FORUM

Using DNA metabarcoding to assess New Zealand's terrestrial biodiversity

Robert J. Holdaway*[1], Jamie R. Wood[1], Ian A. Dickie[2], Kate H. Orwin[1], Peter J. Bellingham[1], Sarah J. Richardson[1], Phil O'B. Lyver[1], Puke Timoti[3] and Thomas R. Buckley[4,5]

[1]Landcare Research PO Box 69040, Lincoln 7640, New Zealand
[2]Bio-Protection Research Centre, Lincoln University, PO Box 85084, Lincoln 7647, New Zealand
[3]Tūhoe Tuawhenua Trust, Private Bag 3001, Ruatāhuna, New Zealand
[4]Landcare Research, Private Bag 92170, Auckland 1142, New Zealand
[5]School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
*Author for correspondence (Email: holdawayr@landcareresearch.co.nz)

Abstract: High throughput DNA sequencing technology has enabled entire biological communities to be characterised from DNA derived from pools of organisms, such as bulk-collected invertebrates, or DNA extracted from environmental samples (e.g. soil). These DNA-based techniques have the potential to revolutionise biodiversity monitoring. One approach in particular, DNA metabarcoding, can provide unprecedented taxonomic breadth at a scale not practically achievable through the morphological identification of individual organisms. Here, we assess the current strengths and weaknesses of DNA metabarcoding techniques for biodiversity assessment. We argue that it is essential to integrate conventional monitoring methods with novel DNA methods, to validate methods, and to better use and interpret data. We present a conceptual framework for how this might be done, explore potential applications within national biodiversity assessment frameworks, Maori biodiversity monitoring and the primary sector, and highlight areas of current uncertainty and future research directions. Rapid developments in DNA sequencing technology and bioinformatics will make DNA-based community data increasingly accessible to ecologists, and there needs to be a corresponding shift in research focus from DNA metabarcoding method development and evaluation to real-world applications that provide rich information for a range of purposes, including conservation planning and land management decisions.

Keywords: biodiversity monitoring; bioinformatics; biosecurity; cultural indicators; ecosystem function; environmental DNA; national framework; species detection; te ao Maori; uncertainty

Introduction

The measurement of biodiversity (i.e. species presence, abundance and change with time) underpins most ecological research, conservation planning and environmental policy (Allen et al. 2003; MacLeod et al. 2012; Pereira et al. 2013; Environmental Reporting Act 2015). Biodiversity data come in various forms, including quantitative records (e.g. defined plots or remote sensing) and qualitative records based on personal experience or traditional ecological knowledge. This reflects the diversity of individuals, groups and organisations who are interested in measuring biodiversity. However, conventional biodiversity assessment techniques are typically constrained by our ability to see and identify biodiversity features of interest and track them over time, which has resulted in a strong taxonomic bias towards large visible organisms such as plants and birds. A side-effect of this bias is that conventional assessments generally exclude many of the taxonomic groups that contain the greatest diversity (e.g. fungi and bacteria), many species of agricultural pests and diseases (e.g. invertebrates, oomycetes and other pathogens), human parasites and diseases (e.g. fungal pathogens, helminth worms), sources of food and medicines, culturally important indicator species (e.g. butterflies), and many threatened species of conservation concern (Hitchmough 2013).

The bias inherent in conventional assessment techniques can now potentially be overcome by using DNA barcoding and metabarcoding to sample entire biological communities. DNA barcoding identifies a species based on its DNA sequence at a specific gene region (Hebert et al. 2003); this information is used to build sequence databases. DNA metabarcoding refers to the process of obtaining gene sequences from these same regions from multiple species at once (e.g. from soil or captured invertebrate samples), and matching these sequences to those in a database to obtain community level information (Figs. 1 & 2). While there are multiple molecular techniques available for analysing environmental samples, metabarcoding is currently the most widely used approach in ecological studies (see Table 1 for New Zealand examples) because it allows specific taxonomic groups to be targeted and identified (e.g. invertebrates, fungi, prokaryotes) within a sample. The potential for DNA metabarcoding techniques to revolutionise biodiversity assessment is widely recognised (e.g. Ficetola et al. 2008; Yu et al. 2012; Bohmann et al. 2014; Gibson et al. 2015; Thomsen & Willerslev 2015; Barnes & Turner 2016; Creer et al. 2016). However, most current applications of DNA metabarcoding have been done independently of conventional monitoring. Integration of DNA metabarcoding with conventional monitoring techniques is essential to validate novel DNA methods, build on existing knowledge, and to
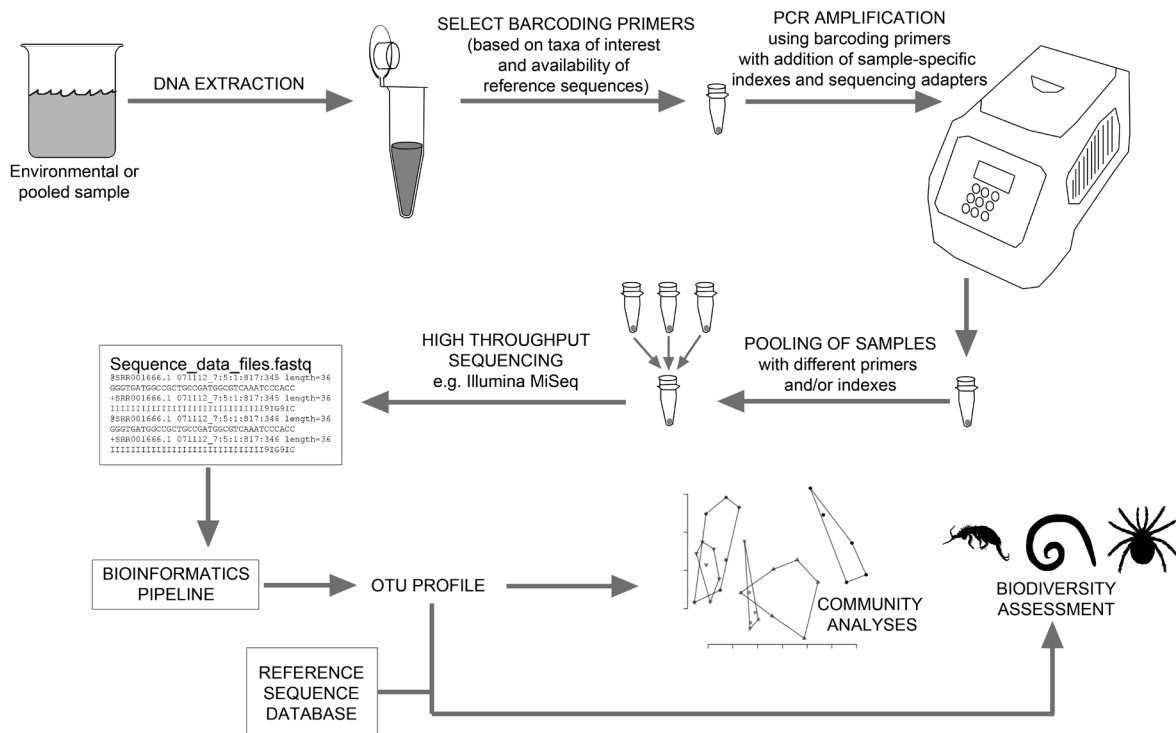
Figure 1. The process of DNA metabarcoding involves extracting community DNA from pools of organisms, such as invertebrates captured in nets or pitfall traps, or DNA from environmental samples (e.g. soil). Standard molecular techniques (e.g. PCR) are used to isolate and amplify the gene region of interest. This amplified community DNA sample is then sequenced using high throughput DNA sequencing technologies. The resulting sequences are clustered into closely related groups known as operational taxonomic units (OTUs). These OTUs are then matched to reference databases of DNA barcodes (e.g. GenBank) to taxonomically identify all the species present.
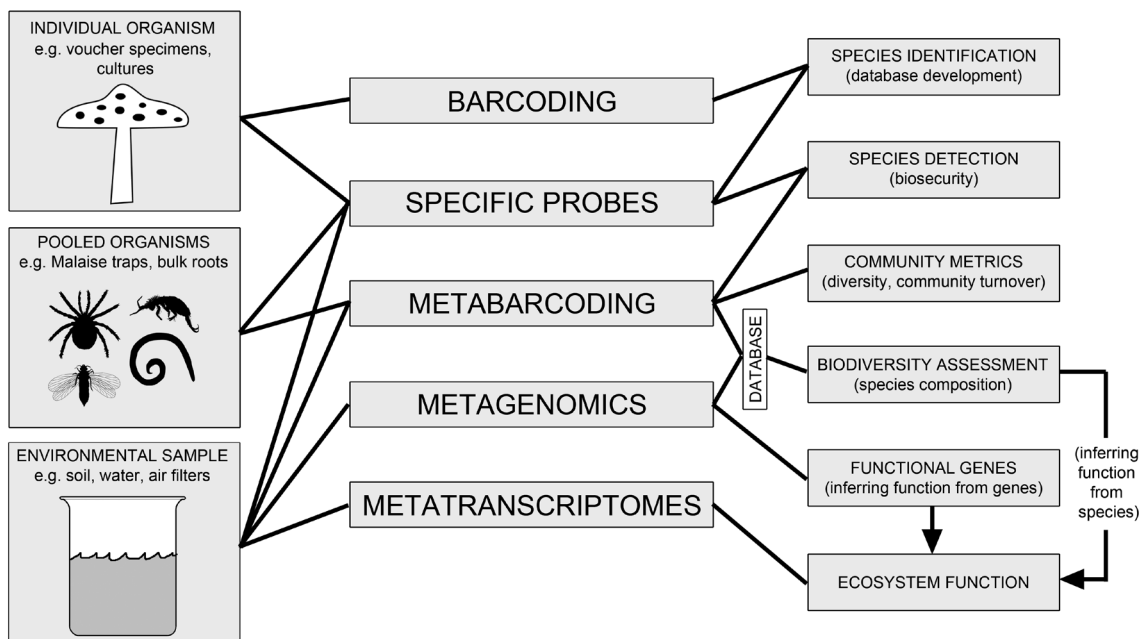

