth, P. M., Li, D.-Z., van der Bank, M., & Twyford, A. D. (2016). Telling plant species apart with DNA: From barcodes to genomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1702), 20150338. https://doi.org/10.1098/rstb.2015.0338

Kier, G., Mutke, J., Dinerstein, E., Ricketts, T. H., Küper, W., Kreft, H., & Barthlott, W. (2005). Global patterns of plant diversity and floristic knowledge. *Journal of Biogeography*, *32*(7), 1107–1116. https://doi.org/10.1111/j.1365-2699.2005.01272.x

Kistler, L., Ware, R., Smith, O., Collins, M., & Allaby, R. G. (2017). A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Research*, *45*(11), 6310–6320. https://doi.org/10.1093/nar/gkx361

Kumpula, K. (2020). *Systematic comparison of the relative accuracy of vegetation surveys and soil DNA metabarcoding: Assessing plant biodiversity at different spatial scales* (Master thesis). UMEA University. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-172130

Lang, A. S., Kruijer, J. D., & Stech, M. (2014). DNA barcoding of Arctic bryophytes: An example from the moss genus Dicranum (Dicranaceae, Bryophyta). *Polar Biology*, *37*(8), 1157–1169. https://doi.org/10.1007/s00300-014-1509-7

Liu, Y., Yan, H.-F., Cao, T., & Ge, X.-J. (2010). Evaluation of 10 plant barcodes in Bryophyta (Mosses). *Journal of Systematics and Evolution*, *48*(1), 36–46. https://doi.org/10.1111/j.1759-6831.2009.00063.x

Matsuhashi, S., Doi, H., Fujiwara, A., Watanabe, S., & Minamoto, T. (2016). Evaluation of the environmental DNA method for estimating distribution and biomass of submerged aquatic plants. *PLoS ONE*, *11*(6), e0156217. https://doi.org/10.1371/journal.pone.0156217

Mauvisseau, Q., Harper, L., Sander, M., Hanner, R. H., Kleyer, H., & Deiner, K. (2021). *The multiple states of environmental DNA and what is known about their persistence in aquatic environments*. Authorea, Inc. https://doi.org/10.22541/au.163638394.41572509/v1

Nagler, M., Insam, H., Pietramellara, G., & Ascher-Jenull, J. (2018). Extracellular DNA in natural environments: Features, relevance and applications. *Applied Microbiology and Biotechnology*, *102*(15), 6343–6356. https://doi.org/10.1007/s00253-018-9120-4

Nilsson, M.-C., & Wardle, D. A. (2005). Understory vegetation as a forest ecosystem driver: Evidence from the northern Swedish boreal forest. *Frontiers in Ecology and the Environment*, *3*(8), 421–428. https://doi.org/10.1890/1540-9295(2005)003[0421:UVAAFE]2.0.CO;2

Økland, R. H. & Eilertsen, O. (1993). *Vegetation-environment relationships of boreal coniferous forests in the Solhomfjellarea, Gjerstad, S Norway*. (No. 16; Sommerfeltia, pp. 1–254.). Botanical garden and museum. University of Oslo.

Økland, R. H., & Eilertsen, O. (1994). Canonical Correspondence Analysis with variation partitioning: Some comments and an application. *Journal of Vegetation Science*, *5*(1), 117–126. https://doi.org/10.2307/3235645

Økland, R. H., & Eilertsen, O. (1996). Dynamics of understory vegetation in an old-growth boreal coniferous forest, 1988–1993. *Journal of Vegetation Science*, *7*(5), 747–762. https://doi.org/10.2307/3236386

Økland, T., Bakkestuen, V., Økland, R. H., & Eilertsen, O. (2004). Changes in forest understorey vegetation in Norway related to long-term soil acidification and climatic change. *Journal of Vegetation Science*, *15*(4), 437–448. https://doi.org/10.1111/j.1654-1103.2004.tb02282.x

Osathanunkul, M., Sawongta, N., Pheera, W., Pechlivanis, N., Psomopoulos, F., & Madesis, P. (2021). Exploring plant diversity through soil DNA in Thai national parks for influencing land reform and agriculture planning. *PeerJ*, *9*, e11753. https://doi.org/10.7717/peerj.11753

Palacios, M., Curd, E., Edalati, K., Renshaw, M. A., Dunn, R., Potter, D., Fraga, N., Moore, J., Saiz, J., Wayne, R., & Parker, S. S. (2021). The utility of environmental DNA from sediment and water samples for recovery of observed plant and animal species from four Mojave Desert springs. *Environmental DNA*, *3*(1), 214–230. https://doi.org/10.1002/edn3.161

Parducci, L., Bennett, K. D., Ficetola, G. F., Alsos, I. G., Suyama, Y., Wood, J. R., & Pedersen, M. W. (2017). Ancient plant DNA in lake sediments. *The New Phytologist*, *214*(3), 924–942. https://doi.org/10.1111/nph.14470

Pärtel, M., Hiiesalu, I., Öpik, M., & Wilson, S. D. (2012). Below-ground plant species richness: New insights from DNA-based methods. *Functional Ecology*, *26*(4), 775–782. https://doi.org/10.1111/j.1365-2435.2012.02004.x

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M. T., Guerri, G., & Nannipieri, P. (2009). Extracellular DNA in soil and sediment: Fate and ecological relevance. *Biology and Fertility of Soils*, *45*(3), 219–235. https://doi.org/10.1007/s00374-008-0345-8

Porté, A. J., Samalens, J.-C., Dulhoste, R., Teissier Du Cros, R., Bosc, A., & Meredieu, C. (2009). Using cover measurements to estimate aboveground understorey biomass in Maritime pine stands. *Annales Des Sciences Forestières*, *66*(3), 307. https://doi.org/10.1051/forest/2009005

Prosser, C. M., & Hedgpeth, B. M. (2018). Effects of bioturbation on environmental DNA migration through soil media. *PLoS ONE*, *13*(4), e0196430. https://doi.org/10.1371/journal.pone.0196430

Qi, Y., Wei, W., Chen, C., & Chen, L. (2019). Plant root-shoot biomass allocation over diverse biomes: A global synthesis. *Global Ecology and Conservation*, *18*, e00606. https://doi.org/10.1016/j.gecco.2019.e00606

R Core Team. (2019). *R: A language and environment for statistical computing* (1.2.5001) [Computer software]. R Foundation for Statistical Computing.

