Standards, methodology and protocols for sampling and identification of grassland fungus species

Using eDNA from soil samples

First published 7th December 2021

Natural England Commissioned Report NECR374

www.gov.uk/natural-england

NATURAL ENGLAND

Natural England Commissioned Report NECR374

Standards, methodology and protocols for sampling and identification of grassland fungus species

Andrew P. Detheridge and Gareth W. Griffith

Published 7th December 2021

ISBN: 978-1-78354-776-0

Project details

This report should be cited as: Detheridge, A.P. and Griffith, G.W., 2021. Standards, methodology and protocols for sampling and identification of grassland fungus species. Natural England Commissioned Reports, Number NECR374.

Natural England Project manager

Debbie Leatherland; Debbie.leatherland@naturalengland.org.uk

Contractor

Institute of Biological, Environmental & Rural Sciences (IBERS), Aberystwyth University

Author

Andrew P. Detheridge and Gareth W. Griffith

Keywords

DNA, environmental DNA (eDNA), grassland fungi, soil, methods, protocols,

Further information

This report can be downloaded from the Natural England Access to Evidence Catalogue: **http://publications.naturalengland.org.uk/** . For information on Natural England publications contact the Natural England Enquiry Service on 0300 060 3900 or e-mail **enquiries@naturalengland.org.uk**.

Acknowledgements

We are grateful to Jan Sherry, Julia Korn and Andrew Lucas at Natural Resources Wales (and its predecessor Countryside Council for Wales) for funding the initial development of our eDNA metabarcoding work and also subsequent eDNA survey studies. We thank the following for access to valuable field sites: National Trust (Mynachdy'r Graig, Lundy), English Heritage (Kristyna Slivova and Antony O'Rourke at Down House) and IBERS Aberystwyth (Brignant). And for provision of fruitbody survey data: Peter Sturgess (Llanishen) and Andy McLay (Hardcastle Crags).

We are grateful to Natural England for funding of this desk study (Debbie Leatherland, Katie Clark and Tim Wilkins) and of several related projects (Hardcastle Crags [Andy McLay], Sussex High Weald [Janet Whitman] and Species Recovery Plan: Earthtongue fungi [Tim Wilkins]).

Contents

Foreword

DNA – based methods offer a significant opportunity to change how we monitor and assess biodiversity. However, for most techniques, there is still much development required before they can be used in routine monitoring.

Natural England has been exploring the further use of these methods for environmental monitoring for several years, delivering a series of reports which focus on the development of DNA-based methods with potential in a particular area.

The ability to use soil samples to identify grassland fungi is an exciting opportunity to add to evidence from traditional fruitbody surveys and not be limited by time of year. This project aimed to write up and publish the standards, methodology and protocols for sampling and identification of grassland fungus species, including discussing issues, limitations and opportunities. Published protocols will result in repeatability and consistency of methodologies, improving the comparability and confidence of results and findings.

1. Introduction

The precipitous loss of permanent semi-natural grasslands across Europe since the advent of mechanised agriculture, has been a major driver of species biodiversity loss during the latter half of the 20th century (Blackstock et al., 1999; Lawton et al., 1995). In the past the focus has been mainly on losses of plant and invertebrate species, but more recently attention has been given to the loss of fungal biodiversity through the destruction of semi-natural grassland habitats (Griffith et al., 2002; Griffith et al., 2013). Fungi are a central component in the functioning of these habitats through the provision of ecosystem services such as decomposition (van der Wal et al., 2013), maintenance of soil-structure (Rillig and Mummey, 2006) and plant nutrition via mycorrhizal associations (Read and Perez-Moreno, 2003).

The agricultural improvement of grasslands and the well-recorded losses in plant and animal biodiversity has more recently increased concerns about the status of the fungi that are associated with these habitats (Newton et al., 2003). Monitoring of macrofungi associated with grasslands and other habitats has traditionally relied on fruitbody (mushroom) surveys. However, fruitbodies are ephemeral, lasting only a few days to a few weeks and also exhibit differing phenological patterns. Additionally, varying meteorological conditions during late summer/autumn can strongly influence the amount and timing of fungal fruiting and such patterns can also change as climate alters (Boddy et al., 2014; Gange et al., 2007). As such, surveys require multiple visits over several years to provide a reasonable assessment of the macrofungal diversity at a given site.

The use of eDNA (environmental DNA), DNA extracted from environmental samples (e.g. soil, water, air, faeces, plant tissues etc.), has proven to be particularly useful for the detection and enumeration of organisms in natural environments. It is particularly useful for those organisms with cryptic (microscopic) phases in their lifecycles, and such approaches have been widely deployed to monitor the biodiversity of a wide range of habitats and species over the past 10-20 years (Deiner et al., 2021). In the case of fungi, such data are now contributing to a global database of fungal species distribution (Větrovský et al., 2020). The potential contribution of eDNA studies to fungal conservation is also recognised (Geml et al., 2014) and recently, evidence from an eDNA study was used to detect and locate fruitbodies of several rare grassland fungi at a UK site (where these species had not previously been observed), leading to its notification by Natural England as a Site of Special Scientific Interest (Griffith et al., 2018; Griffith et al., 2015; Natural_England, 2019). This is the first use globally of eDNA to contribute to the legal protection of a nature site.

eDNA analysis may involve detection/quantification of individual species using amplification via PCR (polymerase chain reaction) using pairs of specific oligonucleotide primers. Such methods are long established and now routinely undertaken for determining the presence of *Triturus cristatus* (great crested newt) via sampled pond water (Harper et al., 2018) or various bat species via faecal samples (multiplexed PCR using several PCR primer pairs, each specific to one bat species)(Harrington et al., 2019).

eDNA metabarcoding is a more recent and advanced approach than the more familiar process of ‘DNA barcoding’. Whilst the latter involves PCR and DNA sequencing of a barcode locus obtained from a single voucher specimen, the former deploys high-throughput (‘NextGen’) sequencing to provide a comprehensive species list for all the organisms within the target group (usually at phylum or kingdom level) from a given environmental sample. In this sense, DNA barcoding is akin to the identification of a single species via a taxonomic key, whereas eDNA metabarcoding equates to the gathering of community level data from a field quadrat (e.g. NVC for plants) where all the species present in the sample are identified and their relative abundance (e.g. % cover) assessed.

eDNA metabarcoding analysis entails collection of an environmental sample, extraction of DNA from all organisms in that sample (the eDNA), amplification using taxon-specific primers of a barcode region (a region of the genome that has been used to identify species) of the group of organisms in question, sequencing the products of this amplification and then processing the sequence data to provide a species list. For animals, the accepted barcode region is the mitochondrial cytochrome oxidase subunit 1 gene, whereas for most other biota various parts of the ribosomal RNA operon are used. It should be noted that the connection from the DNA sequences generated via eDNA metabarcoding to species names is dependent on the existence of published DNA barcode sequences obtained from voucher samples (a specimen of the species, ideally the type specimen, which has been identified by a taxonomic expert). Where no such specimen exists (e.g. for undescribed or poorly studied species), precise identification is not possible but identification to higher taxonomic level (genus/family) can be made.

The application of an eDNA survey method (metabarcoding) to fungi has many potential advantages over traditional survey techniques (Woodcock, 2020). The mycelial network of a fungal species will exist in the soil whether or not the organism is fruiting and so has the potential to be captured through eDNA surveying at all times and climatic conditions. Many species of fungi are hard to identify accurately in the field and require expertise and/or microscopy to differentiate. However, once a database of fungal sequences based on a barcode region is established with correct taxonomic assignments, then field expertise is not required by the person sampling the soil. Establishing such a database, however, is not trivial and requires expertise in fungal taxonomy and the collection of fruitbodies by surveyors to provide the raw material for barcoding. Many species of fungi either do not fruit (e.g. Glomeromycotina the arbuscular mycorrhizal fungi [AMF]) or have very small fruitbodies, such as many species of Ascomycota, and so are hard to spot in the field. In these circumstances an eDNA metabarcoding study may be the only way to efficiently analyse their diversity.

When considering an eDNA metabarcoding study there are aspects to consider, which will affect the outcome of any such analysis:-

- Soil sampling, capturing as close as practically possible to full diversity.
- Soil sample preparation.
- What region of the genome to select for the metabarcoding analysis

- Primer selection
- Laboratory DNA extraction PCR and sequencing methods
- Data processing

Once an eDNA survey is complete then we also need to consider how well the data compares to conventional survey methods, whether important species are missed by either method, and how could our surveying and data processing improve so that species are not missed. Fruitbody and eDNA surveys are never likely to be directly comparable. From eDNA we can determine which species are present and the relative sequence abundance of those species. This relative sequence abundance will not be directly comparable to fruitbody abundance within a quadrat as different factors are in play. For example, the relative sequence abundance will be affected by how many of the cores taken from a quadrat contain a particular species, this will be related to underground biomass but there will be a stochastic element dependent on the shape and interaction of the mycelial networks of different species, leading to the sampling either over or under representing particular species. Fruiting will also not directly correspond to underground biomass, weather and sward length play an important part. Some species that have been observed to fruit may be missed by eDNA if the mycelial network is not cored. Also some species that are found in an eDNA survey may not fruit as the biomass is too low, however, if a species is found in a relatively high sequence abundance (we use a figure of ca. 0.5% and above-see section 4), then we might expect fruiting when conditions are right, this can inform some targeted fruitbody surveys. In conclusion the relative sequence abundance is a useful measure but should not be taken as a definitive measure of relative biomass abundance. In a similar way an abundance of fruitbodies of a particular species in a good year for fruiting of a particular species may not reflect that species below ground biomass.

This report examines the factors that can influence results of metabarcoding surveys and draws upon this to make conclusions about best practice for the monitoring of fungal biodiversity in grassland habitats in the UK and make recommendations for potential future work.

2. Methodology

The eDNA metabarcoding process from start to finish of an eDNA analysis is illustrated in Fig. 1. Each step is analysed and explained below in more detail.

2.1 Sampling strategy, sampling and sampling timing

Sampling strategy is a much neglected aspect of soil metabarcoding analyses and a suitable design is important (Woodcock, 2020). The strategy employed needs to consider the question that is being asked of the analysis. For the purposes of this report, the aim will be to identify, as closely as possible, the complete fungal diversity of the area under consideration, rather than other questions such as mapping the spatial heterogeneity of soil fungi.

As well as variability across an area, fungal communities will also vary with depth (Upton et al., 2020); here we consider the fungal community profile in the top 10 cm of soil, which will encompass the surface fruiting species. However, on some, for example rendzina soils on limestone, the bedrock will often be shallower than this, whilst on others (e.g. Brignant podsol), the depth of topsoil is only ca. 5cm, leading to some sampling of non-organic horizons. The method of soil sampling must also be simple and easy to replicate with minimal training, here we employ a simple gauge auger (17 mm diameter) and a grid-sampling pattern across the area under consideration. Soil collected in this way can then be shipped to a processing laboratory. Soil augers come in many configurations; we consistently use a 17mm T-bar gauge auger made of hardened steel and where third parties undertake the fieldwork, we send an auger for use in sampling. The augers we use are made to order by Brunswick Ironworks Ltd., Caernarfon (http://www.brunswickironworks.co.uk/; cost £80+VAT in 2020) but could be manufactured by many metalwork companies.

Typically, we undertake sampling of 30x30 m quadrats, beginning 3.5 m diagonally in from one corner and taking cores on a 5x5 m grid, providing 36 cores per quadrat, with a fresh weight of pooled cores ranging from 650-850 g depending on the person sampling. We originally adopted the 30x30 m quadrat size from our earlier fruitbody surveying methodology (Griffith et al., 2006; Griffith et al., 2013), with this area being found to strike a good balance between coverage of a moderate area in a ca. 45 min time period. In our experience there are consistent person-to-person differences in the typical pooled-core weight, with inexperienced surveyors tending to collect more soil. When we have tested the effect of sampling of the same quadrat by different surveyors, we have sometimes observed small differences in the macrofungal species detected by different surveyors sampling the same quadrat (unpublished data). We ascribe the differences to the stochastic nature of placement of the auger for taking the cores. For example, if a quadrat contains a single small colony of a given species, 2m in diameter, its biomass is likely to be missed if coring takes place on a 5x5m grid. Different surveyors will vary in how closely they keep to the strict 5x5 m gridline and thus potentially capture or miss particular species which are present only at low abundance.