Figure 2. Illustration of the major ways in which DNA is being used in environmental applications from individual species to ecosystems, putting DNA metabarcoding in the context of other molecular techniques. Depending on the type of sample and the analysis method employed, uses range from identification or detection of individual species up to measurements of ecosystem function. Most of this review focuses on the analysis of pooled organisms and environmental samples using metabarcoding techniques to measure species composition of communities. Ecosystem function can be inferred from the identity of species present but is not directly measured by metabarcoding techniques. Metagenomics and metatranscriptomics measure function directly by targeting functional genes and their expression in the environment. They currently have limited use for biodiversity assessment (especially eukaryotes) and have been most widely applied in studies of prokaryotic communities.

Table 1. Published New Zealand examples illustrating the use of DNA metabarcoding for terrestrial biodiversity assessment.

Application/study description	Sample type	Taxa targeted	Reference
Distribution and diversity of Verrucomicrobia methanotrophs in geothermal and acidic environments	Geothermal soils	Archaea, bacteria	Sharp et al. 2014
Microbial community effects on N_2O emissions	Pasture soils	Archaea, bacteria	Morales et al. 2015
Biodiversity assessment using a multigene approach	Native forest soils	Archaea, bacteria, fungi, eukaryotes	Drummond et al. 2015
Assessing the diet of endangered land snails to facilitate ecological restoration and relocation programme	Snail faeces	Earthworms	Boyer et al. 2013; Waterhouse et al. 2014
Detecting shifts in soil microbial communities following alteration of native tussock grassland ecosystems	Tussock grassland soil	Fungi and bacteria	Wakelin et al. 2013
Assessment of microbial communities in vineyards	Vineyards (soil, bark and fruit)	Fungi	Morrison-Whittle & Goddard 2015
Assessing changes in mycorrhizal fungal community composition during ecosystem development	Roots from native forest	Fungi	Martínez-García et al. 2015

better use and interpret the data generated. Here we describe the general benefits of integrating the two approaches, discuss their application within the context of national biodiversity assessment, Māori biodiversity monitoring and the primary sector, and highlight areas of current uncertainty and future research directions.

General benefits of DNA metabarcoding and its integration with conventional biodiversity assessment

There are a number of general benefits of metabarcoding compared with conventional monitoring techniques. DNA metabarcoding can detect a broad range of taxa from a single environmental sample, removing the need to apply a diverse set of sampling methods in the field. For example, a single soil-based sampling regime could be used to identify locally present above- and below-ground species at multiple trophic levels (e.g. Drummond et al. 2015). Highly trained specialists are not needed for specimen identification, avoiding this typically labour-intensive process (e.g. Yu et al. 2012; Yang et al. 2014). Molecular analyses and bioinformatics, while challenging, can be partially or totally automated to process large datasets rapidly. Further, techniques and expertise are generally transferable across taxa, unlike taxonomic expertise. Cryptic species – those that are hard to identify morphologically and often overlooked in conventional analyses – are readily detected. DNA metabarcoding also allows the incorporation of previously ignored biota such as bacteria and fungi into biodiversity assessments. These biota contribute a large proportion of biodiversity (e.g. fungi encompass c. 23% of New Zealand's described eukaryotic diversity; Gordon 2012) and are critical for many ecosystem functions and services (de Vries et al. 2013; Wagg et al. 2014; Soliveres et al. 2016). Finally, environmental DNA (eDNA) can detect low-density (locally rare) and elusive species (Jerde et al. 2011; Calvignac-Spencer et al. 2013; Doyle et al. 2015). For example, DNA extracted from faeces has been used to model the distribution of invasive red fox (*Vulpes vulpes*) in Tasmania (MacDonald & Sarre 2015).

Given the benefits of metabarcoding techniques, they could be viewed as a replacement for conventional biodiversity monitoring. However, we argue that the two approaches are not interchangeable, and uniting them will result in significant improvements to both fields of research by drawing on their respective strengths (Table 2). For example, conventional approaches are particularly strong in their use of standardised sampling techniques and the analysis, interpretation and presentation of biodiversity datasets (e.g. Allen et al. 2003). Metabarcoding approaches could benefit from drawing on these strengths, facilitating the ecological interpretation of DNA data sets. In addition, comparative analyses applying both DNA and conventional methods to the same taxonomic groups will help test and improve the accuracy and reliability of DNA-based methods. This has begun for some groups such as invertebrates (Watts et al. 2016) and plants (Fahner et al. 2016), but such analyses must be carried out in the context of our ability to detect biologically meaningful patterns. Application of DNA metabarcoding alongside conventional methods of assessment, coupled with a quantitative understanding of the detection probabilities and biases of different methods (e.g. Furlan et al. 2016), could lead to greater overall detection probabilities for target taxa. Further, conventional monitoring techniques provide additional information that cannot be obtained from metabarcoding. These include demography (e.g. tagged individuals in plant population ecology), sex ratios, phenology, and the colours and conditions of individuals. Nevertheless, environmental metabarcoding does have a role to play in assessing taxa where existing methods are non-existent, poor, or too expensive. Overall, there are multiple clear benefits to be obtained through the integration of DNA metabarcoding with conventional biodiversity assessment techniques. These general benefits could be realised in a range of contexts, and we now explore three of these in more detail.