Rodriguez-Ezpeleta, N., Morissette, O., Bean, C. W., Manu, S., Banerjee, P., Lacoursière-Roussel, A., Beng, K. C., Alter, S. E., Roger, F., Holman, L. E., Stewart, K. A., Monaghan, M. T., Mauvisseau, Q., Mirimin, L., Wangensteen, O. S., Antognazza, C. M., Helyar, S. J., de Boer, H., Monchamp, M.-E., … Deiner, K. (2021). Trade-offs between reducing complex terminology and producing accurate interpretations from environmental DNA: Comment on 'Environmental DNA: What's behind the term?' by Pawlowski et al. (2020). *Molecular Ecology*, *30*(19), 4601–4605. https://doi.org/10.1111/mec.15942

Rucińska, A., Świerszcz, S., Nobis, M., Zubek, S., Boczkowska, M., Olszak, M., Kosiński, J. G., Nowak, S., & Nowak, A. (2022). Is it possible to understand a book missing a quarter of the letters? Unveiling the belowground species richness of grasslands. *Agriculture, Ecosystems & Environment*, *324*, 107683. https://doi.org/10.1016/j.agee.2021.107683

Ruppert, K. M., Kline, R. J., & 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*, e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Sepp, S.-K., Davison, J., Moora, M., Neuenkamp, L., Oja, J., Roslin, T., Vasar, M., Öpik, M., & Zobel, M. (2021). Woody encroachment in grassland elicits complex changes in the functional structure of above- and belowground biota. *Ecosphere*, *12*(5), e03512. https://doi.org/10.1002/ecs2.3512

Soininen, E. M., Gauthier, G., Bilodeau, F., Berteaux, D., Gielly, L., Taberlet, P., Gussarova, G., Bellemain, E., Hassel, K., Stenøien, H. K., Epp, L., Schrøder-Nielsen, A., Brochmann, C., & Yoccoz, N. G. (2015). Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. *PLoS ONE*, *10*(1), e0115335. https://doi.org/10.1371/journal.pone.0115335

Soininen, E. M., Zinger, L., Gielly, L., Yoccoz, N. G., Henden, J.-A., & Ims, R. A. (2017). Not only mosses: Lemming winter diets as described by DNA metabarcoding. *Polar Biology*, *40*(10), 2097–2103. https://doi.org/10.1007/s00300-017-2114-3

Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: The principles and practices of statistics in biological research* (3rd ed., p. 880). W. H. Freeman.

Sønstebø, J. H., Gielly, L., Brysting, A. K., Elven, R., Edwards, M., Haile, J., Willerslev, E., Coissac, E., Rioux, D., Sannier, J., Taberlet, P., & Brochmann, C. (2010). Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. *Molecular Ecology Resources*, *10*(6), 1009–1018. https://doi.org/10.1111/j.1755-0998.2010.02855.x

Steinbauer, M. J., Grytnes, J.-A., Jurasinski, G., Kulonen, A., Lenoir, J., Pauli, H., Rixen, C., Winkler, M., Bardy-Durchhalter, M., Barni, E., Bjorkman, A. D., Breiner, F. T., Burg, S., Czortek, P., Dawes, M. A., Delimat, A., Dullinger, S., Erschbamer, B., Felde, V. A., … Wipf, S. (2018). Accelerated increase in plant species richness on mountain summits is linked to warming. *Nature*, *556*(7700), 231–234. https://doi.org/10.1038/s41586-018-0005-6

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

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., & Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*(3), e14. https://doi.org/10.1093/nar/gkl938

Terwayet Bayouli, I., Terwayet Bayouli, H., Dell'Oca, A., Meers, E., & Sun, J. (2021). Ecological indicators and bioindicator plant species for biomonitoring industrial pollution: Eco-based environmental assessment. *Ecological Indicators*, *125*, 107508. https://doi.org/10.1016/j.ecolind.2021.107508

Träger, S., Öpik, M., Vasar, M., & Wilson, S. D. (2019). Belowground plant parts are crucial for comprehensively estimating total plant richness in herbaceous and woody habitats. *Ecology*, *100*(2), e02575. https://doi.org/10.1002/ecy.2575

Uuemaa, E., Mander, Ü., & Marja, R. (2013). Trends in the use of landscape spatial metrics as landscape indicators: A review. *Ecological Indicators*, *28*, 100–106. https://doi.org/10.1016/j.ecolind.2012.07.018

Wang, Y., Pedersen, M. W., Alsos, I. G., De Sanctis, B., Racimo, F., Prohaska, A., Coissac, E., Owens, H. L., Merkel, M. K. F., Fernandez-Guerra, A., Rouillard, A., Lammers, Y., Alberti, A., Denoeud, F., Money, D., Ruter, A. H., McColl, H., Larsen, N. K., Cherezova, A. A., … Willerslev, E. (2021). Late quaternary dynamics of Arctic biota from ancient environmental genomics. *Nature*, *600*(7887), 86–92. https://doi.org/10.1038/s41586-021-04016-x

Wiens, J. J. (2016). Climate-related local extinctions are already widespread among plant and animal species. *PLoS Biology*, *14*(12), e2001104. https://doi.org/10.1371/journal.pbio.2001104

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergård, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L. S., Pearman, P. B., Cheddadi, R., Murray, D., Bråthen, K. A., Yoccoz, N., … Taberlet, P. (2014). Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature*, *506*(7486), 47–51. https://doi.org/10.1038/nature12921

Wilson, J. B. (2011). Cover plus: Ways of measuring plant canopies and the terms used for them. *Journal of Vegetation Science*, *22*(2), 197–206. https://doi.org/10.1111/j.1654-1103.2010.01238.x

Yoccoz, N. G. (2012). The future of environmental DNA in ecology. *Molecular Ecology*, *21*(8), 2031–2038. https://doi.org/10.1111/j.1365-294X.2012.05505.x