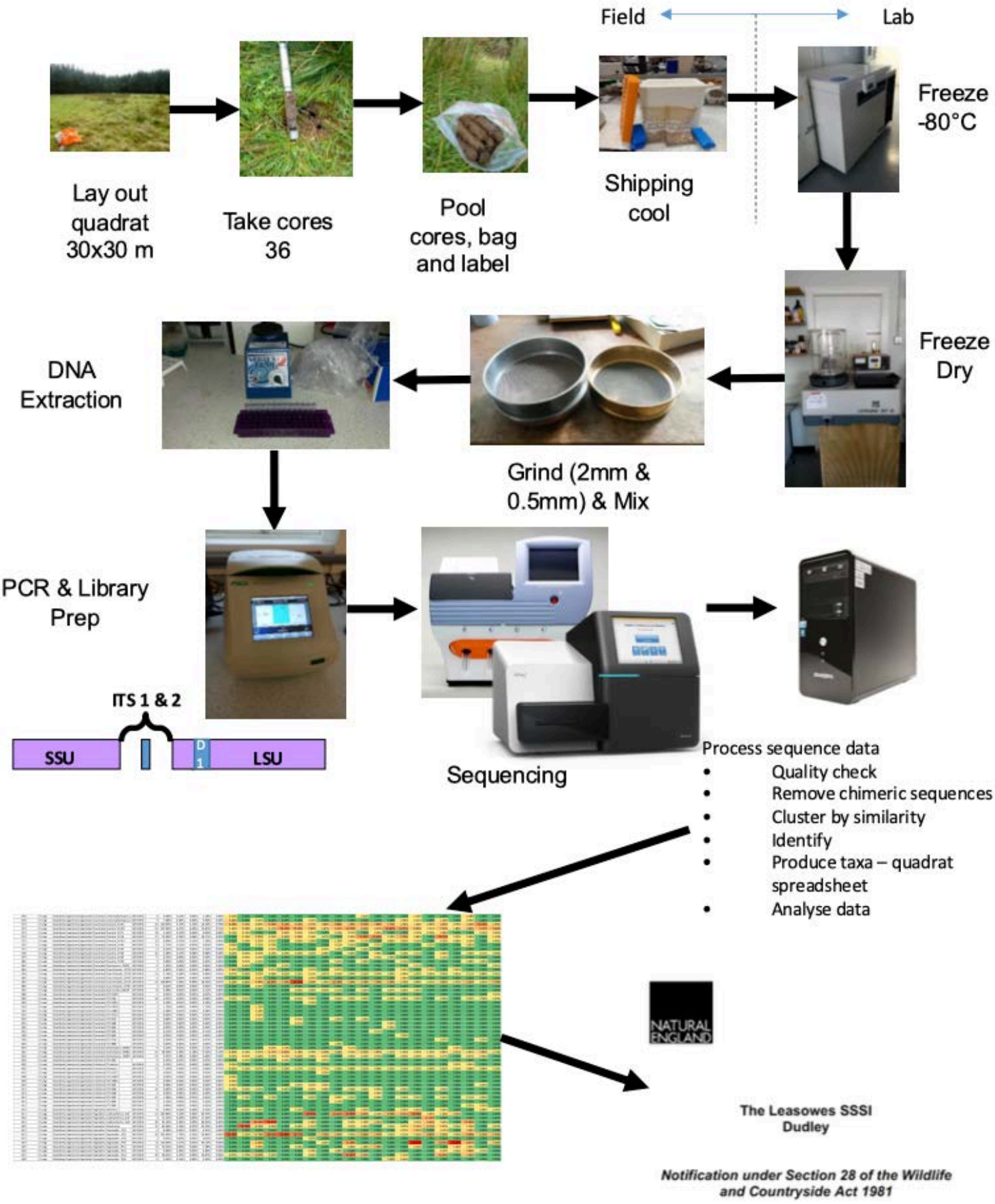

Figure 1 Flow diagram detailing the steps involved in identification of grassland fungi from soil eDNA. From Sampling to report.

The shape and extent of mycelial networks of soil macrofungi are largely unknown. However, many of these fungi can form fairy rings which provide an indication of the distribution of fungal mycelia within the soil. Most obvious are the fungi which form fairy rings where growth of the overlying vegetation is inhibited (type1; Figs. 2A/B; mostly likely by changes to soil hydrology due to secretion of hydrophobic proteins by the fungus) or enhanced (type2; Fig. 2C; likely due to release of nutrients due to the decomposition activity of the mycelia). These vegetational differences are often only seasonally visible and sometimes can show as depressions (Fig. 2B) but they have allowed measurement of ring diameters and growth rates which suggest that fairy rings are long-lived, in some cases 200-700 years (Shantz and Piemeisel, 1917). Additionally, many other macrofungi form type 3 fairy rings with no visible impact on vegetation (Figs. 2D/E/F), and this group includes many of the CHEGD fungi, the prime focus of conservation interest (Griffith et al., 2014). Whilst the mode of nutrition of type 1 and type2 (saprotrophic) is different from the CHEGD fungi (root-associated , likely mycorrhizal) (Halbwachs et al., 2018), examination of the distribution of fairy rings over larger areas of grassland (Figs. 2G/H) provides a better indication of the distribution of the mycelia than does the occurrence of basidiocarps (Miller and Gongloff, 2021; Ramsbottom, 1953), as illustrated by the comparison of the distribution of *Agaricus campestris* mushrooms and its mycelium (Fig. 2I) (Griffith and Roderick, 2008).Thus for type1/2 ring-forming fungi growth is clearly annular, with die-off of mycelia behind the advancing mycelial front and successional processes occurring radially (Edwards, 1988; Zotti et al., 2020). For type3 ring-forming species, a similar growth pattern can be inferred but it has not been confirmed experimentally that they grow in this manner.

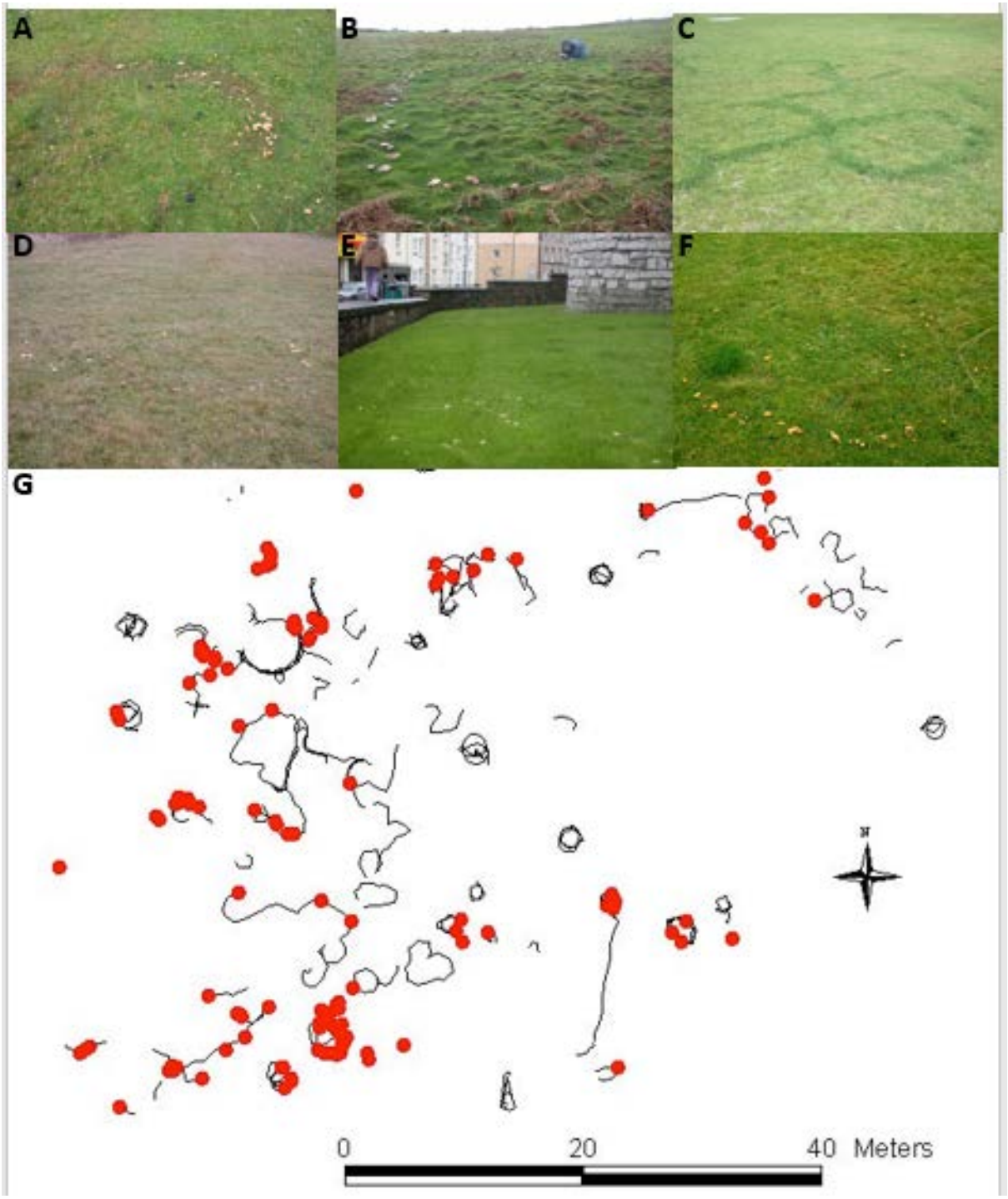

Figure 2. Fungal fairy rings Many grassland macrofungi, notably Marasmius oreades (A) and Lepista spp. (B), form 'type 1' rings, where grass growth may be inhibited (A) but not at all times of year (B), whilst others, for instance Agaricus campestris (C), form 'type 2' rings where grass growth is enhanced (C). Less obvious are the 'type 3' rings formed by many CHEGD fungi, including Cuphophyllus pratensis (D), C. virgineus (E), Hygrocybe chlorophana (F), where no effect on the vegetation is seen (Griffith et al., 2014-Supp2). The visibility of fairy rings is affected by many factors but at landscape scale in undisturbed grasslands, their abundance can be impressive. Type 2 fairy rings (G) A. campestris; Griffith & Roderick, 2008) provide an indication of the extent and heterogeneous distribution of macrofungal mycelia (black lines) relative to basidiocarps (red dots) within grassland soils. Images courtesy of GW Griffith**.**

There is no standardised procedure for sampling soil for eDNA metabarcoding of fungi across different habitats/ecosystems. Different studies adopt different procedures depending on their aims, illustrated here with three contrasting examples: In our 2016 investigation of an arable rotation study, with 90m^2 field plots (7.5x12 m), we took fourteen 17mm cores per plot, whereas for the Tedersoo et al. global soil study (2014) 40 cores (each 50 mm diameter, 50 mm deep) were pooled from a 2500m^2 quadrat, each core taken 1.5 m from the base of a tree (2 cores/tree) ensuring sampled trees were >8m apart within the quadrat. Conversely, for their CEH-funded Wales-wide soil study, George at al. (2019) took a single 150 deep x 40mm diameter soil core at each of the ca. 300 sampling points. One limiting factor that constrains the total weight of soil taken per samples is the ease of storage during transport (e.g. prior to transport; Tedersoo et al. (2014) air-dried their samples whilst George et al. (2019) homogenised frozen-thawed samples through a sieve). In our studies we always freeze-dry and this method avoids any degradation of soil eDNA during initial processing (Weißbecker et al., 2017), whilst permitting very fine grinding and mixing prior to subsampling of the ca. 200mg generally used for DNA extraction (see section 2.2).

To capture exactly the full fungal community profile of an area would require a high level of sampling and high sequencing depth, requiring a lot of effort and expense. At the other end of the scale would be a single sample, which is very unlikely to capture the full range of diversity. Somewhere in between lies a sampling intensity that is practically possible and is likely to capture the majority of species present and especially those important for ecosystem function and of conservation importance.

To test the effect of sampling intensity on the result of metabarcoding analyses we sampled a grassland area of 0.8 ha at a site 8 km south of Aberystwyth known as Mynachdy'r Graig (Fig. 3; unpublished data). Three nested areas were laid out with tapes (Fig. 3). Area A is the whole field (100x80 m) and was sampled every 10m for a total of 99 cores. Area B (30x30 m) was sampled every 5 m for a total of 49 cores. Area C (5x5 m) was sampled every 1 m for a total of 36 cores. After each core was taken the auger was cleaned with a metal spatula to remove soil and each core was bagged separately and frozen the day of collection at -80°C before freeze drying. After freeze drying in the laboratory, each core was ground separately through a 0.5 mm soil sieve and the ground soil thoroughly homogenised before storage at -80°C. The sieve was cleaned between each sample with a wire brush.

Before DNA extraction subsamples from the homogenised individual cores where combined in various proportions to represent different sampling efforts (Table 1; Fig. 4).

All the quadrats show that a higher sampling intensity returns a higher level of diversity. Interestingly the highest diversity from the 30x30 m (Operational Taxonomic Unit [OTU] count = 340) quadrat is very similar to that of the 100x80 m quadrat (OTU count = 340), indicating that within the field the smaller sampling area captures most of the diversity. The 5x5 m quadrat however captures a lower level of diversity even at the highest sampling density with just 1 m between cores. From these data we have established a "standard" 30x30 m quadrat from which we take 36 cores in a 6x6 grid, which takes approximately 40 minutes for one person to sample with a gouge auger. This is a slightly lower level of

sampling than highest test here, but reduces the sampling time to well under 1 hour, with little if any loss of diversity capture (see above). Over a large area we may take several quadrats to cover variations in topography or habitat.

Figure 3. Ground layout of the nested sampling areas at Mynachdy’r Graig.