Case study 1: National Biodiversity Assessment Framework

Multiple agencies within New Zealand are currently working towards implementing a national framework for biodiversity

Table 2. Comparative assessment of the strengths and weaknesses of DNA metabarcoding and conventional biodiversity assessment of terrestrial biodiversity in a New Zealand context.

Aspect		DNA metabarcoding	Conventional assessment
SAMPLING	Field methods	Often poorly defined, inconsistent (-); standard methods could potentially be easily applied across ecosystem types without in-field taxonomic expertise (+)	Generally standardised and well documented (+); optimised for forest ecosystems (+/-); dependent on in-field taxonomic expertise (-).
	Sampling	Single, easily collected sample (+); often poor sample design (-)	Extensive understanding of optimal sample design and use of consistent methods (+)
DATA QUALITY	Quantitative	Limited by biases in DNA extraction, PCR, analysis; empirical evidence suggests semi-quantitative (-)	Counting of individuals or percent cover straightforward (+)
	Taxonomic coverage	High (+)	Low, due to difficulty in identifying some taxonomic groups based on morphology and expense limitations (-)
	Temporal depth	Variable: has potential to provide time-integrated information depending on DNA decay rates (+)	Shallow: often based on a single point in time and may miss transient or ephemeral species (-)
	False positives	High risk of false positives due to incomplete databases, poor taxonomic resolution, sample contamination (-)	Risk of false positives low, depending on quality of expert knowledge (+)
	False negatives	Variable depending on choice of method and sequencing depth (+/-)	Risk of missing species high (-)
UTILITY	Analysis	Difficult: novel methods and indicators need to be developed; results often difficult for end-users to interpret (-)	Easy: multiple existing methods and indicators available (+)
	Relevance of data to primary industry	High: direct links to productivity and ecosystem function (+)	Low–medium: taxa monitored have less direct relevance to industry (-), but biodiversity information can be useful for marketing (+)
	Ability to interpret data in phylogenetic context	High (+)	Medium: limited by taxonomic coverage (-)
	Ability to interpret data in ecological context	Low: ecological function of many cryptic taxa unknown (-)	High (+)
	Ability to interpret data in context of ecological integrity	Med–high for species occupancy and ecosystem representation (+); bio-status hard to determine for some taxa (-)	Generally high (+), although species occupancy and ecosystem representation limited by taxonomic coverage (-)
COSTS	Travel time	Same base cost, but smaller field crew may be sufficient (+)	Same base cost, but requirement for more different experts may increase field crew size (-)
	Time to collect in the field	Variable: generally low for common samples (e.g. soil) but can be higher for targeted sampling (e.g. pooled invertebrate samples) (+)	Medium–high: depends on required taxonomic resolution; can be inefficient and weather dependent (e.g. bird counts); mammal and bird counts require multiple days (-)
	Time to process samples	High, particularly if lab methods need to be optimised (-)	Low (+), although voucher specimen identification is often very time consuming (+/-)
	Cost of species identifications	Based on molecular methods, costs going down (+)	Based on expert knowledge, labour costs going up (-)

assessment (Box 1), based on the concept of ecological integrity (Lee et al. 2005; Allen et al. 2009; Environmental Reporting Act 2015). Ecological integrity is defined as the 'full potential of indigenous biotic and abiotic features and natural processes, functioning in sustainable communities, habitats and landscapes' (Environmental Reporting Act 2015) and has three measurable components: (i) indigenous dominance, (ii) species occupancy and (iii) ecosystem representation. Indigenous dominance refers to the level of indigenous influence on the composition, structure and functioning of a community; species occupancy is the extent to which any species capable of living in a particular ecosystem is actually present; and ecosystem representation is the extent to which a full range of abiotic environments are protected and maintained (Lee et al. 2005). Current field-based biodiversity monitoring within this framework is limited to plants, vertebrates and some large-bodied invertebrates. The integration of DNA metabarcoding approaches into this framework would greatly broaden the taxonomic coverage, improving the assessment of species occupancy and ecosystem representation.

The greater taxonomic coverage provided by metabar-coding will also allow a more comprehensive assessment of ecosystem functions and services tied to community structure than conventional monitoring alone, through the inclusion of taxa that are critical to ecosystem function (e.g. fungi). Therefore, integration of conventional monitoring frameworks with metabarcoding approaches will increase the ability to report on maintenance of ecosystem processes (which is included within the definition of ecological integrity) and will strengthen the relevance of biodiversity data to the primary sector. For example, food web structure and community composition, which can be quantified by DNA metabarcoding, are linked to the stability of communities and the functions they mediate (Rooney & McCann 2012; Oliver et al. 2015), and could provide information on the success of management techniques designed to provide suitable habitats for natural predators of agricultural and horticultural pest species (Haaland et al. 2011). In addition, higher levels of multiple ecosystem functions have been linked to rare species (Soliveres et al. 2016) and diversity across multiple trophic levels (Wagg et al. 2014). Integrating DNA metabarcoding with a national monitoring framework would provide a rich information base to assess relationships among biodiversity, ecosystem function and ecosystem services (Fischer et al. 2010).

Moreover, the benefits of integration go both ways. Conventional biodiversity monitoring approaches within the national framework have invested heavily in optimising sampling design, reporting metrics and uncertainty modelling to provide greater confidence in assessments of biodiversity change over time (e.g. Pereira et al. 2013; Holdaway et al. 2014; Peltzer et al. 2014; Box 1). Applying this same approach to DNA metabarcoding is essential to enable confidence in the use of DNA data for national reporting.

Case study 2: Māori biodiversity monitoring

The alignment of DNA metabarcoding techniques with Māori biodiversity monitoring frameworks offers a cross-cultural approach to monitoring New Zealand's environmental health. Both the metabarcoding and te ao Māori (Māori world) approaches contribute to an understanding of the constituents of New Zealand's ecological 'whakapapa', as well as the

Box 1. Key elements of a proposed national framework for the assessment of biodiversity in New Zealand.

A diverse range of terrestrial biodiversity assessment methods are currently used in New Zealand. Most methods focus on birds, plants and mammals, because they are large, relatively easy to identify, and have strong public/economic interests. While much of the monitoring is disjointed, considerable ongoing efforts aim to improve current monitoring practices and combine them under a national framework (Wiser et al. 2001; Allen et al. 2003; Lee et al. 2005; Bellingham et al. 2016). This framework is an aspirational goal built upon decades of research into methods, sampling, analysis and data management at multiple scales. Key components of the national framework are as follows.

Standard plot/point methods. Well-established methods with significant prior data and known uncertainty are favoured. These include 20 × 20 m vegetation plots, 5-minute bird counts, and traplines/pellet counts. Standard methods allow comparability and scaling across studies.

Multi-tier sampling frameworks. Integrated methods across tiers (e.g. local, regional, national) allow for scalability, ensuring that local and regional data can inform national reporting, while national data provide context for local studies.

National coverage. Uses unbiased sampling of a wide range of taxa across the entire New Zealand landscape, not just public conservation land. Adoption of the DOC sampling framework by regional councils (Bellingham et al. 2016), who also monitor biodiversity within a variety of productive and urban landscapes, is a step towards achieving this.

Information transfer across scales. Coordinated sampling, data quality standards, data repositories (e.g. the New Zealand National Vegetation Survey Databank), and data standardisation promote data access and use at the local, regional, and national scales.

Analysis and reporting frameworks. The concept of ecological integrity can be used to provide a similar set of metrics and indicators for regional and national reporting (Lee et al. 2005; Bellingham et al. 2016). Metrics and indicators are designed to match the scale of the questions being asked and to be robust to data uncertainty.