Yoccoz, N. G., Bråthen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., Von Stedingk, H., Brysting, A. K., Coissac, E., Pompanon, F., Sønstebø, J. H., Miquel, C., Valentini, A., De Bello, F., Chave, J., Thuiller, W., Wincker, P., Cruaud, C., Gavory, F., … Taberlet, P. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, *21*(15), 3647–3655. https://doi.org/10.1111/j.1365-294X.2012.05545.x

Zinger, L., Taberlet, P., Schimann, H., Bonin, A., Boyer, F., De Barba, M., Gaucher, P., Gielly, L., Giguet-Covex, C., Iribar, A., Réjou-Méchain, M., Rayé, G., Rioux, D., Schilling, V., Tymen, B., Viers, J., Zouiten, C., Thuiller, W., Coissac, E., & Chave, J. (2019). Body size determines soil community assembly in a tropical forest. *Molecular Ecology*, *28*(3), 528–543. https://doi.org/10.1111/mec.14919

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: Ariza, M., Fouks, B., Mauvisseau, Q., Halvorsen, R., Alsos, I. G. & de Boer, H. J. (2023). Plant biodiversity assessment through soil eDNA reflects temporal and local diversity. *Methods in Ecology and Evolution*, *14*, 415–430. https://doi.org/10.1111/2041-210X.13865

Supplementary Figures Paper III

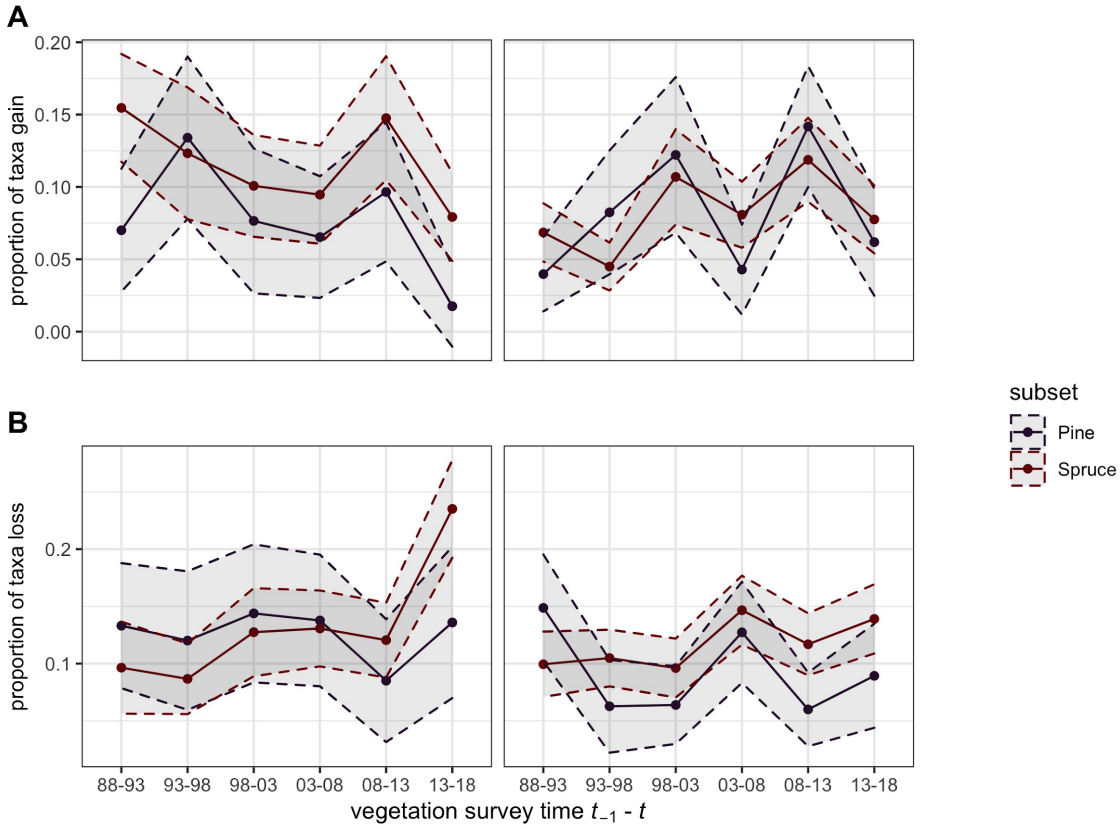

Figure S1 Compositional turnover over five-year periods based upon the vegetation surveys, expressed as averages of plot values for proportions of taxa gain (A) and loss (B) calculated separately for pine and spruce subsets. Confidence intervals (95%) are indicated by grey shaded areas delimited by dotted lines in the respective colours.