Table 1. Mynachdy’r Graig sampling intensity and OTU counts for the different sized nested quadrats

Sample Name	Area	Number of Cores	Distance between Cores	Fungal OTU Count	CHEGD grassland OTUs
A10	A	99	10m	342	60
A20	A	30	20m	336	48
A30	A	12	30m	183	34
A40	A	9	40m	210	37
A50	A	6	50m	268	38
A100	A	4	Max	170	15
B5	B	49	5m	340	60
B10	B	16	10m	295	50
B30	B	4	Max	279	54
C1	C	36	1m	280	52
C2	C	9	2m	191	37
C5	C	4	Max	169	28

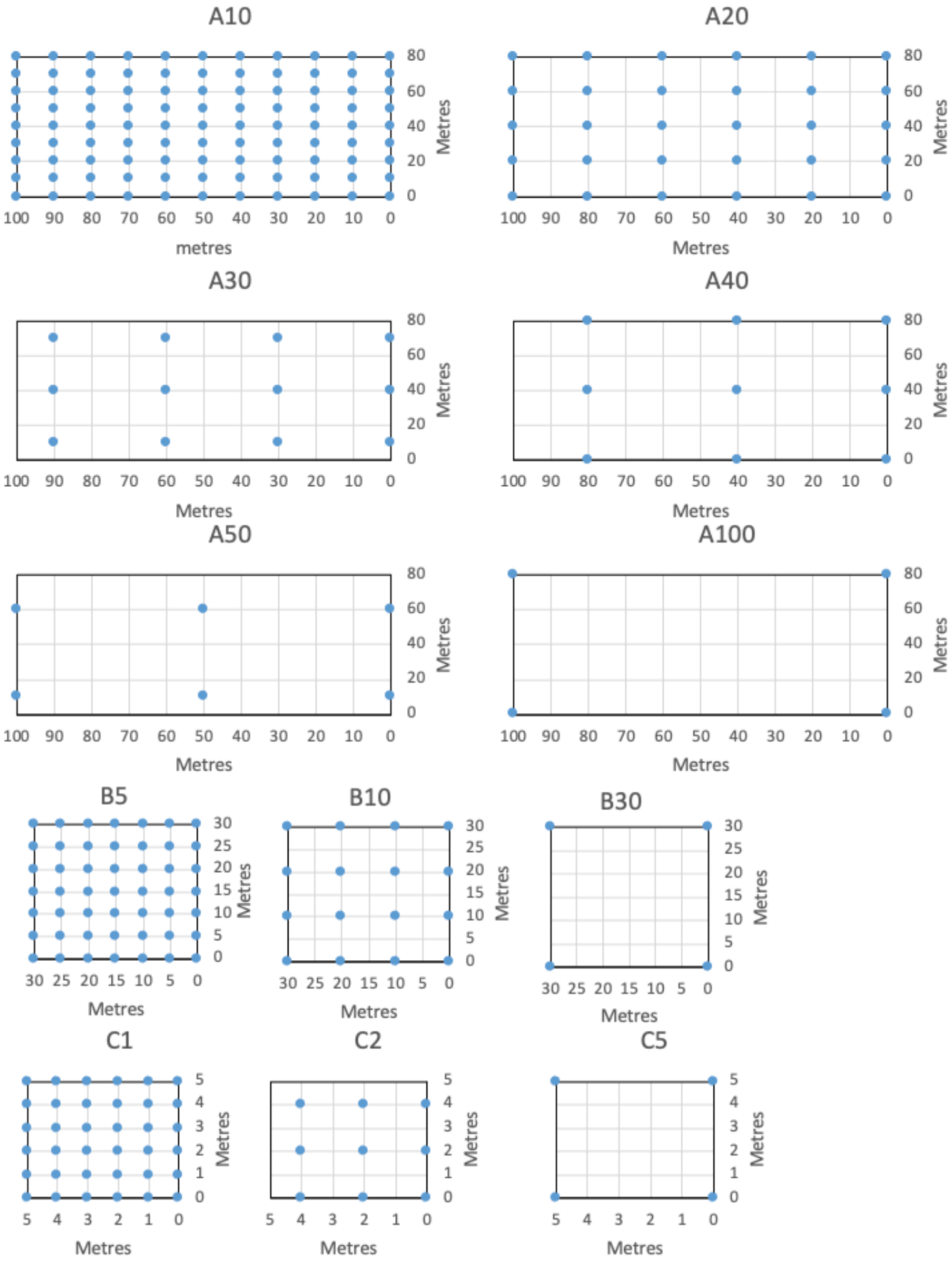

Figure 4. Core density maps of the three nested sampling areas A) 100 x 80 m. B) 30 x 30m and C) 5 x 5 m

When sampling 30x30 m quadrats in the field cleaning the auger between each core is unnecessary, with each core being pushed from the auger directly into a ziplock bag using the handle of a teaspoon or 5 ml plastic pipette tip. Cleaning is necessary between quadrats to avoid cross-contamination; this is done by scraping as much residual soil from the auger as possible using the spoon handle/pipette tip prior to wiping clean with a damp cloth (inside and out) until visible traces of soil are removed. Some cross contamination is possible but any tiny traces that evade visual inspection would not be detectable within the ca. 700 g sample from the next quadrat. However, if switching from a high biomass organic soil to low biomass soil, such as sand dune, here any cross contamination must be avoided by washing with soapy water followed by clean water and drying with a clean tissue, or alternatively a separate clean auger can be used. All soapy water must be retained and removed from site and disposed of in the sewage network. For the same reasons, gloves are not required, unless working on a low biomass soil. Care must be taken moving between sites to meet biosecurity requirements (e.g. clean footwear), as noted at http://www.nonnativespecies.org/checkcleandry/ biosecurity-for-everyone.cfm.

Sampling bags should be zip lock and the air removed as much as possible after sampling. In addition, they need to be able to withstand freezing at -80°C, at least 200 gauge (50 µm thick; e.g. Polybags Ltd.; PAA4; 22.5 x 32 cm; www.polybags.co.uk/shop/grip-seal-bag-wop_p575.htm) is recommended but thicker for very large (> 1 kg) samples. Bags with a write-on panel are also preferred and should be clearly labelled with an indelible marker with sample code (up to 4 letter abbreviation of site and a sample number) and date. A small piece of paper with sample code and sampling date written in pencil should also be placed inside the bag. A spreadsheet should accompany the samples with the codes and metadata such as full site description (See Appendix 1-blank proforma).

A particular advantage of eDNA surveys over traditional fruitbody surveys is that eDNA can detect species when they are not fruiting, especially important for species that are believed to fruit only very occasionally (e.g. *Hygrocybe spadicea*). The question remains however as to the best time to survey, due to changes in the abundance or distribution of mycelia. At Mynachdy'r Graig (Ceredigion; 52.347N,-4.118W) the 30x30 m quadrat was sampled monthly for 1 year. The Hygrophoraceae results from the sequencing of these samples is shown in Table 2. Of the 19 species found 10 were found in all monthly samples, 5 in more than 6 samples and 4 species in 6 or fewer samples. The lowest number of species found was 13 in August and December and the most (18) in April which also returned the highest proportion of Hygrophoraceae. No clear best time emerges from these data, the most abundant species tend to appear in all samplings and the least diverse retrieved 68% of total waxcap diversity. The variability in the data is most probably due to variability in the mycelial networks being cored but this relatively small annual variation in mycelial abundance is consistent with the fact that all these organisms are long-lived (decadal timescales).

Possibly the nationally rarest fungi at Mynachdy'r Graig are the "ballerina" waxcap (*Porpolomopsis calyptriformis)* and the "olive earthtongue" (*Microglossum olivaceum*). These were found in the metabarcoding dataset in the highest density samplings in area A

and area B. In addtion *P. calyptriformis* was found in 8 of the monthly samples whereas *M. olivaceum,* was found in only 4. This is likely due to the lower relative sequence abundance of *M. olivaceum* (average <1%) and hence lower biomass meaning the soil sampling is more likely to miss the mycelial network. As noted above, it is suspected that basidiomycete soil fungi grow in an annular manner (as a fairy ring) but for soil-inhabiting ascomycetes (e.g. *Microglossum*), the nature, extent and longevity of the mycelial systems is even less well-understood.

The aims of a study will affect some of the choices that are made in designing the sampling strategy. For example if the aim is to recover the highest level of diversity in a relatively small area then the sampling strategy may be more intensive than 36 cores we recommend here, which we feel is a good compromise between time/effort and diversity recovery. As we have seen relative abundance of species can vary quite widely between different sampling times and so this data does need to be treated with some caution.

2.2 Shipping and soil sample preparation.

Soil sampling with an auger damages fungal networks and over time the species composition may change in a sample (Clasen et al., 2020). Fast-growing fungi and bacteria will proliferate and degrade the DNA of the organisms initially present unless the sample is stabilised. Therefore the next part of the process is to arrest those changes prior to DNA sampling taking place. Ideally the sample would be frozen at -80°C, within a few hours of it being taken and when completely frozen (1-2 days) processed through a laboratory freeze dryer, which removes water from frozen samples without it going through the liquid phase via sublimation at low pressure. This stabilises the soil in an inactive dry state, without a period of warm drying where fungal communities could change.

We have recently published a detailed study of the effect of different soil storage conditions on the resultant eDNA metabarcoding data. (Clasen et al., 2020). These data showed only small (statistically non-significant) changes in the fungal and plant eDNA later detected from soil samples stored at 4°C for up to 14d. However, frozen samples when allowed to thaw degraded much more rapidly, with air-drying also causing changes in fungal communities. Thus, shipping is best achieved by placing samples in a coolbox/fridge as soon as possible after coring and posting in a bubble-wrapped envelope alongside a frozen coolblock. Under such conditions, 24hr courier service (costing <£10 for 1kg package) would permit the sample to reach the processing lab in good condition, allowing a margin of several days for any potential delays. We also identified particular fungi that proliferated only in poorly stored soil samples, so the presence of these in downstream analysis could provide indication of suspected sample degradation in transit (Clasen et al., 2020)**.**

Once the sample has been frozen and freeze dried the next step is homogenisation. This is another important and neglected step, required because the small sample taken for DNA extraction needs to be representative of the whole. Freeze drying the sample

facilitates the homogenisation as the soil becomes very friable. Our chosen method is to pass the whole sample through a 2 mm soil sieve to remove stones and coarsely grind the soil. This is then mixed thoroughly for 30 seconds by gloved hand and a subsample of approximately 50 g passed through a finer 0.5 mm soil sieve. The remainder of the 2 mm fraction is retained for chemical analyses, such as pH, carbon and available nutrients. Sieves should be cleaned with a wire brush and wiped with a clean tissue between samples; for low biomass samples sieves should be washed with 10% (non-sticky) bleach, rinsed clean and dried between samples. Additional citric acid cleaning (destroys DNA by acid hydrolysis but safe and not corrosive to other materials) of sieves is recommended before and after each batch of soils, but the potential for contamination between samples is extremely low if soil particles from the previous sample have been removed, as the amount of material being ground massively outweighs any potential contamination. The most critical stages for cross contamination are DNA extract and PCR and this is dealt with under those sections. Proprietary soil sieves are recommended for this purpose and are now available for ca. £10 from China (via eBay).

The fine-ground soil is stored at -80°C before DNA extraction and subsequently stored at that temperature as archive material, this ensures that both DNA and RNA are preserved for subsequent analyses. It is worthwhile storing the sample in such a way as to preserve RNA even if not extracted at this stage as it allows for future analyses. Analysis of rRNA (i.e. the transcript rather than the gene which encodes the rRNA) is possible via eDNA metabarcoding (Blazewicz et al., 2013), as detailed below (section 2.4), but procedures for RNA extraction and analysis are more time-consuming/expensive.

2.3 DNA extraction

If the sample has been thoroughly homogenised a single DNA extraction should be representative of the whole. A single extraction reduces the cost of sequencing per sample in terms of both consumables and laboratory time. Multiple extractions from the same sample give very similar results (unpublished data). In addition, as the sampling method is standardised, multiple samplings and extractions should also give similar results. For each DNA extraction procedure, a blank DNA extraction is used which is treated exactly the same but without soil added. This extraction is then added to the following PCR (see section 2.5).

To test that multiple extractions are not necessary, we sampled the Brignant long-term grazing experimental field site (lat/long: 52.3648°N, 3.8214°W; 367 m a.s.l.) near Aberystwyth Wales (Fig. 5), over the course of a full year from 23/04/2015 to 12/05/2016. Here we present data from two of the treatments. Treatment 1 is fertilised and grazed, with fertiliser (60 kg N ha^{-1} and 30 kg P ha^{-1}) applied in May. Treatment 6 is unfertilised with a spring hay cut and aftermath grazing. Each treatment is replicated in 3 blocks (Fig. 5). Each sampling was conducted as mentioned above in a 6x6 grid and the samples were treated as detailed above. Analysis of the sequence data from each sample through a statistical ordination technique known as principal coordinate analysis (PCO) shows a clustering by treatment and also by block, showing how reproducible the results of the

sampling and sample preparation methods are (Fig. 6). From this we conclude that multiple DNA extractions are unnecessary (unpublished data).

One factor that influences the reproducibility of multiple DNA extractions is the recovery of PCR inhibitor-free DNA, which can be influenced by the extraction method employed. Soils are a complex, PCR inhibitor-rich, medium and obtaining consistently clean DNA, free from contaminating inhibitors, in sufficient quantity can be problematical (especially from peaty soils with high organic matter content). Commercial extraction methods can be expensive, but can be cost effective if they save time by providing consistently usable DNA. We use the Powersoil kit (Qiagen), by far the most widely soil DNA extraction kit, costing ca. £5 per sample (£538 per 100) which consistently provides clean DNA that is suitable for downstream applications like PCR and therefore reduces the need for repeated extractions and/or PCR. There were concerns in the user-community that the take-over of MoBio by Qiagen which led to rebranding of the kit might alter product performance but we have found no evidence that this is the case.