Integration of new technologies. Modular design ensures new methods such as emerging DNA techniques and data from remote sensing can be added to improve resolution and coverage without changing the overall framework.

Legislative mandate and public interest. The national framework provides a defensible, replicable means of assessing biodiversity nationally to meet the requirements of the Environmental Reporting Act 2015 and to assist reporting on international agreements (e.g. progress towards the Aichi Targets set for 2020 under the Convention on Biological Diversity). It also provides the evidence base for setting policy and assessing its effectiveness. It could be used to evaluate community conservation projects and to provide the evidence for environmental certification in primary production sectors.

cultural, biological and economic potential of those ecosystems. Whakapapa is a key part of Māori ideology and a way of understanding an ecosystem in its entirety. It is more than just the narrow approach of ancestral lineage between humans: it also refers to the tangible and intangible genealogical connections, relationships, and linkages between the natural environment and the cosmological domain (Harmsworth & Awatere 2013). The construct of ihi (essential force) is used to gauge the vitality of the mana (authority) and mauri (life essence) of whakapapa. Environmental metabarcoding offers a modern-day approach to comprehending the vitality of ecosystems.

DNA metabarcoding provides an opportunity to identify those species contributing to ecosystem health and integrity that are relevant and important to Māori (Lyver et al. 2016). Metabarcoding can provide evidence of organisms relevant to mahinga kai (food species), taonga (highly valued species), kaitiaki (guardian species) or taha wairua (species linked to the spiritual domain) across different habitats and ranges, especially when these species are rare, cryptic, nocturnal or otherwise difficult to assess using conventional methods (e.g. āwheto/ngutara, the caterpillar of one of the forest ghost moths *Dumbletonius* or *Aoraia* spp., that historically was used for dyes in traditional Māori tattooing and has been infected by vegetable caterpillar fungus, *Ophiocordyceps robertsii*). Distributions of species reported by community members as historically abundant and in decline could be assessed using metabarcoding techniques. Tūhoe Tuawhenua elders have reported that harore (i.e. the fruiting body of the *Armillaria novaezelandiae* mushroom) has become harder to find, along with a noticeable dilution of the pungent odour associated with the fungi in the forest during late autumn. However, fruiting bodies are ephemeral and may not reflect the presence of vegetative mycelium. As such, it is unclear if the decline in fruiting bodies reflects a loss of mycelial biomass and ecosystem function. Metabacoding (or species-specific probes) would provide a more complete understanding of species occupancy across all life history stages, can be conducted at any time of the year, and can determine whether observed changes in the abundance of fruiting bodies reflect an overall population decline.

> *In fact this is the time to be collecting harore* [bush mushrooms] *during the winter months, but as you know that sort of food is diminishing, in the past 4 years we have seen a significant decline in harore. There should be consistency every year, in the past it would always return, if that does not happen it is because someone has tampered with it, or it was not picked right, or maybe it wasn't looked after. When my sister rings me, "the harore are ready" you would know through her that the harore are plentiful, but in recent years there has been nothing.* (Spady Kutia, Mātauranga o te taiao interview, 2014).

Metabarcoding techniques could also complement a Māori biodiversity monitoring approach more widely. Treaty of Waitangi settlements have given iwi greater management and reporting responsibilities for regions of high conservation and biodiversity value (e.g. Tūhoe and Te Urewera; Whanganui and Te Awa Tupua). Alongside their own biodiversity monitoring approaches, metabarcoding approaches could assist iwi and hapū with biodiversity inventories within their rohe and national environmental reporting requirements (e.g. Environmental Reporting Act 2015). Metabarcoding approaches could also provide community-level information on the impact of species that can change the āhua (character or nature) of an ecosystem.

Members of the Tūhoe Tuawhenua community reported concern that wasps have changed the āhua of their forest by preying heavily on its invertebrate fauna. The 'language' of the forest has changed, with a substantial increase in the sound associated with wasps, combined with a reduction in the sound associated with other invertebrate fauna. Monitoring of eDNA within an invasive wasp nest (*Vespula* spp.) could identify the invertebrate species (e.g. kākāpohae, the bush dragonfly *Antipodochlora braueri*; or whē, the stick insects *Tectarchus ovobessus*, *Argosarchus horridus*) most at risk from wasp predation in different habitats at different times of the year.

Case study 3: Primary sector and biosecurity outcomes

Metabarcoding data sets are able to provide information on the entire trophic web, which means they offer novel opportunities to determine how ecosystem functions relate to patterns in biota, particularly soil biota. This knowledge has the potential to contribute directly to primary sector outcomes such as improved farm productivity, profitability and resilience. For example, growing evidence links soil biology with critical ecosystem functions such as crop yield and nutrient cycling (Bender et al. 2015), disease suppression (Sánchez-Moreno & Ferris 2007), terroir (Gilbert et al. 2014), and resistance to drought stress (Orwin et al. 2015). This has led to increasingly widespread agricultural application of live biological cultures, for example Bayer Crop Science's Serenade® which is a patented strain of *Bacillus subtilis* bacterium. Use of metabarcoding techniques by the primary sector could provide data validating biological product effectiveness or underpinning the development of sustainable management strategies that support natural ecological processes and their resilience (de Vries et al. 2013; Nielsen et al. 2015; Delgado-Baquerizo et al. 2016). A dual benefit of primary sector adoption of a metabarcoding approach is that the same data sets would also provide information about biodiversity on production land. Such data are currently lacking for many environments in New Zealand and would inform national (e.g. New Zealand Ministry for the Environment State of the Environment annual reports), sector (e.g. The New Zealand Sustainability Dashboard Project reports), and farm-scale environmental reporting, with opportunities to garner market advantage and support the primary sectors sustainable licence to operate.

DNA metabarcoding has proven biosecurity application in freshwater ecosystems (e.g. Collins et al. 2013), and has the potential to be an important terrestrial biosecurity tool. For example, metabarcoding could be applied to screen bulk samples collected from light traps (e.g. Yu et al. 2012) at biosecurity incursion hotspots such as maritime ports and airports. The ability to screen for multiple pest species within a single metabarcoding data set would be a critical advance in the early detection process, even though more specific verification tests will be required to provide the level of evidence necessary for decisions around market access. Initial integration within current biosecurity monitoring networks is important to ensure metabarcoding is leveraging off existing resources, and that it is suitable for widespread field implementation. The broad application of DNA metabarcoding by a range of researchers and organisations will provide additional biosecurity benefits by increasing our knowledge of what species are already present in New Zealand. These data (both sequence data and

archived DNA samples) are a baseline from which to screen for new organisms, and to test species-specificity of DNA-based diagnostic tools. Post-border biosecurity applications are likely to be focused on detecting the distribution and spread of species that are difficult to otherwise detect using conventional methods (e.g. Macdonald & Sarre 2015). Therefore, there are significant opportunities for DNA metabarcoding to contribute to positive biosecurity outcomes.

Methodological uncertainties associated with DNA metabarcoding

Although there are clear benefits to combining conventional and DNA metabarcoding biodiversity assessment techniques, there are caveats. The detection of individual species in environmental or pooled-organism samples depends on several assumptions, including efficient and consistent polymerase chain reaction (PCR) amplification across all species in the sample and sufficient sequencing depth to detect all species represented within the amplified DNA. The detection of individual species also assumes that DNA sequences can be confidently linked to species names. Many ecological applications also require the assignment of DNA sequences to species or some other consistent taxonomic grouping. The assignment of sequences to species requires an adequate taxonomic framework for all species in the sampled community and the genetic characterisation of those species for the target gene region(s). For groups in which it is difficult to study individual species (e.g. fungi) or those that are hyperdiverse (e.g. fungi, invertebrates), complete species-level classifications are unlikely ever to be developed (e.g. Carbayo & Marque 2011; Buckley et al. 2015). Although partial reference data exist for metabarcoding of some taxonomic groups in New Zealand (e.g. vertebrates), most New Zealand biodiversity has not been genetically characterised. A further issue is reference data quality, as curated databases exist for bacteria and fungi, but not for all groups, leading to a high risk of errors (Nilsson et al. 2006; Bidartondo 2008).