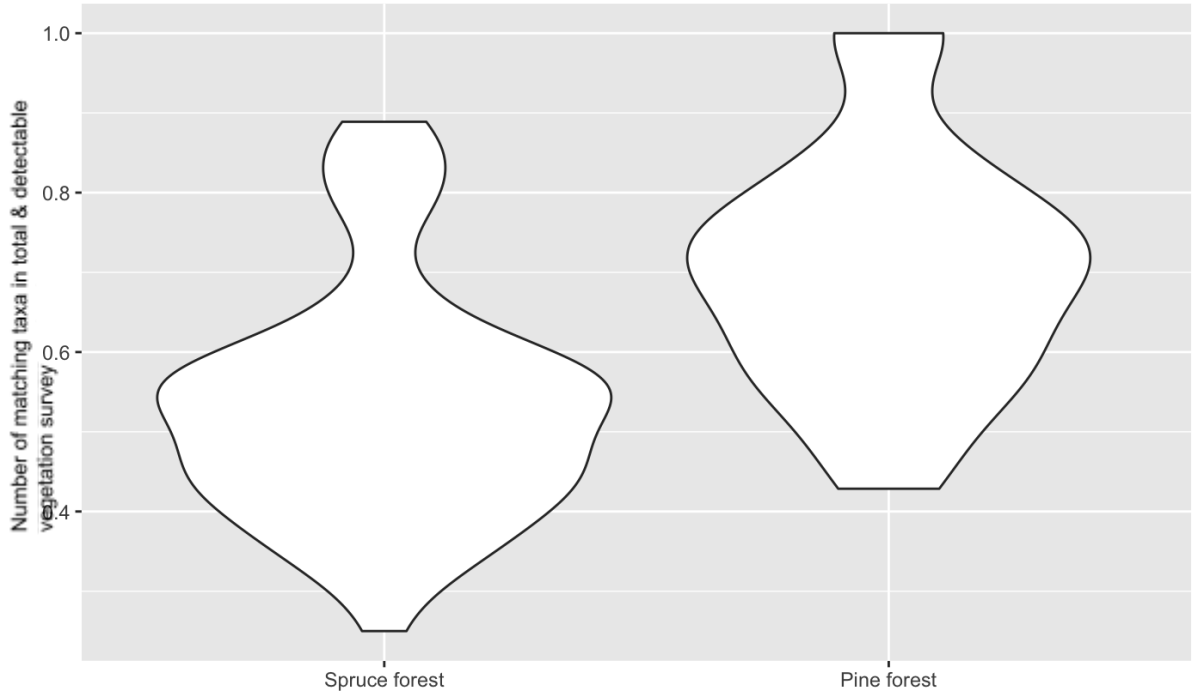

Figure S2 Number of matching taxa in total & detectable vegetation survey in plots of the spruce or pine subset. From the total plot composition registered across years of vegetation survey, the proportion of number of matching taxa in total & detectable vegetation survey (2018) is summarised for plots located in forest dominated by spruce or pine. Violin charts display the distribution of the data, where wider shapes indicate data abundance and narrow shapes vice versa. Proportions were compared between plots with a Wilcoxon rank sum test; with a resulting p-value>0.05.

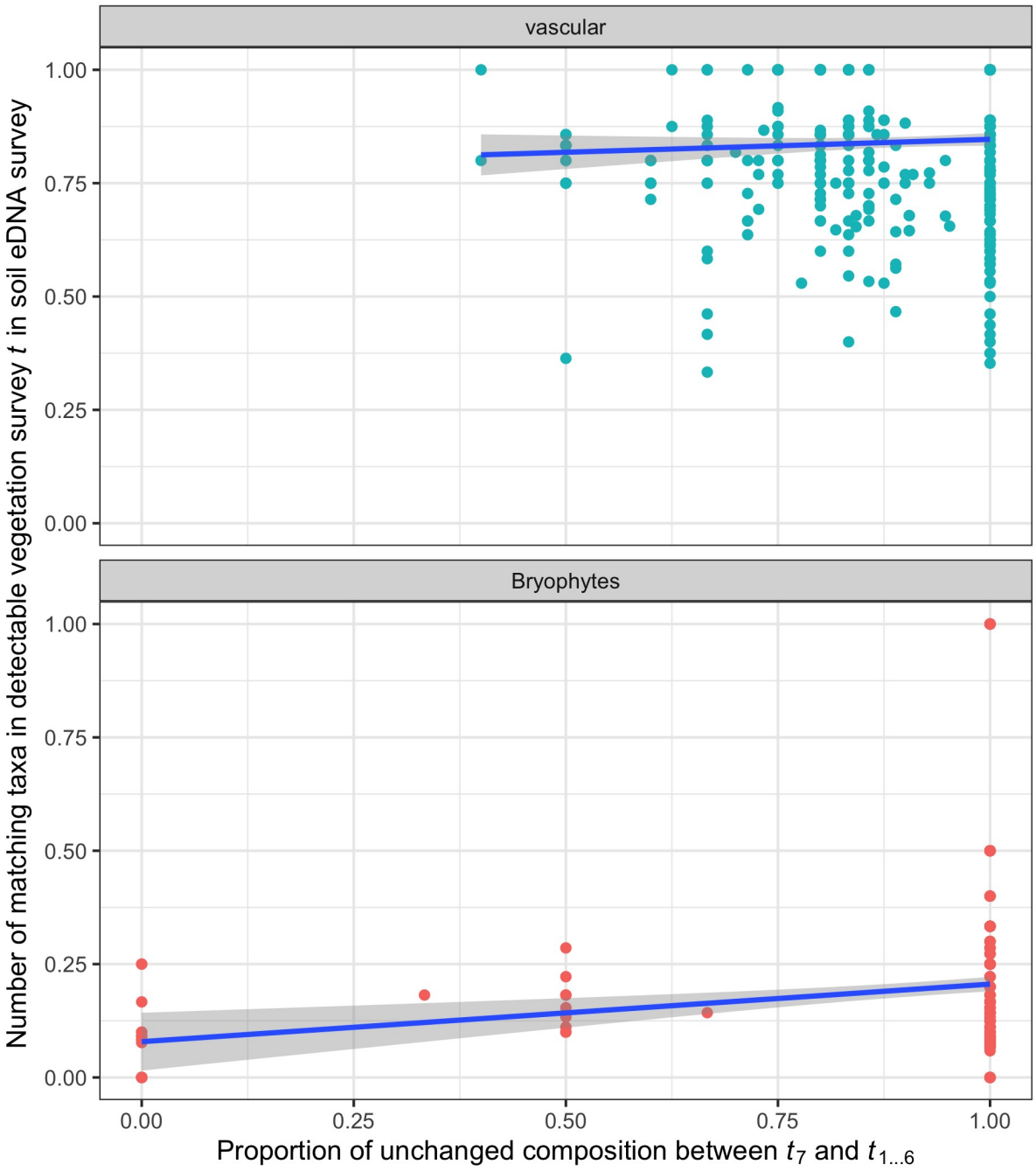

Figure S3 Number of matching taxa in each detectable vegetation survey and the relation to the proportion of unchanged composition out of the *taxon count* t_{bdvs}+ *t*. The proportion of unchanged taxa between vegetation survey t_{bdvs} and t_1, ...,t_6 is calculated as the fraction of identical taxa between t_{bdvs} and t_1, ...,t_6 out of the sum of number of taxa in t_{bdvs} and t_1, ...,t_6.Thus, in each plot six iterations are

calculated for this variable. Each one is correlated to the number of matching taxa in each detectable vegetation survey t_1, ...,t_6, accounting for 600 points of data pairing in total. Correlations are shown by a blue line and tested with a Spearman test (vascular R: -0.041, p value> 0.05; bryophytes R: 0.22, p value<0.05) whereas shaded colour indicates 95% confidence for predictions from a linear model.