The potential for contamination and cross contamination is greatly increased when dealing with the small sample weights used for DNA extraction and so great care needs to be taken. All extractions need to take place under a laminar flow hood that has been decontaminated before use (U.V. for 5 mins and wipe down with ethanol) and filter tips must be used to remove potential contaminants from inside the pipette. In addition to labcoat/gloves, wearing of a shower cap to avoid sample contamination by skin flakes containing *Malassezia* (dandruff fungus) is recommended.

Figure 5 Aerial photograph of the field experiment at Brignant, with the sampled areas labelled.

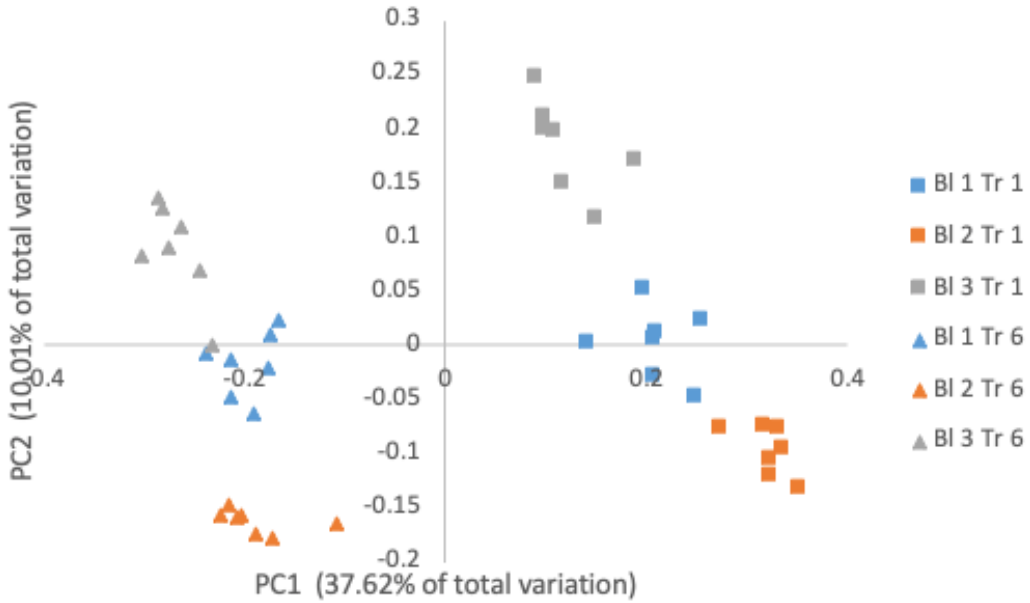

Figure 6 Principal coordinate ordination of the fungal sequencing data from Brignant from seven sampling times. Data clusters by treatment but also by block highlighting the reproducibility of multiple samplings and DNA extractions. The two primary axes account for 47% of the variability in the data.

2.4 Target region selection

There are many factors that need to be considered when choosing a region of the genome for metabarcode sequencing and choice will be influenced by the intended application and the required level of taxonomic discrimination. The ribosomal RNA (rRNA) operon was the first locus to be used for molecular phylogenetics, almost 50 years ago (Woese and Fox, 1977; Woese et al., 1990) and widely used for the identification of fungi since the 1990's (White et al., 1990a). Only much later was the term DNA barcode coined by the zoologist Paul Hebert (2003). The rRNA operon consists of three genes, the small subunit (SSU; 16S/18S) the 5.8S subunit and the large subunit (LSU; 26S/28S); these are transcribed into RNA as constituent parts of the ribosome separated by two intronic spacer regions that are transcribed but subsequently excised from the mature transcript (internal transcribed spacers (ITS) 1 and 2)(Fig. 7).

The two ITS spacer regions along with the 5.8S, as originally suggested by White et al. (1990a), have been accepted by the Consortium for the Barcode of Life (CBOL) as the primary barcode for fungi (Schoch et al., 2012), due to the sufficient variability in the ITS1 & 2 regions to discriminate between species (the ITS regions have high evolutionary rates (Begerow et al., 2010)), ease of amplification due to multiple copies and conserved flanking regions in the SSU and LSU regions for PCR priming sites. As a result of this adoption by CBOL a large collection of sequences has developed on GenBank, generally sequenced using the ITS1F (Gardes and Bruns, 1993) primer in the SSU with the ITS4 (White et al., 1990b) primer in the LSU. The taxonomic reliability of public databases such as GenBank, can be poor since uploaded sequences are not curated by taxonomic experts (Kõljalg et al., 2013; Meiklejohn et al., 2019). This highlights the need for curated databases with names assigned to sequences and clusters of sequences where reference sequences are available. Best known of the curated databases is the Barcode of Life Data Systems (BOLD) (Meiklejohn et al., 2019) but this database is highly animal-focused and not widely used by mycologists, since it does not address certain taxonomic issues important in mycology. The UNITE database (Abarenkov et al., 2010; Kõljalg et al., 2013) was developed specifically for ITS fungal sequences (Nilsson et al., 2019) although it is now expanding to include all Eukaryotic taxa. UNITE groups sequences into sequence hypotheses (SHs) at different levels of similarity (99.5% to 97% at 0.5% intervals) which broadly correspond to species. Species hypotheses are assigned taxonomies by mycologists with expertise in particular groups. There are currently 120,447 fungal SHs at a 1.5% distance and 565,915 ITS2 sequences.

The database of LSU sequences is much smaller for fungi than that of the ITS dataset. The Ribosomal Database Project (RDP) released a database of 62,860 fungal LSU sequences in 2013 (Cole et al., 2014). Since the release we have updated and expanded our copy of this database with recent changes to taxonomic classifications and also supplemented it with barcodes we have collected from grassland ecosystems especially within the CHEGD group of fungi, giving very good taxonomic coverage of this group.

As a result of this expanding database of published ITS sequences, the ITS region has become the region of choice for metabarcoding studies from environmental DNA (eDNA).

However, there were additional issues that need to be taken into consideration when developing a metabarcoding protocol over a more straightforward barcoding protocol. The first problem we addressed was that the full ITS1, 5.8S and ITS2 sequence was too long for much of the high-throughput sequencing (HTS) technologies and so a choice between ITS1 and ITS2 needed to be made. Fortunately the 5.8S conserved region sits between the two allowing for the development of further primer sequences. Given that suitable primers exist for amplification of ITS1 and ITS2 independently, the question remains as to what region is better suited to a metabarcoding analysis.

On average, for fungal species, the variability of ITS1 exceeds that of ITS2, although there are exceptions and about 34% of species have more variability in ITS2 (Nilsson et al., 2008). This would indicate that in terms of sequence diversity alone the ITS1 region would be the better target. This is not always borne out by comparative analyses, for example Bazzicalupo et al. (2013), found more variability in an ITS2 dataset of leaf-associated fungal communities, whereas Monard et al. (2013) found that sequencing of the two regions led to similar results at the fungal community level. In contrast an in silico analysis (Bellemain et al., 2010) recommended the use of ITS1 over ITS2 to avoid some of the length variability in ITS2 which biases against some of the longer Basidiomycota sequences. However, Nilsson et al. (2009) recommend the use of ITS2 as a metabarcoding region because of the extra taxonomic resolution provided by the start of the LSU region included in the ITS2 sequence and a more comprehensive database of ITS2 sequences. A more in depth analysis of multiple primer pairs covering variable regions of the SSU and LSU as well as ITS (Tedersoo et al., 2015), concluded that all have biases for some groups of fungi but through the use of multiple forward primer combinations (Table 3) and a slightly degenerate reverse primer based on the ITS4 primer, these can for the most part be overcome for ITS2.

The analysis of Tedersoo et al. (2015) also indicate that, despite lower taxonomic recovery, the variable regions of the LSU (D1 & D2) were more efficient in identifying trends in fungal composition due to environmental gradients. Another advantage of the LSU variable regions is their constrained length when compared to the much more variable ITS1/2. Longer amplicons are biased against by PCR as shorter amplicons amplify much more readily and may be missed altogether by the sequencing technology if the length exceeds the sequencing limit. These advantages suggest that for some applications, one or more of the LSU variable regions would give complementary information about fungal communities to data generated from ITS2 sequencing. From the analysis of Tedersoo et al. (2015) the D1 variable region gives the most taxonomic resolution. At Aberystwyth we have developed primers that amplify the D1 region of the LSU across the taxonomic range of the fungal kingdom (Table 3). This was initially in response to the technology available to us at the time (200 bp sequencing), but has subsequently allowed us to compare datasets from ITS2 and LSU.

The variability in the D1 region is sufficient within the “waxcap” (Hygrophoraceae) and the “fairy club” (Clavariaceae) fungi to discriminate to species level and our database has been updated to reflect this. This gives us the ability to use this region to survey

grasslands, one of the foremost ecosystems in Wales and a nationally important habitat for waxcap fungi (Griffith et al., 2013).

In 2019, English Heritage permitted us to sample the lawns and meadow at Down House. On December 31st 2019 we sampled three areas known as the 'fungal lawn', the 'orchid lawn' and the 'meadow' with 25 cores from the fungal lawn and the meadow and 16 from the orchid lawn (Fig. 8). These samples were prepared as outlined above and then amplified and sequenced using the LSU primers as above and also the primer mix for ITS2 as recommended by Tedersoo et al. (2014). Both datasets were dominated by the Dikarya (Ascomycota/Basidiomycota) but one striking difference was the increased proportion of the Ascomycota in the ITS2 dataset (Table 4).

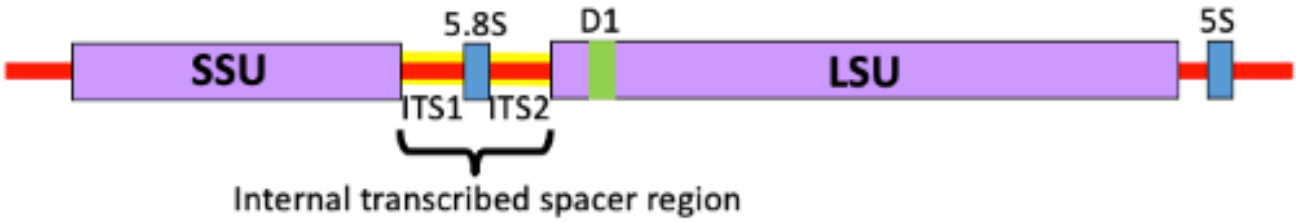

Figure 7 Schematic diagram of the ribosomal RNA operon, showing the locations of the small (SSU; 18S; ca. 1850 bp) and large (LSU; ca. 4800 bp) ribosomal subunit genes. Two smaller rRNA subunits (5.8S and 5S) are also shown. The internal transcribed spacer region is shown in yellow. This intronic region is spliced out during maturation of the transcript and since these two regions do not contribute to ribosome function they are under reduced selective pressure, allowing more rapid evolution. Thus levels of taxonomic variation in the ITS spacers are greater than in adjacent regions of the operon which are more evolutionarily constrained. The original primers for PCR amplification of the ITS region (White et al., 1990) were designed based on conserved regions of the adjacent SSU and LSU genes. The length of the ITS1/2 PCR amplicon (600-700 bp) was well suited for Sanger-based dideoxy DNA sequencing. For eDNA metabarcoding, however, shorter amplicons are generally used (3 – 500bp) and amplification of the ITS region (using conserved flanking primers in the 5.8S and LSU region) are widely used. Also shown is the D1 variable region of the LSU, which also contains sufficient variation to allow species-level discrimination of most fungi, including CHEGD taxa.

Figure 8 Sampling locations at Down House.