The lack of stable classifications with accompanying genetic data for much of the biota means that DNA metabarcoding approaches will depend, at least in part, on clustering DNA sequences into operational taxonomic units (OTUs) (e.g. Kõljalg et al. 2005; Ratnasingham & Hebert 2013). The rules for grouping sequences into OTUs can be applied consistently across studies and varied retrospectively, which gives the OTU approach of DNA metabarcoding an advantage over the 'recognisable taxonomic units' or RTUs approach currently used by many ecological studies of invertebrates, which sorts specimens into RTUs based on their morphology (Ward & Stanley 2004). OTUs with DNA sequences similar to archived reference sequences can be identified to varying taxonomic levels. As the genetic characterisation of biota proceeds, and classifications are revised and linked with genetic information, more and more OTUs will be associated with reference DNA data and provided with taxonomic names. Despite debates about 'DNA taxonomy' (e.g. Tautz et al. 2003; Ratnasingham & Hebert 2013; Hibbett 2016), this approach is likely to become the norm, driven by the accelerating rate at which metabarcoding data are collected. The amount, quality, and accessibility of reference data deposited in international databases need significant improvement to more fully exploit the power of DNA to assess biodiversity at a relevant taxonomic resolution, and to detect invading species.

Quality reference data are particularly essential for optimal biosecurity outcomes in order to mitigate false positive matches to unwanted organisms. A widely publicised example is the putative detection of plague and anthrax in the New York subway (Afshinnekoo et al. 2015), which gained widespread media coverage but with further analysis was shown to be incorrect (Gonzalez et al. 2016). Currently, incomplete reference data sets and uncertain bioinformatics standards can result in best species matches for DNA sequences that are the most-studied organisms within a group, and these are often pathogens or pest species. For example, DNA reference data for many fly species in New Zealand are lacking, which causes many fly DNA sequences to match well-studied organisms such as Mediterranean fruit fly (*Ceratitis capitata*, an unwanted organism). Although such matches have a relatively low certainty, this raises the possibility of false positive detections when the data are not properly interpreted. Another biosecurity issue is the detection of trace amounts of an unwanted organism and the inability of metabarcoding techniques to prove that that organism was recently alive and/or functionally active. These biosecurity issues need to be managed in current and future DNA metabarcoding studies in New Zealand to prevent the occurrence of unnecessary and costly trade barriers.

In the analysis of DNA data, 97% sequence similarity is often used as a threshold for clustering amplified DNA sequences into 'taxa' (OTUs), yet for the gene regions most commonly used in DNA metabarcoding studies (e.g. 16S, 18S, ITS, COI), this may not discriminate all species. Taxonomic resolution can be improved by tailoring similarity thresholds for different taxa (Kõljalg et al. 2013), but even within taxonomic groups the correlation between species status and sequence divergence is at best weak (Wiemers & Fiedler 2007). For this reason, molecular systematics has turned to more sophisticated methods for species delimitation, including the use of information from more than gene region to define a species (Rannala 2015). However, it is not yet possible to apply these approaches to pooled samples. Increased taxonomic resolution may come through technological advances (e.g. longer read lengths), improved bioinformatics pipelines, and the development of primers better tailored for resolving taxa of interest. However, the most significant and immediate gains are to be made by increasing the taxonomic coverage of reference sequence databases through barcoding initiatives, particularly focussing on poorly characterised taxa (e.g. Hebert et al. 2013). Nevertheless, such issues indicate that DNA metabarcoding studies are currently better suited to the analysis of broad ecological and community-level patterns. Analyses that require more accurate species-level resolution are likely to require targeted improvements to current reference databases.

Methodological issues may also bias the interpretation and accuracy of DNA metabarcoding data sets. Biases in the efficacy of DNA isolation (Kim & Bae 2011), because not all DNA can be extracted from environmental samples (e.g. for certain soils, Huang et al. 2016), is an often undocumented but important consideration. PCR biases, including factors such as nucleotide variation at primer binding sites and variable sequence length, mean that not all targeted DNA has the same probability of amplification (Clarke et al. 2014). Furthermore, some 'universal' primers fail to amplify entire groups of taxa, e.g. ferns are typically not amplified by some 'plant' primers that amplified the chloroplast tRNA gene (trnL; Martínez-García et al. 2015); arbuscular mycorrhizal fungi are not well amplified by the most widely used primers for amplification of the fungal internal transcribed spacer (ITS1 – 5.8S – ITS2

regions; Dickie & St John 2016); and little is known about the extent to which these biases might affect the characterisation of communities of poorly studied taxa. Further biases arise from sequencing errors (e.g. favouring of sequences with shorter read lengths by Illumina), and the treatment of chimeric and low-copy sequences during bioinformatic processing (Lindgreen et al. 2016). These represent significant sources of uncertainty when it comes to interpreting the ecological patterns within and across DNA metabarcoding data sets.

The longevity of DNA in environmental samples is a key consideration. Under frozen conditions DNA molecules can be recovered from environmental samples (e.g. ice or permafrost) that are hundreds of thousands of years old (Willerslev et al. 2007). Even in temperate and aerobic substrates such as New Zealand forest soils, DNA signals can persist for thousands of years after a species has vanished from a particular site (Wilmshurst et al. 2014; Huang et al. 2016). Legacy DNA may also be an issue with recent land use change (Barnes & Turner 2016). The impact of legacy DNA on metabarcoding-based biodiversity assessments is poorly understood (Carini et al. 2016), but can be reduced in practice. In particular, DNA fragments become shorter over time, and thus legacy DNA will be less likely to amplify with generally used primers. Further, DNA from current biodiversity is likely to be significantly more abundant. Other strategies, such as using soil samples from the uppermost layers and increasing the length of target sequences, will reduce the legacy DNA detected. Predictable patterns of damage seen in ancient DNA molecules can also be assessed in metagenomic data sets (Ginolhac et al. 2011) to help exclude legacy DNA.

Finally, as DNA metabarcoding starts to be used to address ecological questions, sample design becomes even more important, because this limits interpretability. Initial metabarcoding applications have mostly tackled the substantial technical hurdles of sampling DNA under field conditions – the 'proof of concept' (e.g. Drummond et al. 2015). To now benefit from the rich information DNA metabarcoding can provide, robust sampling designs are needed, built on a quantitative understanding of spatial variability among samples, detection probabilities (e.g. Furlan et al. 2016), repeatability by field teams, and uncertainty and statistical power to determine effect sizes that can be interpreted ecologically. Most current metabarcoding studies lack these design elements (e.g. uncertainties are large and the ability to robustly detect change is unproven), and when coupled with the challenges of analysing 'big data' this limits the ecological questions that can be addressed.

While the uncertainties mentioned here are important to consider when applying DNA metabarcoding approaches, they are likely to be overcome with ongoing research effort. For example, taxonomic classifications and the coverage and depth of reference databases are continually improving (Ratnasingham & Hebert 2013), and the quantification of uncertainty in sampling, molecular and bioinformatics methods is well underway (e.g. Furlan et al. 2016; Lindgreen et al. 2016). Furthermore, an increasing uptake of metabarcoding approaches by ecological researchers (Bohmann et al. 2014; Creer et al. 2016) is likely to improve the design and ecological interpretation of DNA metabarcoding data sets. Nevertheless, the current uncertainties around DNA approaches add further weight to our argument that the integration of DNA approaches and conventional biodiversity techniques is critical to enable assessments of data quality and ensure that data are used appropriately.