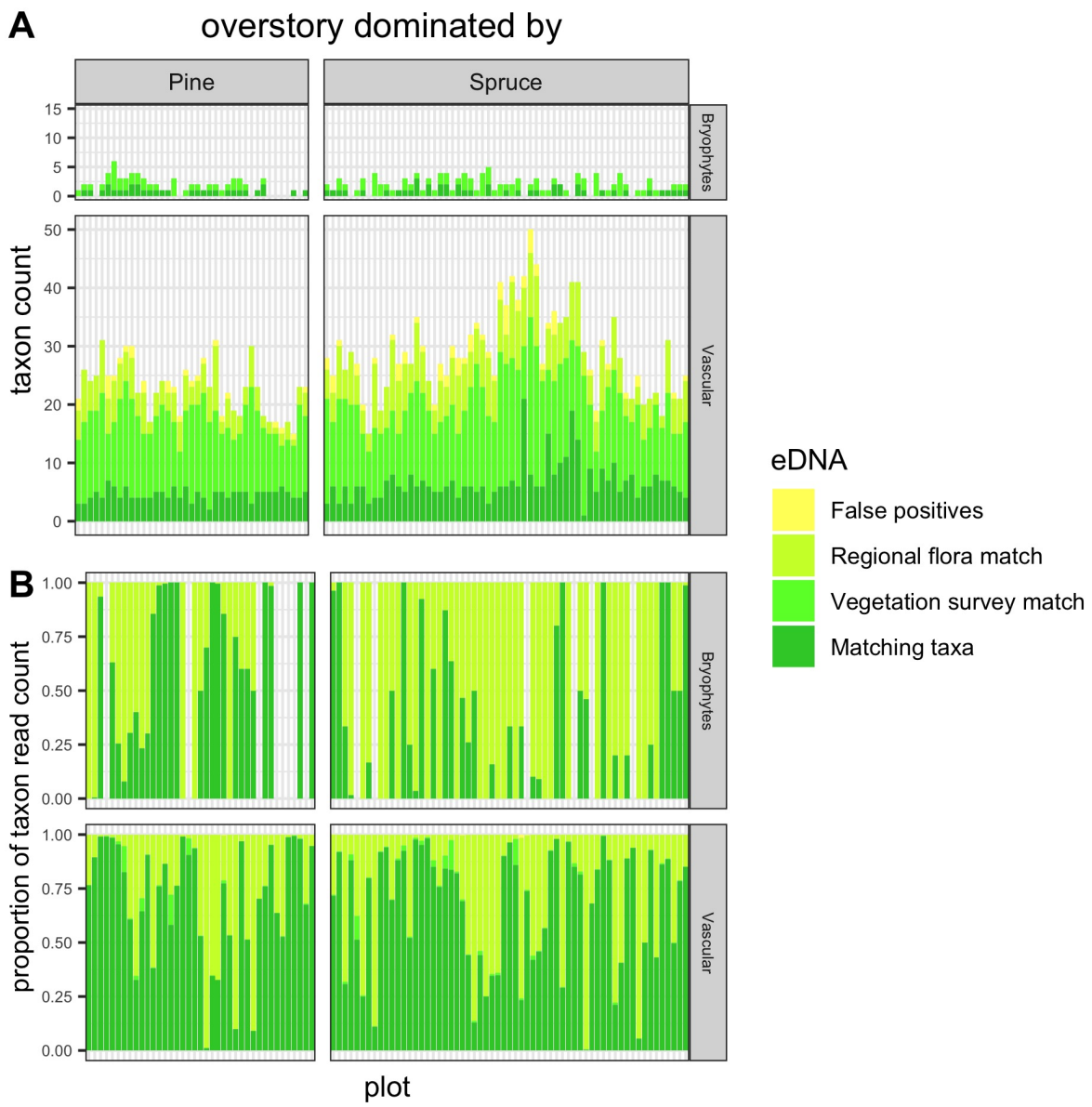

Figure S4 Spatial resolution of the soil eDNA survey, assessed by categorising plant taxa detected with a single soil eDNA sample from each plot into: matching taxa (taxa registered in the plot in the best detected vegetation survey t_7=2018); vegetation survey match (taxa recorded in any other plot in the best detected vegetation survey t_7=2018); regional flora match (taxa recorded in the Solhomfjell area outside plots); and false positives (taxa not recorded in the Solhomfjell area or with higher mean read count in PCR negatives than across

samples and replicates). Results are shown for (A) taxon counts and (B) proportions of taxon read count from each category. For bryophytes, no taxon categorised as false positives or regional flora match were found.