Table 2. Relative abundances of Hygrophoraceae species from monthly sampling at Mynachdy'r Graig

Species	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Count	Min	Max
Cuphophyllus pratensis	0.36%	5.52%	16.18%	24.57%	13.23%	4.99%	14.79%	6.52%	7.74%	15.75%	5.63%	16.90%	12	0.36%	24.57%
Cuphophyllus russocoriacea	2.24%	0.07%	4.78%	1.69%	4.63%	4.97%	0.02%	0.00%	0.04%	0.05%	2.46%	1.51%	12	0.00%	4.97%
Cuphophyllus virgineus	5.68%	0.33%	9.19%	3.35%	1.12%	1.75%	0.08%	1.45%	3.65%	3.63%	5.39%	4.00%	12	0.08%	9.19%
Gliophorus irrigatus	0.84%	0.25%	0.63%	0.36%	0.20%	0.01%	0.41%	0.15%	3.00%	0.52%	1.09%	1.26%	12	0.01%	3.00%
Gliophorus psittacinus	0.67%	0.00%	1.71%	0.20%	7.33%	1.16%	0.54%	0.04%	0.32%	0.15%	0.07%	3.14%	11	0.00%	7.33%
Hygrocybe cantharellus AFF	0.00%	0.00%	0.27%	0.00%	0.07%	0.01%	0.01%	0.04%	0.00%	0.00%	0.00%	0.31%	6	0.00%	0.31%
Hygrocybe ceracea	1.28%	1.98%	1.52%	0.92%	0.40%	0.83%	0.08%	0.33%	0.74%	0.28%	1.66%	0.07%	12	0.07%	1.98%
Hygrocybe cf. acutoconica	5.39%	10.65%	3.09%	3.28%	1.72%	2.99%	0.00%	0.00%	0.00%	3.65%	4.29%	0.00%	8	0.00%	10.65%
Hygrocybe chlorophana	4.12%	1.91%	0.95%	0.01%	3.01%	9.24%	0.50%	1.89%	2.73%	8.96%	12.37%	3.43%	12	0.01%	12.37%
Hygrocybe citrinovirens	1.69%	0.00%	0.00%	4.87%	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3	0.00%	4.87%
Hygrocybe coccinea	15.87%	1.05%	2.38%	2.09%	2.45%	2.23%	10.67%	12.20%	10.33%	2.25%	7.88%	3.17%	12	1.05%	15.87%
Hygrocybe conica	0.00%	0.00%	0.00%	0.01%	0.22%	0.00%	0.02%	0.06%	0.08%	0.01%	0.21%	0.22%	8	0.00%	0.22%
Hygrocybe glutinipes	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2	0.00%	0.02%
Hygrocybe mucronella	0.09%	0.04%	0.71%	0.24%	0.44%	0.53%	0.12%	0.06%	2.23%	0.12%	2.58%	0.36%	12	0.04%	2.58%

Hygrocybe punicea	0.29%	0.50%	1.35%	12.59 %	4.73%	12.09 %	0.04%	3.24%	0.05%	6.31%	0.61%	0.31%	12	0.04 %	12.59 %
Hygrocybe quieta	3.45%	0.08%	4.04%	0.67%	6.75%	4.71%	5.73%	1.64%	1.78%	2.41%	4.93%	1.01%	12	0.08 %	6.75%
Hygrocybe reidii	0.00%	0.08%	0.83%	0.10%	0.00%	0.00%	0.72%	0.00%	0.00%	0.21%	2.13%	0.00%	6	0.00 %	2.13%
Hygrocybe spadicea	1.26%	1.98%	1.00%	0.90%	0.36%	0.53%	0.00%	0.00%	0.00%	0.07%	1.04%	0.00%	8	0.00 %	1.98%
Porpolomopsis calyptriformis	1.43%	0.06%	1.13%	0.33%	0.66%	0.65%	0.76%	0.00%	0.00%	0.12%	0.00%	0.00%	8	0.00 %	1.43%
% Hygrophoraceae	44.65 %	24.49 %	49.76 %	56.19 %	47.36 %	46.70 %	34.51 %	27.62 %	32.69 %	44.50 %	52.36 %	35.69 %			
No. Hygrophoraceae Species	15	14	16	18	17	15	15	13	12	17	15	13			

Table 3 ITS and LSU primer sequences

Amplicon	Type	Primer name	Primer sequence	Matching organisms
ITS2	Fwd	ITS3NGS1	CATCGATGAAGAACGCAG	Ca. 95% of all fungi
ITS2	Fwd	ITS3NGS2	CAACGATGAAGAACGCAG	Chytridiomycota
ITS2	Fwd	ITS3NGS3	CACCGATGAAGAACGCAG	Sebacinales p.parte
ITS2	Fwd	ITS3NGS4	CATCGATGAAGAACGTAG	Glomeromycota (also plant)
ITS2	Fwd	ITS3NGS5	CATCGATGAAGAACGTGG	Sordariales p.parte
ITS2	Fwd	ITS3NGS10	CATCGATGAAGAACGCT	G Stramenopila
ITS2	Rev	ITS4NGS	TCCTSCGCTTATTGATATGC	>99% fungi, plants, most protists
LSU	Fwd	D1F2	CYYAGTARCTGCGAGTGAAG	>99% fungi, some plants & protists
LSU	Rev	NLC2	GAGCWGCATTCCCAAACWA	>99% fungi, some plants & protists

On the fungal lawn both datasets were dominated by the Hygrophoraceae, 71.88% ITS2 and 76.92% LSU with *Hygrocybe punicea* the most dominant species, accounting for approximately 50% of total fungal sequence abundance in both datasets (Table 4). One difference in Hygrophoraceae species abundance between the two datasets is for *Porpolomopsis calyptriformis.* This species was found in the highest abundance on the Orchid lawn in the LSU dataset at 11% and at 3.7% on the fungal lawn. In contrast ITS2 abundances for this species were very much lower (1.84% and 0.52% respectively), indicating the relative long amplicon length of the ITS2 region for this species (407 bp between the priming sites) in comparison to *Hygrocybe punicea* (359 bp between the priming sites), was inhibiting either PCR amplification or sequencing. There were no primer mismatches for either species for either the ITS2 or LSU amplicons species. However, the LSU amplicons were very similar in length for both species (178bp and 181 bp respectively) (Table 4).

Another site where we have data for both ITS2 and LSU from the same extract is the Llanishen and Lisvane Reservoir Embankments SSSI in Cardiff (Fig. 9) (51.528N,-3.175W). This site has been designated a Site of Special Scientific Interest (SSSI) on account of the “waxcap” fungal species found on the embankments surrounding the reservoirs. It has been regularly surveyed for fungal fruitbodies since 2004 so we can also compare eDNA data to traditional fruitbody surveys. Twelve quadrats were sampled in November 2018 using a gouge auger. In each, a total of 36 soil cores were taken from just below the turf surface (each approximately 100x17 mm) at 5 m intervals. Table 5 shows the comparison for Hygrophoraceae using the two different amplicons. In terms of species there is a good concordance between the two datasets, with one species, *Hygrocybe splendidissima*, only present in the ITS2 dataset and four species (*Cuphophyllus fornicatus, Hygrocybe cantharellus, Hygrocybe insipida, Hygrocybe reidii*), at very low relative abundances and only present in the LSU dataset

The main disparity between the ITS2 and LSU datasets is the relative sequence abundance, with the proportion of Hygrophoraceae sequences in the LSU dataset being consistently higher. The ratio of ITS2:LSU abundance for Hygrophoraceae does vary between quadrats but only one species, *G. psittacinus*, had lower abundance in the LSU dataset than ITS2 (Table 5). Examination of the primer binding sites of published *G. psittacinus* sequences from GenBank reveals that there is a mismatch in LSU forward primer region for *G. psittacinus* sequences (position 11, primer: T, sequence: G) and this may therefore explain the low representation.

The accepted primary barcode for fungi is the ITS region, with the LSU region as a secondary barcode region (Schoch et al., 2012). For this reason there are more published barcode sequences for ITS for most taxa. Clavariaceae would be an unusual exception to this (Birkebak et al., 2016)). Thus no reference LSU barcodes are available for some species (e.g. *H. splendidissima* and *C. colemannianus*) but we are working to rectify this. The varieties of *C. virgineus* (e.g. vars. *virgineus*, *fuscescens* and *ochraceopallida* and *pratensis* (vars. *pratensis* and *ochraceopallida*) do not vary across the LSU region.

As mentioned above, the ITS2 sequence is variable in length and PCR is biased towards shorter amplicons, so a potential explanation of the variation in the ratio of ITS2:LSU abundances is that Hygrophoraceae ITS2 amplicons for Hygrophoraceae are longer that those of other fungi. To see if a relationship exists between these two factors, the average total ITS2 sequence length for all fungi and the average sequence length for Hygrophoraceae of each quadrat was calculated and the ratio plotted against the ratio of ITS2:LSU abundances (Fig. 10), with good correlation between these two variables (R = 0.86). When the average total fungal and Hygrophoraceae sequence lengths are closer (a ratio close to 100%), then the difference between the LSU and ITS2 abundances tends to be smaller. The explanation for this is that the shorter sequences are preferentially amplified and sequenced over the longer sequences, so when there are shorter sequences in the dataset, the longer Hygrophoraceae sequences tend to be lower in abundance.

Figure 9 Map of Llanishen reservoir showing positions of quadrats for soil sampling

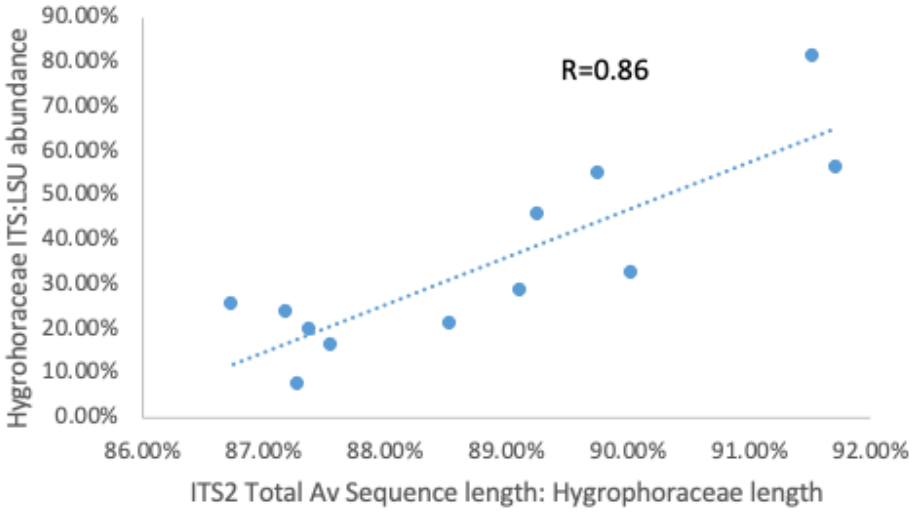

Figure 10. Relationship between the ratio of ITS to LSU Hygrophoraceae sequence abundance and the ratio of the average Hygrophoraceae ITS sequence length to total average ITS sequence length

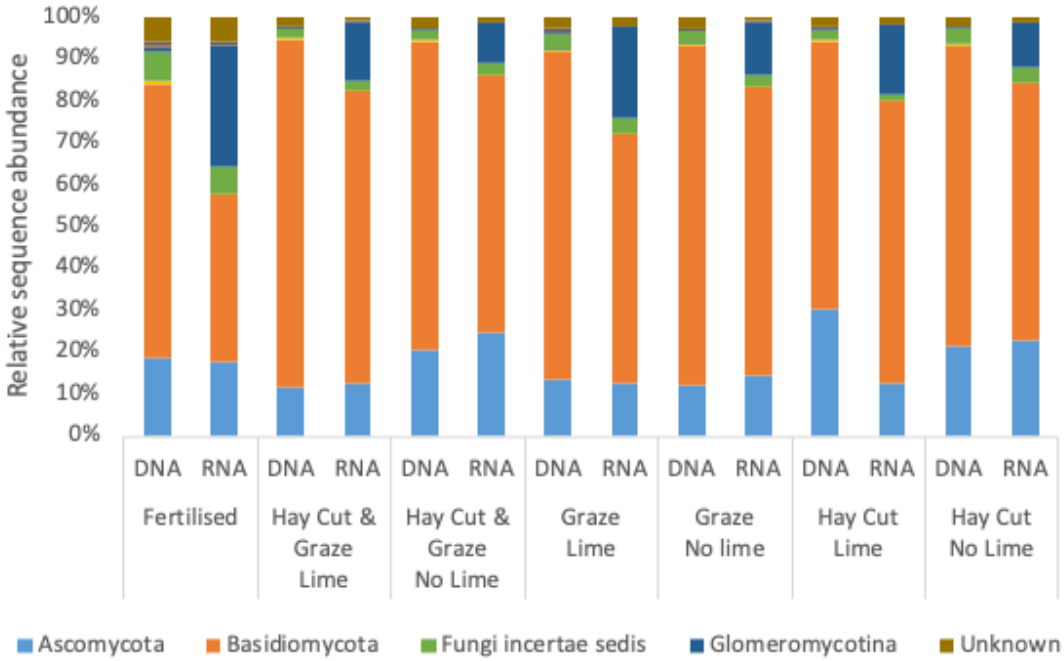

Figure 11 Stacked barcharts of the relative abundance of fungal phyla by treatment and nucleic acid type. Basidiomycota have lower relative abundances in RNA sequencing across all treatments, whereas the Glomeromycotina have higher abundances with RNA sequencing

Table 4. Dominant phyla (upper) and waxcap species (lower) from the sampled areas at Down House, showing relative abundance values obtained for the LSU and ITS datasets. Data are expressed as a percentage of total fungal reads

Taxon	Fungal Lawn		Meadow		Orchid Lawn	
	ITS2	LSU	ITS2	LSU	ITS2	LSU
Ascomycota	13.17%	6.00%	59.73%	40.08%	34.69%	22.63%
Basidiomycota	82.19%	92.63%	32.72%	50.81%	57.06%	72.36%
Geoglossaceae	**0.39%**	**0.15%**	**0.61%**	**0.41%**	**4.96%**	**2.64%**
Clavariaceae	2.07%	2.52%	18.09%	22.16%	30.29%	22.56%
Dermaloma	1.15%	1.23%	0.00%	0.00%	0.00%	0.00%
Entolomataceae	1.17%	2.06%	0.29%	0.74%	0.73%	0.99%
Hygrophoraceae	71.88%	76.92%	6.99%	14.37%	16.18%	26.90%