Future research directions

It is important that the needs of end users are addressed in future research so that the key barriers to the uptake of metabarcoding are removed. The Biological Heritage National Science Challenge held a series of workshops in 2015 and 2016 involving representatives from research organisations, central and local government, and primary sector groups. These workshops identified a number of pressing research questions that need addressing if DNA metabarcoding is to be successfully integrated with conventional biodiversity assessment practices (Table 3). The uncertainties identified above limit both our ability to use the method for ecological research and its practical uptake by end-users. Perhaps the most pressing issue, in terms of biodiversity assessment, is the as-yet-unproven ability of DNA metabarcoding to detect ecologically meaningful changes in communities and populations over time. Resolving these uncertainties is a key objective of the Biological Heritage National Science Challenge, and requires significant investment in taxonomy and barcoding to improve reference databases. Comparative studies of DNA metabarcoding alongside conventional monitoring are required to validate the ability of DNA metabarcoding, particularly those that investigate the relationship between metabarcoding data and alternative measures of species abundance, and those that look at how amounts of legacy DNA change through time. Coordinated research addressing a range of specific questions relating to sampling design, scale, uncertainty and data interpretation is also required (see Table 3).

Rapidly changing DNA technologies influence methods for monitoring long-term changes in biodiversity, and consistent methods need to be applied at multiple time periods to reliably distinguish real change from methodological variation. For example, in the near future non-PCR metagenomic techniques may be used to sequence and assemble diverse environmental genomes. This rapid advance is both an asset, because it expands the horizon of what is possible, but also a challenge, because it can make existing methods appear obsolete and detection of temporal trends difficult. However, there are ways to manage such technological change, the most critical being clear sample design and appropriate sample archiving (physical samples and extracted DNA). This allows re-analysis of samples using standard techniques in the future. Archiving DNA sequence data will also allow revisiting of data sets as reference databases improve, providing a significant long-term data resource (Cary & Fierer 2014).

Expectations must be managed to closely match what DNA-based biodiversity assessment approaches can actually deliver. Molecular approaches possess their own set of challenges including uncertainties resulting from biases associated with PCR, different DNA extraction methods and bioinformatics pipelines (Clarke et al. 2014; Lindgreen et al. 2016). Moreover, reliable assignment of DNA sequences to named taxonomic units (Schloss & Westcott 2011; Creer et al. 2016) and ecological interpretation of the resulting highly complex community data (e.g. Ma et al. 2016) also present challenges. As is typical for the initial uptake of any technology, early applications (e.g. Drummond et al. 2015) have focused on exploring the method. As our understanding of the properties of these methods evolves, there is an increasing need to move beyond this and into applications that provide rich information for a range of purposes, including conservation planning and land management decisions (Hajibabaei et al. 2016). Nevertheless, the potential benefits of a framework for

Table 3. Key end-user uncertainties and outstanding research questions derived from proceedings of four DNA metabarcoding workshops hosted by the Biological Heritage National Science Challenge.

Key end-user uncertainties

What are the means to ensure protection of Māori intellectual property with respect to indigenous organisms and associated data and metadata?

Can DNA sampling be integrated with (rather than replace) existing monitoring programmes?

What is the cost/benefit of DNA techniques compared to other existing tools?

Is DNA metabarcoding able to provide information on species abundance?

Can DNA metabarcoding be used to detect temporal changes in species/communities?

Does the precision and variability of DNA metabarcoding make it a useable measure for trend analysis of biodiversity and abundance?

How do we manage data, data use rights, data security, attribution, and information release?

What are the potential risks to markets and trade as new organisms are identified, and how will such information be handled?

Are we able to analyse the huge data sets in a timely and efficient way and summarise the results for other users?

How can the outcomes of DNA metabarcoding research be used to inform management perspectives or actions?

Key future research questions

What is the variability of environmental DNA across space and time (including legacy DNA), and what are the implications of this for sampling design?

Can primer and PCR biases be quantified and methods developed to minimise their effects?

How do you optimise DNA extraction for different environmental samples in a way that preserves comparability across samples?

How does choice of bioinformatics pipeline affect the results and conclusions?

How is DNA data (sequence read count) related to in-field abundance?

How do we best share, archive and analyse DNA metabarcoding data sets to address broad-scale ecological questions?

---

the monitoring of all taxa are vast (Fischer et al. 2010), and alignment of DNA metabarcoding and conventional methods offers a way forward to that goal.

Acknowledgements

References

Afshinnekoo E, Meydan C, Chowdhury S, Jaroudi D, Boyer C, Bernstein N, Maritz Julia M, Reeves D, Gandara J, Chhangawala S, Ahsanuddin S, Simmons A, Nessel T, Sundaresh B, Pereira E, Jorgensen E, Kolokotronis S-O, Kirchberger N, Garcia I, Gandara D, Dhanraj S, Nawrin T, Saletore Y, Alexander N, Vijay P, Hénaff Elizabeth M, Zumbo P, Walsh M, O'Mullan Gregory D, Tighe S, Dudley Joel T, Dunaif A, Ennis S, O'Halloran E, Magalhaes Tiago R, Boone B, Jones Angela L, Muth Theodore R, Paolantonio Katie S, Alter E, Schadt Eric E, Garbarino J, Prill Robert J, Carlton Jane M, Levy S, Mason Christopher E 2015. Geospatial resolution of human and bacterial diversity with city-scale metagenomics. Cell Systems 1: 72–87.

Allen RB, Bellingham PJ, Wiser SK 2003. Developing a forest biodiversity monitoring approach for New Zealand. New Zealand Journal of Ecology 27: 207–220.

Allen RB, Bellingham PJ, Forsyth DM, MacLeod CJ, Wright E 2009. Implementing an inventory and monitoring programme for the Department of Conservation's Natural Heritage Management System. Landcare Research Contract Report LC0809/154, prepared for the Department of Conservation. 40 p.

Barnes MA, Turner CR 2016. The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics 17: 1–17.

Bellingham PJ, Overton JM, Thomson FJ, MacLeod CJ, Holdaway RJ, Wiser SK, Brown M, Gormley AM, Collins D, Latham DM, Bishop C, Rutledge DT, Innes JG, Warburton B 2016. Standardised terrestrial biodiversity indicators for use by regional councils. Landcare Research Contract Report LC2109, for Regional Councils' Biodiversity Monitoring Working Group, Auckland Council. 404 p.

Bender SF, van der Heijden MGA 2015. Soil biota enhance agricultural sustainability by improving crop yield, nutrient uptake and reducing nitrogen leaching losses. Journal of Applied Ecology 52: 228–239.

Bidartondo MI 2008. Preserving accuracy in GenBank. Science 319: 1616a.

Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, Yu DW, de Bruyn M 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution 29: 358–367.

Boyer S, Wratten SD, Holyoake A, Abdelkrim J, Cruickshank RH 2013. Using next-generation sequencing to analyse the diet of a highly endangered land snail (*Powelliphanta augusta*) feeding on endemic earthworms. PLoS ONE 8: e75962.

Buckley TR, Krosch M, Leschen RAB 2015. Evolution of New Zealand insects: summary and prospectus for future research. Austral Entomology 54: 1–27.

Calvignac-Spencer S, Merkel K, Kutzner N, Kühl H, Boesch C, Kappeler PM, Metzger S, Schubert G, Leendertz FH 2013. Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. Molecular Ecology 22: 915–924.