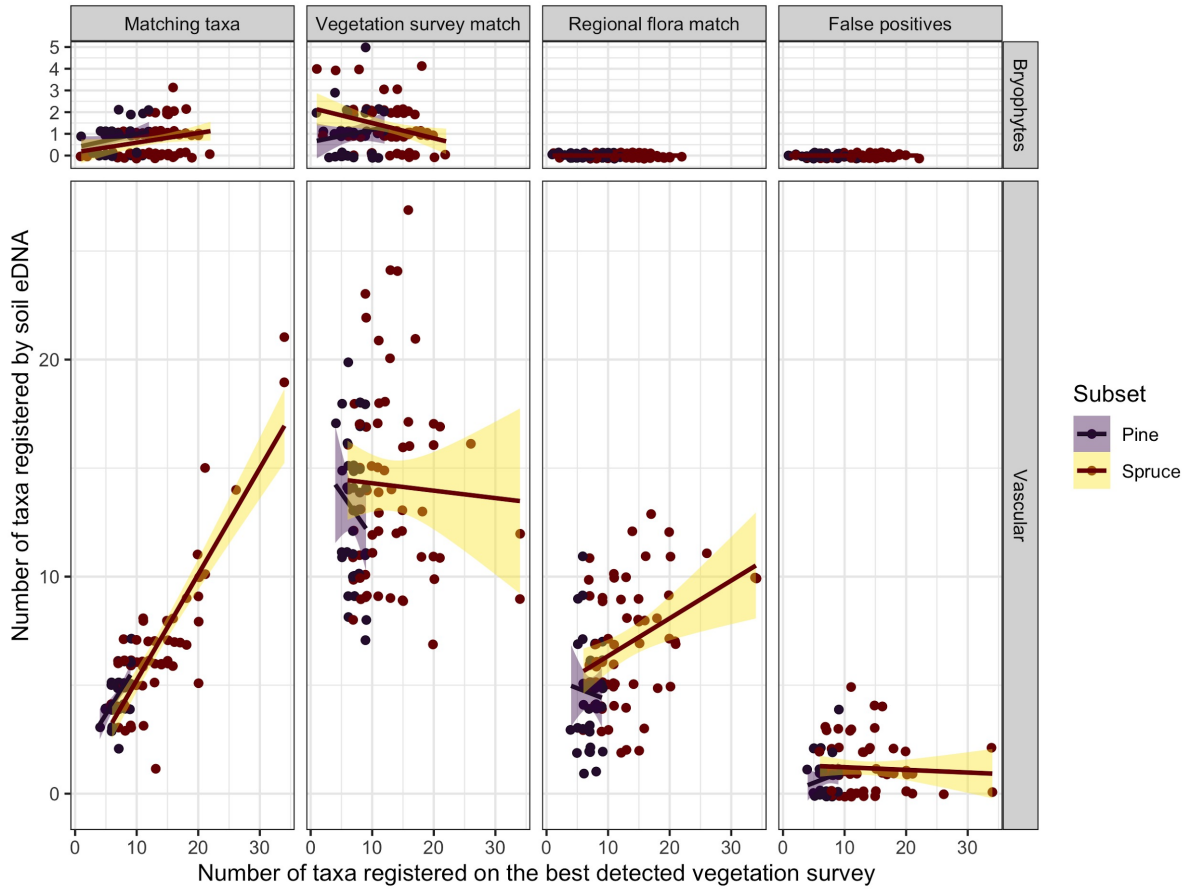

Figure S5 Soil eDNA taxa categories and the relation to the best detected vegetation survey. For each plot (points) from pine or spruce subsets, counts of taxa categories registered by soil eDNA assessment are shown against the number of taxa registered on the best detected vegetation survey (2018). Point positions are jittered 0.15 in both axis to avoid overlap. Correlations between inventories are fitted to a linear model and indicated by a line, with shaded colours that indicate 95% confidence level interval for predictions from a linear model.