	Fungal Lawn		Meadow		Orchid Lawn	
	ITS2	LSU	ITS2	LSU	ITS2	LSU
Cuphophyllus pratensis	2.54%	2.42%	0.00%	0.00%	0.00%	0.00%
Cuphophyllus russocoriaceus	0.34%	0.44%	1.42%	2.09%	11.21%	11.59%
Gliophorus irrigatus	0.12%	0.63%	0.00%	0.00%	0.00%	0.00%
Gliophorus psittacinus	4.52%	0.74%	0.00%	0.00%	1.82%	0.56%
Hygrocybe chlorophana	0.01%	0.07%	0.34%	0.68%	0.00%	0.00%

	Fungal Lawn		Meadow		Orchid Lawn	
Hygrocybe conica	0.00%	1.43%	1.36%	5.91%	0.45%	1.20%
Hygrocybe citrinovirens	3.04%	3.67%	3.85%	6.20%	0.00%	0.00%
Hygrocybe intermedia	0.97%	5.06%	0.00%	0.01%	0.00%	0.01%
Hygrocybe coccinea	4.54%	9.87%	0.00%	0.01%	0.00%	0.01%
Hygrocybe punicea	55.28%	48.39%	0.03%	0.05%	0.02%	0.02%
Hygrocybe insipida	0.00%	1.92%	0.00%	0.00%	0.00%	0.00%
Hygrophorus	0.00%	0.00%	0.00%	0.00%	0.85%	1.40%
Porpolomopsis calyptriformis	0.52%	3.69%	0.00%	0.02%	1.84%	11.01%

Table 5. Hygrophoraceae species and relative sequence abundance at Llanishen for LSU and ITS amplicons

Species	Amp.	Quadrat											
		1	2	3	4	5	6	7	8	9	10	11	12
Cuphophyllus flavipes	LSU	10.78%	0.48%	1.12%	1.93%	0.02%	0.17%	0.00%	0.00%	0.00%	7.17%	0.02%	0.65%
	ITS	0.00%	0.11%	0.00%	0.12%	0.00%	0.06%	0.00%	0.00%	0.00%	0.31%	0.00%	0.00%
Cuphophyllus fornicatus	LSU	0.00%	0.00%	0.00%	0.00%	0.28%	0.05%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%
	ITS	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Cuphophyllus pratensis	LSU	1.20%	8.96%	4.06%	14.47%	0.02%	0.00%	0.00%	0.04%	0.00%	0.05%	0.00%	0.06%
	ITS	0.13%	1.92%	1.03%	8.31%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Cuphophyllus sp.	LSU	0.69%	0.12%	6.85%	6.77%	33.82%	11.59 %	28.95 %	0.06%	4.66%	0.77%	0.15%	0.47%
	ITS	0.02%	0.04%	2.81%	5.71%	8.78%	4.80%	17.41 %	0.00%	2.42%	0.18%	0.03%	0.00%

Gliophorus laetus	LSU	0.00%	0.00%	0.00%	1.63%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	ITS	0.00%	0.00%	0.00%	1.33%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Gliophorus psittacinus	LSU	0.00%	0.30%	0.00%	0.00%	0.00%	0.00%	0.00%	0.22%	0.12%	1.05%	0.00%	0.00%
	ITS	0.00%	1.49%	0.00%	0.00%	0.00%	0.00%	0.00%	0.92%	0.22%	2.32%	0.00%	0.00%
Gliophorus irrigatus	LSU	0.10%	0.00%	0.14%	0.01%	0.00%	0.00%	0.00%	0.89%	2.79%	1.38%	0.00%	0.00%
	ITS	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.29%	1.05%	0.53%	0.00%	0.00%
Hygrocybe acutoconica	ITS	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.69%	0.03%	0.00%	0.00%	0.00%	0.00%
	LSU	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Hygrocybe cantharellus	ITS	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%
	LSU	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Hygrocybe chlorophana	LSU	3.87%	0.19%	0.00%	1.15%	0.02%	0.00%	0.00%	0.00%	13.55%	0.00%	0.00%	0.00%

	ITS	0.64%	0.09%	0.00%	0.37%	0.00%	0.00%	0.00%	0.00%	7.67%	0.00%	0.00%	0.00%
Hygrocybe citrinovirens	LSU	3.21%	6.22%	9.55%	1.28%	0.01%	24.84 %	0.01%	21.82%	1.30%	18.17%	0.94%	2.12%
	ITS	0.36%	1.57%	2.11%	0.47%	0.00%	8.12%	0.00%	19.72%	0.43%	5.42%	0.09%	0.30%
Hygrocybe coccinea	LSU	0.05%	0.00%	0.53%	0.00%	0.01%	0.00%	0.00%	1.58%	5.47%	1.31%	14.64%	0.00%
	ITS	0.00%	0.00%	0.19%	0.00%	0.00%	0.00%	0.00%	0.70%	1.62%	0.40%	2.54%	0.00%
Hygrocybe conica	LSU	0.99%	0.90%	0.03%	0.00%	0.18%	2.06%	2.40%	0.35%	4.68%	0.05%	0.00%	2.15%
	ITS	0.09%	0.12%	0.00%	0.00%	0.00%	0.00%	0.29%	0.00%	2.01%	0.00%	0.00%	0.10%
Hygrocybe glutinipes	LSU	0.58%	0.00%	0.00%	13.52%	0.00%	0.00%	0.00%	0.00%	0.64%	0.00%	7.19%	11.45%
	ITS	0.13%	0.00%	0.00%	10.65%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.12%	2.49%
Hygrocybe insipida	ITS	0.47%	0.05%	0.00%	0.44%	0.00%	0.00%	0.13%	0.01%	0.00%	0.22%	1.28%	0.00%
	LSU	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Hygrocybe intermedia	LSU	15.62%	19.96%	15.87 %	18.75%	0.01%	26.00 %	2.60%	19.29%	6.72%	6.39%	0.00%	4.22%
	ITS	1.20%	3.48%	2.89%	9.00%	0.00%	9.30%	0.81%	12.75%	2.22%	1.82%	0.00%	0.67%
Hygrocybe mucronella	LSU	0.00%	0.70%	0.00%	0.00%	1.25%	3.82%	0.10%	0.02%	0.00%	0.00%	0.00%	0.00%
	ITS	0.00%	0.00%	0.00%	0.00%	0.24%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Hygrocybe punicea	LSU	2.43%	7.18%	0.01%	0.00%	0.04%	0.00%	0.00%	22.23%	0.25%	0.00%	0.00%	0.00%
	ITS	0.44%	0.63%	0.00%	0.00%	0.00%	0.00%	0.00%	20.34%	0.11%	0.00%	0.00%	0.00%
Hygrocybe quieta	LSU	0.09%	2.38%	0.01%	0.58%	0.00%	0.00%	0.74%	0.02%	1.15%	0.00%	0.01%	1.22%
	ITS	0.00%	0.83%	0.00%	0.31%	0.00%	0.00%	0.62%	0.00%	0.61%	0.00%	0.00%	0.16%
Hygrocybe reidii	LSU	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	ITS	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Hygrocybe splendidissima	LSU	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

	ITS	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.16%	0.00%	0.00%	0.00%
Neohygrocybe ingrata	LSU	0.00%	0.00%	0.00%	1.71%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	ITS	0.00%	0.00%	0.00%	1.62%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%
Porpolomopsis calyptriformis	LSU	0.21%	2.88%	0.00%	10.57%	0.01%	0.00%	0.00%	2.88%	2.46%	6.84%	1.87%	0.57%
	ITS	0.00%	0.28%	0.00%	2.84%	0.00%	0.00%	0.00%	1.57%	0.54%	1.32%	0.12%	0.05%
Total	LSU	39.82%	50.27%	38.18 %	72.38%	35.42%	68.47 %	34.80 %	69.39%	43.79%	43.19%	24.81%	22.92%
	ITS	2.99%	10.55%	9.02%	40.74%	9.02%	22.28 %	19.13 %	56.28%	20.06%	12.33%	4.91%	3.77%
ITS:LSU ratio		7.52%	20.99%	23.63 %	56.28%	25.46%	32.54 %	54.96 %	81.11%	45.80%	28.54%	19.77%	16.45%

The data above (Fig. 10, Tables 4&5) indicates that where a soil sample contains fungi with relatively long ITS2 sequences, then PCR will bias relative abundances to species with shorter lengths and so using the LSU region may give a more accurate picture of the actual fungal population and relative abundances, in some cases species could be missed altogether from the ITS2 dataset because of sequence length. For example, some *Boletus* species have a sequence length (between ITS3 and ITS4 priming sites) up to 580bp. There is bias against these amplicons not only during the PCR reaction (with shorter amplicons being amplified more efficiently) but potentially during sequencing, where such lengths are at the limits of many current high-throughput technologies. The sequencing of mock communities constructed from a mix of DNA extracts from dried mushrooms and pure culture, gave very different answers using LSU and ITS2 metabarcoding on the Ion Torrent (Hopwood et al., unpublished data) due to very poor representation of some *Boletus* spp. (with long ITS2 spacer regions) in ITS datasets. However, this problem is somewhat counter-balanced by the greater taxonomic resolution and better barcode coverage of the ITS2 region. The two regions provide complementary information but to analyse both loci for all samples would have significant cost implications. Therefore, the choice of locus needs to be based on the scientific question that is being asked with regard to the relative importance of precise identification vs, accurate quantification, and also the group of fungi being studied.

For grassland surveys where the main question is one of which species of conservation importance are present and in what approximate relative abundances (i.e. as % of the total fungal population), then the LSU is more appropriate, since it provides good species-level identification for CHEGD fungi. For other ecosystems, such as woodland soils, LSU may not provide the required taxonomic resolution and for some groups of fungi reference DNA barcode coverage is poorer. Thus for woodland soils, sequencing of the ITS2 region would be recommended in order to obtain accurate taxon identification but at the expense of accurate relative quantification of the different species present. The ITS2 region also has the advantage of having a good database coverage of plants/green algae (kingdom Plantae), so by including a plant specific primer in our mix we can also identify plant species from a soil sample, as we have demonstrated previously (Clasen et al., 2020; Detheridge et al., 2020). Future developments in high-throughput sequencing, for example via Oxford Nanopore's MinION device, may permit simultaneous sequencing of both ITS2 and LSU barcode loci within a single ca. 700bp amplicon (since they are adjacent).

The discussion above has focused only analysis of eDNA but it is also possible to undertake similar analyses for eRNA to provide different information about fungal communities in soil. The target locus for eDNA analyses is the rRNA operon whose transcripts later form part of the ribosome, 'protein factory' of the cell, where messenger RNA molecules (mRNA) are translated into proteins (LaRiviere et al., 2006). By specific extraction of total RNA from soil and use of the enzyme reverse transcriptase, it is possible to synthesize DNA from rRNA molecules (RT-PCR) and thereby quantify the number of ribosomes present in a cell (as opposed to the number of gene copies of the rRNA gene). Since active cells contain more ribosomes than cells in a resting state (e.g. spores), it is

possible to identify which fungal species are most active in the soil community, rather than which species are present at highest biomass.

The structure of the rRNA operon is described above in section 2.4 (Fig. 7). A single transcript is initially transcribed but the (intronic) ITS regions of the raw rRNA transcript are excised and degraded during transcript maturation (Houseley and Tollervey, 2009), with the remaining SSU/LSU/5S/5.8S portions being incorporated into ribosomes. Thus RNA of the ITS sequences are present only transiently wherease the rRNA elements contributing to ribosome structure are much more persistent (Houseley and Tollervey, 2009; Karnahl and Wasternack, 1992). Thus eRNA metabarcoding of the LSU transcript will provide an estimate of the relative abundance (i.e. activity) of different species within the soil, thereby identifying those species which contribute most to soil processes rather than those which are simply present as less active hyphae or resting structures. As noted by Blazewicz et al (2013), ribosomes and their constituent rRNA molecules may persist in dead cells which have lost viability but retain cellular integrity but the same is true of DNA molecules.

Our study of RNA:DNA relative abundance was undertaken at the 27 year-old Brignant long-term grass land extensification experiment (Aberystwyth; 52.365N,-3.831W; Fig 3), where the replicated treatments permitted comparison of different grassland management regimes. The site was sampled on the 12th May 2016 using the standard protocol outlined above. To amplify the RNA molecule, the first step after extraction is to reverse transcribe the RNA into complimentary DNA (cDNA) using the enzyme reverse transcriptase, as the enzyme used in PCR can only amplify using DNA as a template. The reverse transcription was carried out using the LSU specific reverse primer extended at the 5' end with an adaptor sequence (P1). The PCR was then performed with the forward LSU primer and P1, this ensures that only the cDNA created by the reverse transcription step is amplified.

There are seven treatments in total at Brignant, the fertilised and grazed treatment described above, then six non fertilised treatment: hay cutting only plus and minus lime; hay cutting with aftermath grazing plus and minus lime; grazing only plus and minus lime. Data for RNA and DNA sequencing at phylum level are shown as stacked bars (Fig. 11). The clear difference between the two datasets is the increased relative abundance of the arbuscular mycorrhizal fungi (AMF) of the sub phylum Glomeromycotina from RNA data and the reduced relative abundance of the Basidiomycota, this result is most pronounced on the fertilised plots (Basidiomycota DNA 65.09% ± 21.33%; RNA 39.88% ± 10.63%: Glomeromycotina DNA 1.18% ± 1.14%; RNA 29.1% ± 4.52%). The Hygrophoraceae are most dominant on the hay cut with aftermath grazing without lime, 42.38% ± 14.73%, which falls to 19.05% ± 10.16% relative sequence abundance in the RNA data, whereas the total Basidiomycota relative abundance on this same treatment falls from 73.74% ± 14.36% to 61.62% ± 7.4%.