Carbayo F, Marque AC 2011. The cost of describing the entire animal kingdom. Trends in Ecology & Evolution 26: 154–155.

Carini P, Marsden PJ, Leff JW, Morgan EE, Strickland MS, Fierer N 2016. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. Nature Microbiology 2: 16242.

Cary SC, Fierer N 2014. The importance of sample archiving in microbial ecology. Nature Reviews Microbiology 12: 789–790.

Clarke LJ, Soubrier J, Weyrich LS, Cooper A 2014. Environmental metabarcodes for insects: in silico PCR reveals potential for taxonomic bias. Molecular Ecology Resources 14: 1160–1170.

Collins RA, Armstrong KF, Holyoake AJ, Keeling S 2013. Something in the water: biosecurity monitoring of ornamental fish imports using environmental DNA. Biological Invasions 15: 1209–1215.

Creer S, Deiner K, Frey S, Porazinska D, Taberlet P, Thomas K, Potter C, Bik H 2016. The ecologist's field guide to sequence-based identification of biodiversity. Methods in Ecology and Evolution 7: 1008–1018.

de Vries FT, Thébault E, Liiri M, Birkhofer K, Tsiafouli MA, Bjørnlund L, Bracht Jørgensen H, Brady MV, Christensen S, de Ruiter PC, d'Hertefeldt T, Frouz J, Hedlund K, Hemerik L, Hol WHG, Hotes S, Mortimer SR, Setälä H, Sgardelis SP, Uteseny K, van der Putten WH, Wolters V, Bardgett RD 2013. Soil food web properties explain ecosystem services across European land use systems. Proceedings of the National Academy of Sciences of the USA 110: 14296–14301.

Delgado-Baquerizo M, Maestre FT, Reich PB, Jeffries TC, Gaitan JJ, Encinar D, Berdugo M, Campbell CD, Singh BK 2016. Microbial diversity drives multifunctionality in terrestrial ecosystems. Nature Communications 7: 10541.

Dickie IA, St John MG 2016. Second-generation molecular understanding of mycorrhizas in soil ecosystems. In: Martin F ed. Molecular Mycorrhizal Symbiosis. Hoboken, NJ, USA, John Wiley & Sons. Pp. 473–491.

Doyle EJ, Hogg ID, Banks JC 2015. Finding whio: detection of blue duck using environmental DNA. Genome 58: 213.

Drummond AJ, Newcomb RD, Buckley TR, Xie D, Dopheide A, Potter BCM, Heled J, Ross HA, Tooman L, Grosser S, Park D, Demetras NJ, Stevens MI, Russell JC, Anderson SH, Carter A, Nelson N 2015. Evaluating a multigene environmental DNA approach for biodiversity assessment. GigaScience 4: 46.

Environmental Reporting Act 2015. Public Act 2015 No 87. New Zealand. www.legislation.govt.nz (Accessed 21 April 2017).

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.

Ficetola GF, Miaud C, Pompanon F, Taberlet P 2008. Species detection using environmental DNA from water samples. Biology Letters 4: 423–425.

Fischer M, Bossdorf O, Gockel S, Hänsel F, Hemp A, Hessenmöller D, Korte G, Nieschulze J, Pfeiffer S, Prati D, Renner S, Schöning I, Schumacher U, Wells K, Buscot F, Kalko EKV, Linsenmair KE, Schulze E-D, Weisser WW 2010. Implementing large-scale and long-term functional biodiversity research: the Biodiversity Exploratories. Basic and Applied Ecology 11: 473–485.

Furlan EM, Gleeson D, Hardy CM, Duncan RP 2016. A framework for estimating the sensitivity of eDNA surveys. Molecular Ecology Resources 16: 641–654.

Gibson JF, Shokralla S, Curry C, Baird DJ, Monk WA, King I, Hajibabaei M 2015. Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. PLoS ONE 10: e0138432.

Gilbert JA, van der Lelie D, Zarraonaindia I 2014. Microbial terroir for wine grapes. Proceedings of the National Academy of Sciences of the USA 111: 5–6.

Ginolhac A, Rasmussen M, Gilbert MT, Willerslev E, Orlando L 2011. MapDamage: testing for damage patterns in ancient DNA sequences. Bioinformatics 27: 2153–2155.

Gonzalez A, Vázquez-Baeza Y, Pettengill JB, Ottesen A, McDonald D, Knight R 2016. Avoiding pandemic fears in the subway and conquering the platypus. mSystems 1: e00018-16.

Gordon DP 2012. New Zealand inventory of biodiversity (Volumes 1–3). Christchurch, New Zealand, Canterbury University Press. 566 p.

Haaland C, Naisbit RE, Bersier LF 2011. Sown wildflower strips for insect conservation: a review. Insect Conservation and Diversity 4: 60–80.

Hajibabaei M, Baird DJ, Fahner NA, Beiko R, Golding GB 2016. A new way to contemplate Darwin's tangled bank: how DNA barcodes are reconnecting biodiversity science and biomonitoring. Philosophical Transactions of the Royal Society B 371: 20150330.

Harmsworth GR, Awatere S 2013. Indigenous Māori knowledge and perspectives of ecosystems. In: Dymond J ed. Ecosystem services in New Zealand: conditions and trends. Palmerston North, New Zealand, Manaaki Whenua Press. Pp. 274–286.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society B 270: 313–321.

Hebert PDN, deWaard JR, Zakharov EV, Prosser SWJ, Sones JE, McKeown JTA, Mantle B, La Salle J 2013. A DNA 'Barcode Blitz': rapid digitization and sequencing of a natural history collection. PloS One 8: e68535.

Hibbett D 2016. The invisible dimension of fungal diversity. Science 351: 1150–1151.

Hitchmough R 2013. Summary of changes to the conservation status of taxa in the 2008–11 New Zealand Threat Classification System listing cycle. New Zealand Threat Classification Series 1. Wellington, Department of Conservation. 20 p.

Holdaway R, McNeill S, Mason NWH, Carswell F 2014. Propagating uncertainty in plot-based estimates of forest carbon stock and carbon stock change. Ecosystems 17: 627–640.

Huang YT, Lowe DJ, Zhang H, Cursons R, Young JM, Churchman GJ, Schipper LA, Rawlence NJ, Wood JR, Cooper A 2016. A new method to extract and purify DNA from allophanic soils and paleosols, and potential for

paleoenvironmental reconstruction and other applications. Geoderma 274: 114–125.

Jerde CL, Mahon AR, Chadderton WL, Lodge DM 2011. 'Sight-unseen' detection of rare aquatic species using environmental DNA. Conservation Letters 4: 150–157.

Kim K-H, Bae J-W 2011. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. Applied and Environmental Microbiology 77: 7663–7668.

Kõljalg U, Larsson KH, Abarenkov K, Nilsson RH, Alexander IJ, Eberhardt U, Erland S, Høiland K, Kjøller R, Larsson E, Pennanen T, Sen R, Taylor AFS, Tedersoo L, Vrålstad T 2005. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. New Phytologist 166: 1063–1068.

Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Dueñas M, Grebenc T, Griffith GW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lücking R, Martín MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Põldmaa K, Saag L, Saar I, Schüßler A, Scott JA, Senés C, Smith ME, Suija A, Taylor DL, Telleria MT, Weiss M, Larsson K-H 2013. Towards a unified paradigm for sequence-based identification of fungi. Molecular Ecology 22: 5271–5277.

Lee W, McGlone M, Wright E 2005. Biodiversity inventory and monitoring: a review of national and international systems and a proposed framework for future biodiversity monitoring by the Department of Conservation. Landcare Research Contract Report LC0405/122 prepared for the Department of Conservation. 213 p.