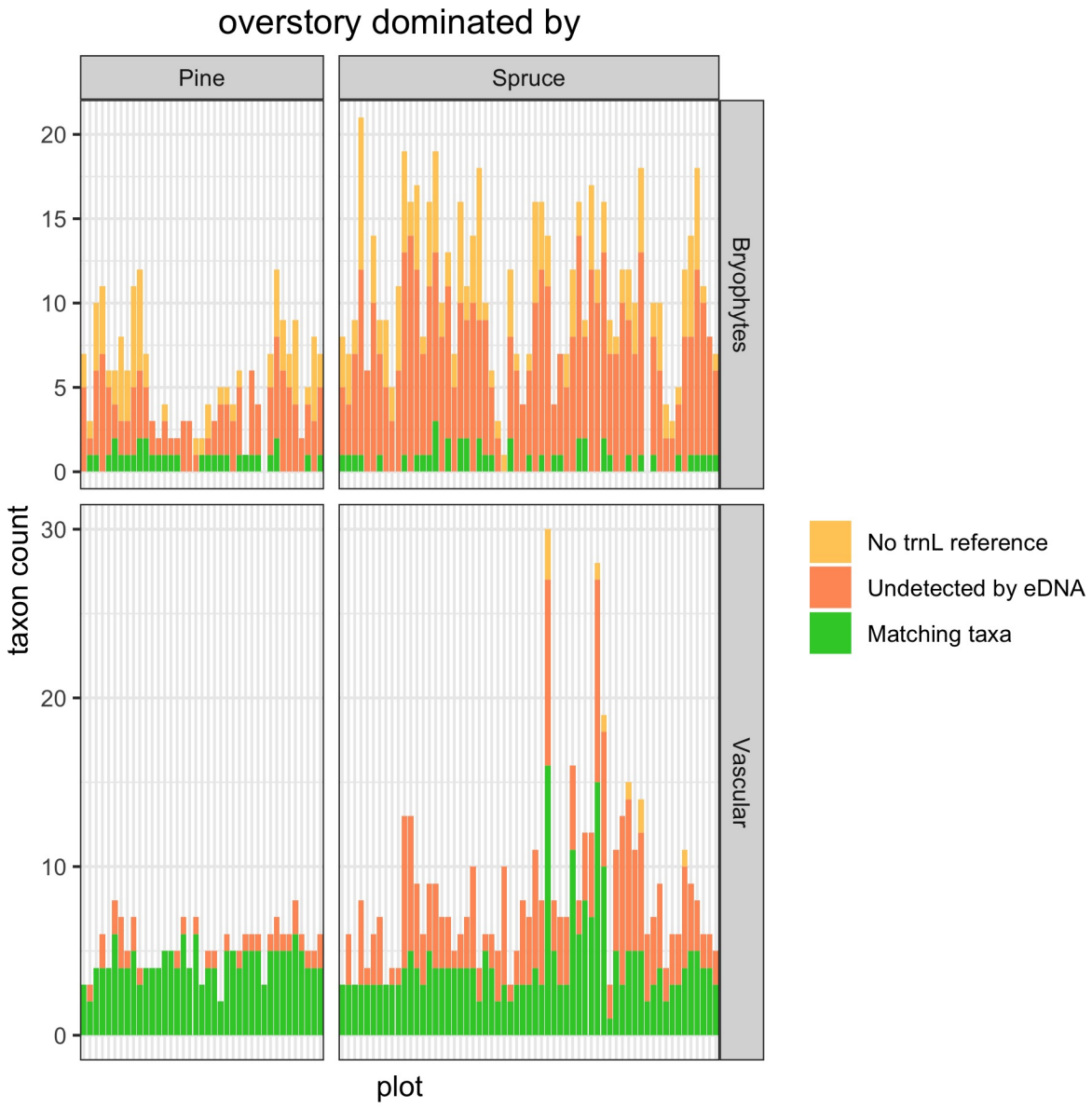

Figure S6 Spatial resolution of the soil eDNA survey and best detected vegetation survey. Plant taxa recorded in each plot at the best detected vegetation survey t_7=2018 are categorised into: matching taxa (taxa registered in the plot also in the soil eDNA survey); undetected by eDNA (trnL (UUA) reference sequence available but taxa were not registered in the soil eDNA survey); and no trnL reference (no reference sequence was available for identification with eDNA data). Taxon counts are shown separately for

combinations of vascular plants and bryophytes, and spruce and pine forest subsets.

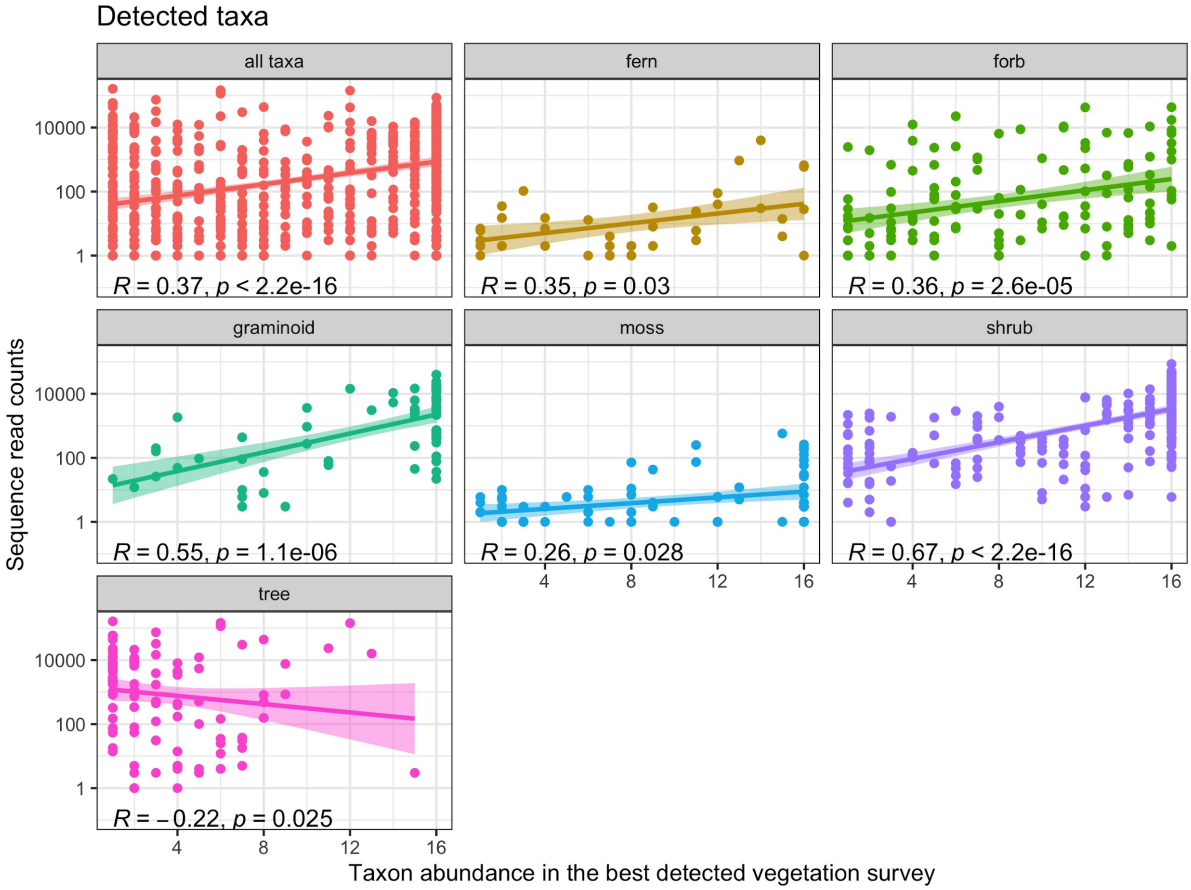

Figure S7 Concordance between read counts and taxa abundance registered in the best detected vegetation survey (2018). Sequence read counts from matching taxa registered in the best detected vegetation survey within all growth forms are log 10 transformed. Subplot frequency (proxy for biomass) corresponds to taxon presence in 625 cm^2 grids. Spearman correlations between both variables are displayed by a line and shaded colour indicate 95% confidence level interval for predictions from a linear model with p values.