Sequence data comparing activity levels suggests that in early spring the AMF are much more active than the slower-growing basidiomycetes especially the Hygrophoraceae, and provides further detail on the ecology of different fungal groups (Detheridge et al., in prep). It would be interesting to compare these data to other seasons to see if activity levels change seasonally, since some fungi may be active at only certain time of year; for such

fungi changes in abundance not observed in temporal patterns for eDNA could be elucidated.

It is possible to quantify the abundance of the ITS regions transiently present in the raw rRNA transcript but this would quantify the rate of ribosome synthesis rather than the number of active ribosomes, correlating more closely with mycelial growth rate rather than activity (since cells may be metabolically active but not growing). This type of analysis was performed recently (Adamo et al., 2021) on different ecosystems, grassland soil, woodland soil and decaying wood from four different sites in northern Italy, sampled between February and June. Their results were consistent with our data for Brignant, with Glomeromycotina increasing from a relative abundance of 0.63% in eDNA to 10% in eRNA. Unfortunately, they did not present the data for each site separately so it is not possible to determine whether the time of sampling influenced the results which they reported.

2.5 Polymerase chain reaction (PCR)

PCR permits DNA molecules to be copied (amplified) in an exponential manner ("chain reaction'), allowing very small amounts (potentially single molecules) to be detected. It is a relatively simple reaction requiring an enzyme that synthesizes DNA (a DNA polymerase), synthetic oligonucleotide primers which bind the target DNA and initiate DNA synthesis and the 'building blocks' for DNA (i.e. A,C,G and T deoxyribonucleotide triphosphates [dNTP]). The reaction is taken through a series of heating/cooling steps to denature double stranded DNA, permit primer binding to the target molecule and to synthesize new DNA . Repetition of these temperature conditions (25-35 times leads to DNA amplification, with the amount of the target DNA potentially doubling in each cycle.

There are a number of factors to consider when amplifying DNA from environmental samples for metabarcoding studies. Firstly the annealing temperature for primers; higher temperature makes primer binding more specific, if too high it may exclude some organisms of interest that may differ at certain positions along the primer, too low and organisms not of interest may be amplified, or worse still, primers may bind to regions of DNA outside that we are trying to amplify. Primer binding temperatures can be predicted to a certain degree based on primer sequence but the best temperature is usually determined empirically. For ITS2 we use 55°C and LSU 52°C. The number of cycles is also important, too few and not enough product is generated and too many increases the number of errors in the sequence, including chimeric amplicons (amplicons made from DNA from more than one species see 2.7 below). It is important therefore to use a good quality polymerase that can reduce the number of cycles required.

A further consideration of PCR is whether to include any necessary sequencing adaptors and indexes necessary to identify samples with the amplifying primers. A DNA sequence is directional and has a 5' (5 prime) end (a phosphate group attached to the 5 carbon of the ribose backbone) and a 3' (3 prime) end (unmodified from the ribose -OH substituent on the 3 carbon). Primers bind to DNA from the 5' end and DNA polymerase extends DNA from the 3' end. Therefore, it is possible to attach an additional short DNA sequence (an

adapter) to the 5’ end of the primer which should not interfere with amplification. These adapter sequences are designed to be compatible for the particular sequencing platform being used and also for each DNA sample the adapter region contains a unique index to permit identification of amplicons originating from that particular samples at the end of the sequencing process (confusingly these index sequences are sometimes referred to as barcodes). The advantage of this is that only one step is required to amplify product for sequencing although there is evidence that there could be some interference from these primer tails (O’Donnell et al., 2016). The alternative would be to ligate the adaptors at a later stage (an inefficient reaction) or have a smaller extension at the 5’ end (10 bases) that allows for a second round PCR amplification that adds adaptors and indexes. A further advantage of this is that the first round PCR is not sequencing platform specific.

The PCR reaction itself needs to be set up in a different laboratory than that used for DNA extraction because again a small level of contaminant fungal DNA can lead to large differences in the end result. A PCR cabinet is the preferred option for setting up the reaction that has been UV decontaminated first and as with DNA extraction, filter tips are essential. Negative controls are needed in the PCR to determine if contamination has taken place during DNA extraction or preparing the PCR reaction. The first negative includes the blank DNA reaction and the second negative includes U.V. irradiated water. Both negatives should appear blank on an electrophoresis gel, if the first negative is positive on the gel but not the second then contamination has occurred during DNA extraction, which needs to be repeated. If the second negative shows positive on the gel then contamination has occurred to the reagents or PCR hood and replacing reagents and/or cleaning the hood is necessary.

2.6 Sequencing technology

For all of the sequencing work here we have used the Ion Torrent platform which is the technology with which we are most familiar; however there are other options and the most common are briefly described below. Currently for eDNA metabarcoding the ‘industry-standard’ is Illumina MiSeq and that would be the recommended method. As other methods such as Nanopore, improve metagenomic sequencing direct from eDNA extracts becomes more of a possibility.

Ion Torrent PGM/S5

Ion Torrent is a "sequencing by synthesis" method, where a complementary strand is built based on the template sequence. It uses a semiconductor based technology and works on the principle that when a nucleotide is joined to the end of the synthesising DNA sequence a proton is released. By measuring the charge when particular nucleotide is added to the reaction it is possible to determine if a base or bases have been added and hence determine the sequence of the original complimentary template strand. When Ion Torrent was originally released there was a problem with homopolymers (runs of the same base within a sequences) and determination of the exact number of bases present; this problem has largely been resolved as the sequencing chemistry and base calling software have

improved. Ion torrent chips are of different sizes, for the PGM machine the largest is the 318 generating up to 10 million sequences (each up to 400 bp long) and a cost per run of £1200 pounds and 100 samples per run.

Illumina MiSeq

Illumina is another “sequencing by synthesis” method, but uses fluorescently tagged nucleotides, each fluorescing at a different wavelength that is detected by a camera. In contrast to the Ion Torrent system, only 1 base is added at each round of synthesis as the base has a blocking group attached to the 3’ end, which is chemically removed before the next round of synthesis. This means that the Illumina method accurately sequences homopolymeric sections of DNA. Sequences are generated from both ends of the DNA to generate paired end reads that are combined to generate a single sequence. MiSeq offers a read length of up to 300bp giving a combined read length of approximately 550bp with a 50bp overlap. This method has become the ‘industry standard’ for eDNA metabarcoding and the cost per experiment is about 1.5 times that than of 318chip IonTorrent run. However, the amount of sequence data obtained per run is approximately two fold greater. Run times are significantly longer taking up to 65 hours, compared to 18-24 hours for Ion Torrent

Pac Bio

Pac Bio utilises single molecule real time sequencing and does not require amplification of each DNA molecule to be sequenced as does Ion Torrent and Illumina, but still captures sequence information during the replication process of the target DNA molecule and hence is sequencing by synthesis. It is possible to sequence very long molecules and so is more applicable to genome sequencing but can be used for longer metabarcoding regions such as ITS2 and LSU D1 combined. Accuracy used to be a problem with error rates of up to 15% but this has improved with the latest chemistry and chips which can sequence up to 8 million DNA molecules (Castaño et al., 2020).

Nanopore

Unlike Ion Torrent, Illumina and Pac-Bio, Nanopore is not sequencing by synthesis. The method uses a lipid membrane embedded with protein nanopores. The membrane separates two chambers with a charge gradient across it. Electrodes then measure the charge flow through the nanopores and as a DNA strand enters the pore this flow is restricted by an amount that differs depending on the base inside the pore. By measuring the charge differences as the DNA strand moves through the pore the DNA can be sequenced. The cost of nanopore sequencing is lower than either Ion Torrent or Illumina but at the moment the basecalling accuracy is not good enough for amplicon sequencing. Its main advantage lies in its portability and ability to sequence very long strands that can be very helpful in genome sequencing and assembly when combined with shorter fragment Illumina sequencing.

2.7 Data processing

The processing of data from a sequencing run is an important aspect of metabarcoding analyses that can take time and effect the final result if not done correctly (Anslan et al., 2018; Pauvert et al., 2019). The process that takes raw sequences from the sequencer to the final output is referred to as a pipeline, which is a series of steps that the sequence data moves through to get to the final goal of a report. Several software packages have been developed to provide the necessary functionality, so the pipeline in essence strings these together and manages the datasets at each stage. Several pipelines exist for sequence processing such as QIIME2 (Caporaso et al., 2010), PipeCraft (Anslan et al., 2017) and PIPITS (Gweon et al., 2015), however at Aberystwyth we have developed an in-house pipeline that allows fine tuning of the processing and report output. The steps involved in a metabarcoding pipeline are fairly standard and are discussed briefly below.

Quality checking and sequence trimming. The first step in any pipeline is to remove poor quality sequences. Each sequence is assigned a score by the base-calling algorithm of the particular sequencing platform, these can be checked and any sequences that fall below a quality threshold are either removed or truncated from where the quality falls. In addition to quality scores a further check should be performed to see if the primer sequences are present. Quality checked sequences are split into samples based on the short identifying sequence incorporated into sequences via PCR. ITS2 sequences will include regions of the 5.8S and LSU that do not offer much taxonomic information, it is possible to trim these regions to leave just the ITS2 using ITSx (Bengtsson-Palme et al., 2013), this affects OTU generation but seems to have little impact on correct taxonomic assignment.

Chimera removal. Chimera sequences can be formed during PCR when a single sequence is formed from multiple parent sequences. These can be recognised by dividing the sequence into sections and matching these either to a database (multiple matches indicates a chimera) or during clustering where OTUs themselves form the database against which each new OTU is checked (Edgar, 2013).

Clustering (remove low count variants). The Initial level of clustering is at 100% (i.e. identical sequences are clustered together) with the removal of unique sequences, likely to be due to sequencing errors. The secondary level of clustering forms operational taxonomic units (OTUs). The Level of clustering is important. For ITS2, a 98.5% identity threshold is used to match UNITE species hypotheses and we the same or higher degree of clustering for LSU. This level of clustering is performed on all sequences for all samples. Individual sample sequences are then compared to the OTU file to determine which OTU they belong to, at this stage we also remove small clusters (i.e. singletons / doubletons containing only 1 or 2 sequences) or at a higher threshold if we are not interested in very rare cluster counts for example if we consider they are not large enough to be major players or able to form fruitbodies (section 4).

Taxonomic assignment. The key to a correct taxonomic assignment is a high quality sequence database to match metabarcoding data against. It is also important that this database is flexible and new names can be assigned and sequences added as more

taxonomic studies are performed and relationships between species become clearer. For the UNITE database sequences are added regularly and curated by mycologists with interests in the taxonomy of particular groups, and this filters down to the database used for taxonomic assignment. No such mechanism exists for LSU so we have taken as a starting point the database developed for RDP and added sequences to it from GenBank and those generated at Aberystwyth to improve the taxonomic assignment of grassland fungi especially those found in the UK.

The software used to match sequence data to the database can also affect the result and a number of options exist. Since metabarcoding assigns a taxonomy to many sequences using a large database, performance as well as accuracy is a consideration. We have chosen the RDP project's Naïve Bayesian Classifier (Wang et al., 2007). This tool builds into the database the taxonomic classification at different levels and will return for each sequence the probability score for each level. Therefore if the probability is too low to make an assignment at the highest level (species for ITS2; genus [and species for some groups] for LSU), we can examine the probability at family level and so on until an assignment can be made with certainty.

Data presentation. Once sequences have been through the above processes the data needs to be formatted as a table, with taxonomic assignments and OTUs as rows and the samples (e.g. sites, quadrats or sampling time) as columns with the relative sequence abundance as the data in this matrix (e.g. Table 2). We have chosen to output this data in Microsoft Excel format as the software is ubiquitous and widely known. In addition to the matrix above, summary statistics at different taxonomic levels, alpha diversity measures and beta diversity comparisons in the form of a Bray-Curtis distance matrix are also automatically calculated.

3. Comparisons of eDNA methods with fruitbody surveys

3.1 Welsh grasslands ((Natural Resources Wales/Countryside Council for Wales)

Between 2003 and 2005 50 semi-natural grassland across Wales were selected and during that period surveyed twice for CHEGD fruitbodies using a 900 m^2 quadrat (Griffith et al., 2006). Based on Hygrophoraceae species counts after the two surveys, it was found that 31% of the sites were ranked as of national importance (≥17 spp.) and 8% of international importance (≥22 spp.) according to the classification of Rald (1985). In October 2012 a subset of these sites was surveyed again and at the same time the quadrats were soil cored by the surveyor (36 cores in a grid pattern), using a small domestic apple corer. The same quadrats were repeat cored one month later to provide a comparison and indicate the reproducibility of the sequencing results. The soil sample from the apple corer was shorter (7cm depth; 17.5 mm diameter) than that from later sampling using the now standard 17 mm gouge auger (200 g in these 'apple-corer' surveys against 500 – 700 g in later surveys). DNA extraction, PCR (using LSU primers), sequencing and bioinformatics was conducted as above, and the sequence data was reanalysed against the latest updated LSU database.