Lindgreen S, Adair KL, Gardner PP 2016. An evaluation of the accuracy and speed of metagenome analysis tools. Scientific Reports 6: 19233.

Lyver PO'B, Timoti P, Jones CJ, Richardson SJ, Tahi BL, Greenhalgh S 2016. An indigenous community-based monitoring system for assessing forest health in New Zealand. Biodiversity and Conservation: DOI: 10.1007/s10531-016-1142-6.

Ma B, Wang H, Dsouza M, Lou J, He Y, Dai Z, Brookes PC, Xu J, Gilbert JA 2016. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. ISME Journal 10: 1891–1901.

MacDonald AJ, Sarre SD 2015. Species assignment from trace DNA sequences: an in silico assessment of the test used to survey for foxes in Tasmania. Journal of Applied Ecology 52: 1649–1655.

MacLeod CJ, Affeld K, Allen RB, Bellingham PJ, Forsyth DM, Gormley AM, Holdaway RJ, Richardson SJ, Wiser SK 2012. Department of Conservation biodiversity indicators: 2012 assessment. Landcare Research Contract Report LC1102. Lincoln, Landcare Research. 63 p.

Martínez-García LB, Richardson SJ, Tylianakis JM, Peltzer DA, Dickie IA 2015. Host identity is a dominant driver of mycorrhizal fungal community composition during ecosystem development. New Phytologist 205: 1565–1576.

Morales SE, Jha N, Saggar S 2015. Biogeography and biophysicochemical traits link N_2O emissions, N_2O emission potential and microbial communities across New Zealand pasture soils. Soil Biology and Biochemistry 82: 87–98.

Morrison-Whittle P, Goddard MR 2015. Quantifying the relative roles of selective and neutral processes in defining eukaryotic microbial communities. ISME Journal 9: 2003–2011.

Nielsen UN, Wall DH, Six J 2015. Soil biodiversity and the environment. Annual Review of Environment and Resources 40: 63–90.

Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson K-H, Kõljalg U 2006. Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. PLoS ONE 1: e59.

Oliver TH, Isaac NJB, August TA, Woodcock BA, Roy DB, Bullock JM 2015. Declining resilience of ecosystem functions under biodiversity loss. Nature Communications 6: 10122.

Orwin KH, Dickie IA, Wood JR, Bonner KI, Holdaway RJ 2015. Soil microbial community structure explains the resistance of respiration to a dry–rewet cycle, but not soil functioning under static conditions. Functional Ecology 30: 1430–1439.

Peltzer DA, Allen RB, Bellingham PJ, Richardson SJ, Wright EF, Knightbridge PI, Mason NWH 2014. Disentangling drivers of tree population size distributions. Forest Ecology and Management 331: 165–179.

Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, Coops NC, Dulloo E, Faith DP, Freyhof J, Gregory RD, Heip C, Höft R, Hurtt G, Jetz W, Karp DS, McGeoch MA, Obura D, Onoda Y, Pettorelli N, Reyers B, Sayre R, Scharlemann JPW, Stuart SN, Turak E, Walpole M, Wegmann M 2013. Essential biodiversity variables. Science 339: 277–278.

Rannala B 2015. The art and science of species delimitation. Current Zoology 61: 846–853.

Ratnasingham S, Hebert PD 2013. A DNA-based registry for all animal species: the barcode index number (BIN) system. PLoS One 8: e66213.

Rooney N, McCann KS 2012. Integrating food web diversity, structure and stability. Trends in Ecology & Evolution 27: 40–46.

Sánchez-Moreno S, Ferris H 2007. Suppressive service of the soil food web: effects of environmental management. Agriculture, Ecosystems & Environment 119: 75–87.

Schloss PD, Westcott SL 2011. Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. Applied and Environmental Microbiology 77: 3219–3226.

Sharp CE, Smirnova AV, Graham JM, Stott MB, Khadka R, Moore TR, Grasby SE, Strack M, Dunfield PF 2014. Distribution and diversity of Verrucomicrobia methanotrophs in geothermal and acidic environments. Environmental Microbiology 16: 1867–1878.

Soliveres S, Manning P, Prati D, Gossner MM, Alt F, Arndt H, Baumgartner V, Binkenstein J, Birkhofer K, Blaser S, Blüthgen N, Boch S, Böhm S, Börschig C, Buscot F, Diekötter T, Heinze J, Hölzel N, Jung K, Klaus VH, Klein A-M, Kleinebecker T, Klemmer S, Krauss J, Lange M, Morris EK, Müller J, Oelmann Y, Overmann J, Pašalić E, Renner SC, Rillig MC, Schaefer HM, Schloter M, Schmitt B, Schöning I, Schrumpf M, Sikorski J, Socher SA, Solly EF, Sonnemann I, Sorkau E, Steckel J, Steffan-Dewenter I, Stempfhuber B, Tschapka M, Türke M, Venter P, Weiner CN, Weisser WW, Werner M, Westphal C, Wilcke W, Wolters V, Wubet T, Wurst S, Fischer M,

Allan E 2016. Locally rare species influence grassland ecosystem multifunctionality. Philosophical Transactions of the Royal Society B 371: 20150269.

Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP 2003. A plea for DNA taxonomy. Trends in Ecology & Evolution 18: 70–74.

Thomsen PF, Willerslev E 2015. Environmental DNA – an emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183: 4–18.

Wagg C, Bender SF, Widmer F, van der Heijden MGA 2014. Soil biodiversity and soil community composition determine ecosystem multifunctionality. Proceedings of the National Academy of Sciences of the USA 111: 5266–5270.

Wakelin SA, Barratt BIP, Gerard E, Gregg AL, Brodie EL, Andersen GL, DeSantis TZ, Zhou J, He Z, Kowalchuk GA, O'Callaghan M 2013. Shifts in the phylogenetic structure and functional capacity of soil microbial communities follow alteration of native tussock grassland ecosystems. Soil Biology and Biochemistry 57: 675–682.

Ward DF, Stanley MC 2004. The value of RTUs and parataxonomy versus taxonomic species. New Zealand Entomologist 27: 3–9.

Waterhouse BR, Boyer S, Wratten SD 2014. Pyrosequencing of prey DNA in faeces of carnivorous land snails to facilitate ecological restoration and relocation programmes. Oecologia 175: 737–746.

Watts C, Holdaway R, Davis C, Wood J, Dickie I, Thomson F, Thornburrow D 2016. Invertebrate services within

Cape to City and comparison of environmental DNA with conventional invertebrate community monitoring: research synthesis 2015/2016. Landcare Research Contract Report LC2541, for Hawke's Bay Regional Council. 59 p.

Wiemers M, Fiedler K 2007. Does the DNA barcoding gap exist? A case study in blue butterflies (Lepidoptera: Lycaenidae). Frontiers in Zoology 4: 1–16.

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, Muscheler R, Barker J, Sharp M, Penkman KEH, Haile J, Taberlet P, Gilbert MTP, Casoli A, Campani E, Collins MJ 2007. Ancient biomolecules from deep ice cores reveal a forested southern Greenland. Science 317: 111–114.

Wilmshurst JM, Moar NT, Wood JR, Bellingham PJ, Findlater A, Robinson J, Stone C 2014. Pollen and ancient DNA provide conservation baselines for offshore islands in New Zealand. Conservation Biology 28: 202–212.

Wiser S, Bellingham PJ, Burrows LE 2001. Managing biodiversity information: development of New Zealand's National Vegetation Survey databank. New Zealand Journal of Ecology 25: 1–17.

Yang C, Wang X, Miller JA, de Blécourt M, Ji Y, Yang C, 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.

Yu DW, Ji Y, Emerson BC, Wang X, Ye C, Yang C, Ding Z 2012. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods in Ecology and Evolution 3: 613–623.