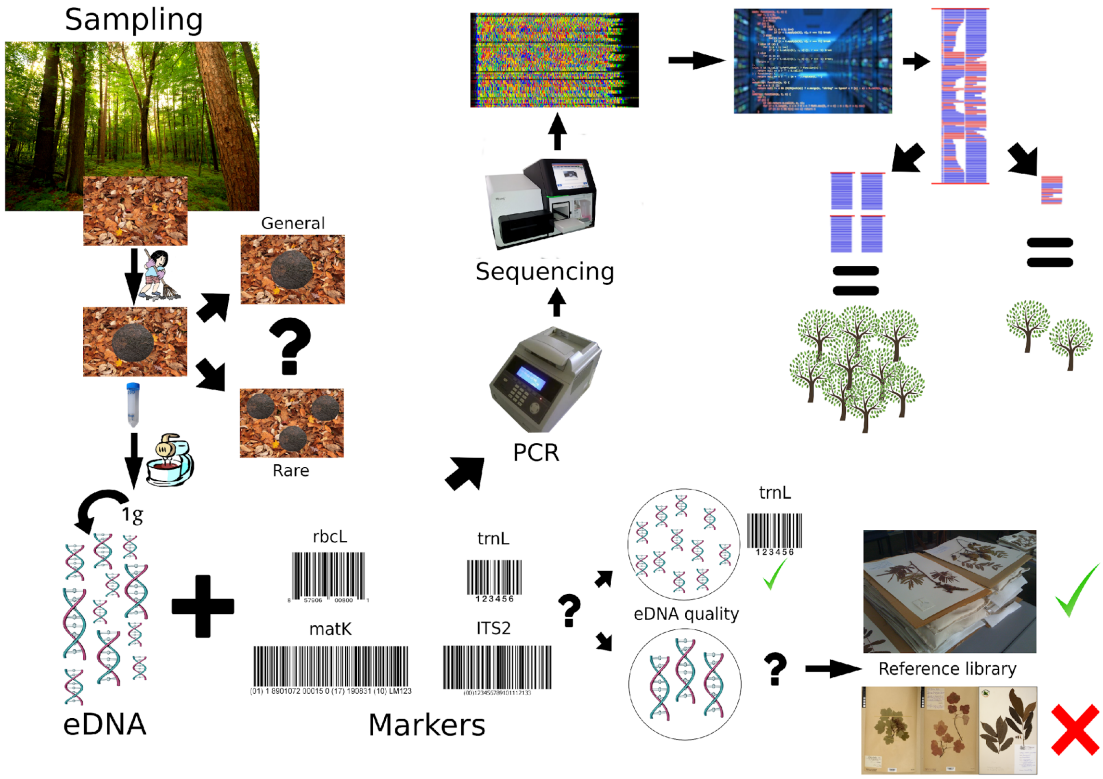

Figure 8 Key methodological steps in a soil eDNA plant assessment. Area for soil eDNA sampling must be cleared out from debris and litter, and number of eDNA subsamples will correspond to the aims of the assessment. We recommend hammering a falcon tube of 50 mL onto the soil to avoid major disruptions of plots and underground ecosystems. After homogenization, soil eDNA (sub)samples can be aliquoted up to 1 gram and we recommend those DNA extractions protocols that maximise the division between the organic phase and aqueous phase. When choosing a marker, one must consider both the DNA quality in the sample and if the flora to be assessed is sufficiently catalogued in sequence reference libraries. We recommend the chloroplast marker trnL (UUA) intron P6 loop when working with degraded samples and particularly in northern areas. Multiple PCR replicates are recommended, and equimolar

normalisation of samples should be carried before sequencing. We recommend a strict threshold of sequence match to reference libraries when working with short markers. Finally, sequence abundance may be used as a proxy for relative abundance.

Supplementary Tables Paper III

Table 1 Summary of taxa detected across all vegetation surveys

Taxa	Overstory dominated by	Number of plots	Taxa count* (plot mean)	Taxa count (plot standard deviation)
Vascular	Pine	39	7	1,29
Vascular	Spruce	61	13,31	6,35
Non-vascular	Pine	39	10,15	4,66
Non-vascular	Spruce	61	19,55	6,83

***taxa counts comprise those that are detectable with the trnL (UUA) intron *g h* primers

Table 2 Summary of taxa detected in each vegetation survey

Taxa	Survey year	Overstory dominated by	Number of plots	Taxa count* (plot mean)	Taxa count (plot standard deviation)
vascular	1	pine	39	5,54	1,39
vascular	2	pine	39	4,92	1,46
vascular	3	pine	39	5,03	1,46
vascular	4	pine	39	5,36	1,33
vascular	5	pine	39	4,87	1,30
vascular	6	pine	39	5,38	1,44
vascular	7	pine	39	5,21	1,44
vascular	1	spruce	61	10,59	5,45
vascular	2	spruce	61	10,21	5,36
vascular	3	spruce	61	9,54	5,16
vascular	4	spruce	61	9,66	5,21
vascular	5	spruce	61	9,00	4,98
vascular	6	spruce	61	8,89	4,90
vascular	7	spruce	61	8,36	4,79
non-vascular	1	pine	39	5,10	2,26
non-vascular	1	spruce	58	8,83	2,85
non-vascular	2	pine	39	4,79	2,07
non-vascular	2	spruce	58	9,71	3,50
non-vascular	3	pine	39	4,97	2,40
non-vascular	3	spruce	60	9,67	3,62

non-vascular	4	pine	39	4,74	2,11
non-vascular	4	spruce	59	9,46	3,47
non-vascular	5	pine	38	4,45	1,78
non-vascular	5	spruce	60	9,02	3,11
non-vascular	6	pine	38	4,42	1,95
non-vascular	6	spruce	61	9,11	3,55
non-vascular	7	pine	38	3,82	1,71
non-vascular	7	spruce	59	8,10	3,17

*****taxa counts comprise those that are detectable with the trnL (UUA) intron *g h* primers

Paper IV

Contributions not included in this dissertation

Chua, P., Polling, M., Lynggaard, C., **Ariza, M.**, Bohmann, K (2022) Chapter 11. Amplicon metabarcoding. In: de Boer H, Rydmark MO, Verstraete B, Gravendeel B (Eds) *Molecular identification of plants: from sequence to species.* Advanced Books. https://doi.org/10.3897/ab.e98875

Polling, M., ter Schure, A.T.M., van Geel, B., van Bokhoven, T., Boessenkool, S., MacKay, G., Langeveld, B.W., **Ariza, M.**, van der Plicht, H., Protopopov, A.V., Tikhonov, A., de Boer, H., Gravendeel, B., 2021. Multiproxy analysis of permafrost preserved faeces provides an unprecedented insight into the diets and habitats of extinct and extant megafauna. Quat. Sci. Rev. 267, 107084. doi:10.1016/j.quascirev.2021.107084

Quintana, I., Cifuentes, E.F., Dunnink, J.A., **Ariza, M.**, Martínez-Medina, D., Fantacini, F.M., Shrestha, B.R., Richard, F.-J., 2022. Severe conservation risks of roads on apex predators. Sci. Rep[14]. 12, 2902. doi:10.1038/s41598-022-05294-9

[14] One of the Top 100 downloaded Earth, Environment and Ecology Scientific Reports papers published in 2022