The comparison between the original 2003-5 surveys and the sequence data for Hygrophoraceae is shown in Table 6 (unpublished data). The first thing to note is that there is a significantly higher (ANOVA $P<0.001$) average number of species found through eDNA (18) compared to fruitbody surveys (11). Of all the species/site observations, 155 were for both fruitbody and eDNA, 154 eDNA only and 35 fruitbody only. The high prevalence of eDNA only observations is potentially due to a number of reasons. For example, the mycelial network may not be large enough to support fruiting, conditions may not be suitable for fruiting of that particular species and fruiting may happen earlier or later than the survey date, even though in this case there were two dates, and fruiting frequency may vary between species.

P. calyptriformis showed higher abundance in fruitbody surveys compared to eDNA. It is possible that the distribution of this species may have changed from the original 2003-5 surveys to the eDNA survey in 2012. Fruitbody surveys were also conducted at the time of soil collection in 2012 but fruitbody counts were low. For the 2012 fruitbody and eDNA surveys three sites had both fruitbody and eDNA data and only 1 site had just fruitbody data for *P. calyptriformis,* indicating the possibility of a change in distribution. Another potential explanation could be a relatively small and localised mycelial network that reduces the chances of it being sampled by coring.

Table 6. Fruitbody surveys in 2004 – 2005 compared to eDNA analysis from 2011-12 for permanent quadrats in Wales. BOTH indicates species found in surveys from both dates, ONE from only one date. Only one survey date at Caeau Llety Cybi

		Allt goch		Blaen Nedd Q1		Blaen Nedd Q2		Blaen Nedd Q3		Caeau Llety Cybi		Great Orme Q1		Great Orme Q2		Great Orme Q3	
Notes	Species	FB	DNA	FB	DNA	FB	DNA	FB	DNA	FB	DNA	FB	DNA	FB	DNA	FB	DNA
	Chromosera citrinopallida																
	Cuphophyllus colemannianus			BOTH				ONE						ONE			
	Cuphophyllus flavipes	ONE	BOTH	ONE	ONE				ONE				ONE				
	Cuphophyllus fornicatus		ONE		ONE		ONE	ONE	BOTH	YES	YES						
	Cuphophyllus pratensis	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	ONE	ONE	YES		BOTH	BOTH			BOTH	BOTH
No Variation	Cuphophyllus pratensis v pallida			ONE													
	Cuphophyllus russocoriaceus		BOTH	ONE	BOTH		BOTH	ONE	BOTH		YES	ONE	BOTH	BOTH	BOTH		BOTH
	Cuphophyllus virgineus	ONE	BOTH	ONE	ONE		BOTH	BOTH	BOTH	YES	YES		BOTH	BOTH	BOTH	BOTH	BOTH
No variation	Cuphophyllus virgineus var. fuscescens			ONE				ONE									
No variation	Cuphophyllus virgineus var. ochraceopallida											YES					
	Gliophorus irrigatus				ONE		ONE		ONE			BOTH	BOTH				

	Gliophorus laetus					ONE	ONE	ONE	ONE								
	Gliophorus psittacinus	ONE		BOTH	BOTH	BOTH	ONE	ONE		YES			ONE	ONE		BOTH	ONE
	Gliophorus reginae												ONE				
	Gloioxanthomyces vitellina																
	Hygroaster												ONE		ONE		
	Hygrocybe acutoconica				ONE												
No LSU	Hygrocybe aurantiosplendens											YES					
	Hygrocybe cantharellus		BOTH		ONE										ONE		
	Hygrocybe ceracea				ONE				ONE	YES					ONE	BOTH	ONE
	Hygrocybe cf. acutoconica		ONE				ONE				YES						
	Hygrocybe chlorophana	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	ONE				BOTH	ONE			BOTH	
	Hygrocybe citrinovirens				ONE		ONE										
	Hygrocybe coccinea			BOTH	BOTH	ONE	ONE	BOTH	BOTH			BOTH	ONE	BOTH	BOTH	BOTH	BOTH
	Hygrocybe conica	BOTH	BOTH	BOTH	BOTH		BOTH	BOTH	BOTH	YES	YES	BOTH	ONE	BOTH	BOTH		BOTH
	Hygrocybe constrictospora												ONE	BOTH	BOTH		
	Hygrocybe glutinipes		BOTH		ONE		ONE		ONE				ONE		ONE		ONE
	Hygrocybe helobia																

	Hygrocybe insipida	ONE	ONE		BOTH		ONE	BOTH	BOTH			ONE	BOTH	BOTH	ONE	BOTH	BOTH
	Hygrocybe intermedia	ONE	ONE	ONE	ONE								ONE				
	Hygrocybe miniata																
	Hygrocybe mucronella		BOTH		BOTH		BOTH	ONE	BOTH	YES	YES		BOTH		BOTH		BOTH
	Hygrocybe noninqAffBrignant		ONE		ONE				ONE				ONE		ONE		ONE
	Hygrocybe phaeococcinea		ONE								YES						
	Hygrocybe punicea		ONE		BOTH		ONE		BOTH			ONE	BOTH	ONE	BOTH	BOTH	BOTH
	Hygrocybe quieta	BOTH	ONE	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	YES	YES		BOTH	BOTH	BOTH	BOTH	ONE
	Hygrocybe reidii	BOTH	ONE		ONE	BOTH	BOTH	BOTH	BOTH				BOTH		ONE		ONE
	Hygrocybe spadicea		ONE				ONE		ONE		YES						
NO LSU	Hygrocybe splendidissima	BOTH		ONE													
	Neohygrocybe ingrata																
	Neohygrocybe nitrata																
	Neohygrocybe ovina										YES						
	Porpolomopsis calyptriformis											ONE					

	JQ657783.1 Hygrocybe sp. TU112116																
	Total	11	19	13	22	7	19	14	18	8	10	11	20	9	15	9	14
	BOTH	5	9	6	10	5	8	6	10	0	0	5	9	7	8	9	8
	% Both surveys	45.45 %	47.37 %	46.15 %	45.45 %	71.43 %	42.11 %	42.86 %	55.56 %	0.00 %	0.00 %	45.45 %	45.00 %	77.78 %	53.33 %	100.00%	57.14 %

Table 6 (cont. 1)

		Hay Common Q1		Hay Common Q2		Hay Common Q3		Mynachdy'r Graig Q1		Mynachdy'r Graig Q2		Maes Caradog		Somerton Farm	
Notes	Species	FB	DNA	FB	DNA	FB	DNA	FB	DNA	FB	DNA	FB	DNA	FB	DNA
	Chromosera citrinopallida														
	Cuphophyllus colemannianus														
	Cuphophyllus flavipes			ONE	ONE								BOTH		ONE
	Cuphophyllus fornicatus		ONE						ONE	ONE		ONE			ONE
	Cuphophyllus pratensis	ONE	BOTH	BOTH	BOTH	BOTH	BOTH	ONE	BOTH	ONE	BOTH	BOTH	BOTH	BOTH	BOTH
No variation	Cuphophyllus pratensis v pallida	ONE													
	Cuphophyllus russocoriaceus		BOTH		BOTH	BOTH	BOTH	ONE	BOTH		BOTH		BOTH		BOTH
	Cuphophyllus virgineus	BOTH	ONE	BOTH		BOTH	BOTH		ONE		ONE			BOTH	ONE
No variation	Cuphophyllus virgineus var. fuscescens														
No variation	Cuphophyllus virgineus var. ochraceopallida														
	Gliophorus irrigatus	ONE	ONE		ONE	ONE	BOTH				BOTH		BOTH	BOTH	ONE
	Gliophorus laetus	ONE	ONE	ONE		ONE	ONE	ONE	ONE			BOTH			
	Gliophorus psittacinus	BOTH	ONE	ONE	ONE	BOTH	BOTH	ONE		ONE	BOTH	BOTH	ONE	BOTH	ONE
	Gliophorus reginae														

	Gloioxanthomyces vitellina												ONE		
	Hygroaster														
	Hygrocybe acutoconica														
No LSU	Hygrocybe aurantiosplendens														
	Hygrocybe cantharellus		BOTH		BOTH		BOTH		ONE			ONE	BOTH	ONE	ONE
	Hygrocybe ceracea	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	ONE	BOTH	ONE	BOTH	BOTH	BOTH		BOTH
	Hygrocybe cf. acutoconica						BOTH				ONE				ONE
	Hygrocybe chlorophana	BOTH	ONE	BOTH	ONE	BOTH	BOTH	ONE		ONE	BOTH	ONE	BOTH	BOTH	ONE
	Hygrocybe citrinovirens										ONE		ONE		ONE
	Hygrocybe coccinea	ONE	BOTH	ONE	BOTH	BOTH	BOTH	ONE	BOTH	ONE	BOTH		ONE	ONE	BOTH
	Hygrocybe conica	ONE	ONE	ONE	BOTH	ONE	ONE					ONE	ONE		BOTH
	Hygrocybe constrictospora														
	Hygrocybe glutinipes		ONE	ONE	BOTH		BOTH				ONE		BOTH	BOTH	BOTH
	Hygrocybe helobia														
	Hygrocybe insipida	BOTH	BOTH		BOTH	ONE	BOTH		ONE		BOTH	BOTH	ONE	ONE	ONE
	Hygrocybe intermedia				ONE		ONE						ONE	ONE	ONE
	Hygrocybe miniata		ONE		ONE		ONE								
	Hygrocybe mucronella		BOTH			ONE	ONE				BOTH		ONE		BOTH
	Hygrocybe noninqAffBrignant				ONE		ONE						BOTH		BOTH

	Hygrocybe phaeococcinea														
	Hygrocybe punicea	ONE	ONE	ONE	BOTH		ONE	ONE	BOTH				ONE		ONE
	Hygrocybe quieta	BOTH	ONE	ONE	ONE	ONE	BOTH			ONE	BOTH	ONE	BOTH	ONE	BOTH
	Hygrocybe reidii	BOTH	BOTH	ONE	BOTH	ONE	BOTH	ONE	BOTH	ONE	BOTH	BOTH	BOTH		BOTH
	Hygrocybe spadicea				ONE		BOTH				ONE		ONE		BOTH
NO LSU	Hygrocybe splendidissima					BOTH									
	Neohygrocybe ingrata						ONE								
	Neohygrocybe nitrata														
	Neohygrocybe ovina														ONE
	Porpolomopsis calyptriformis							ONE			ONE	ONE		ONE	
	JQ657783.1 Hygrocybe sp. TU112116		ONE												
	Total	14	20	13	19	15	23	10	11	8	17	12	21	12	24
	BOTH	7	8	4	10	8	15	0	6	0	11	6	11	6	11
	% Both surveys	50.00%	40.00%	30.77%	52.63%	53.33%	65.22%	0.00%	54.55%	0.00%	64.71%	50.00%	52.38%	50.00%	45.83%

Table 6 (con't 2)

		Trawscoed Q1		Trawscoed Q2		Total		BOTH	FB	DNA
Notes	Species	FB	DNA	FB	DNA	FB	DNA	DNA & FB	ONLY	ONLY
	Chromosera citrinopallida		ONE			0	1	0	0	1
	Cuphophyllus colemannianus					3	0	0	3	0
	Cuphophyllus flavipes	ONE	BOTH		ONE	4	9	4	0	5
	Cuphophyllus fornicatus					4	8	2	2	6
	Cuphophyllus pratensis	BOTH	BOTH	BOTH	BOTH	16	15	15	1	0
No Variation	Cuphophyllus pratensis v pallida					2	0	0	2	0
	Cuphophyllus russocoriaceus		BOTH		ONE	6	17	6	0	11
	Cuphophyllus virgineus					10	13	9	1	4
No variation	Cuphophyllus virgineus var. fuscescens					2	0	0	2	0
No variation	Cuphophyllus virgineus var. ochraceopallida					1	0	0	1	0
	Gliophorus irrigatus	BOTH	BOTH	BOTH	BOTH	6	12	6	0	6
	Gliophorus laetus	ONE	ONE		ONE	8	7	6	2	1
	Gliophorus psittacinus	BOTH	BOTH	BOTH	BOTH	16	12	11	5	1
	Gliophorus reginae					0	1	0	0	1
	Gloioxanthomyces vitellina					0	1	0	0	1

	Hygroaster					0	2	0	0	2
	Hygrocybe acutoconica					0	1	0	0	1
No LSU	Hygrocybe aurantiosplendens					1	0	0	1	0
	Hygrocybe cantharellus		ONE	ONE	BOTH	3	11	3	0	8
	Hygrocybe ceracea		ONE			8	12	7	1	5
	Hygrocybe cf. acutoconica					0	6	0	0	6
	Hygrocybe chlorophana	ONE	ONE	BOTH	ONE	15	12	12	3	0
	Hygrocybe citrinovirens				ONE	0	6	0	0	6
	Hygrocybe coccinea				ONE	12	14	12	0	2
	Hygrocybe conica		ONE	ONE	ONE	11	15	11	0	4
	Hygrocybe constrictospora					1	2	1	0	1
	Hygrocybe glutinipes	ONE	BOTH	ONE	BOTH	4	15	4	0	11
	Hygrocybe helobia	ONE		ONE		2	0	0	2	0
	Hygrocybe insipida	ONE	ONE		ONE	10	16	10	0	6
	Hygrocybe intermedia					3	7	3	0	4
	Hygrocybe miniata	ONE				1	3	0	1	3
	Hygrocybe mucronella				ONE	3	14	3	0	11
	Hygrocybe noninqAffBrignant				ONE	0	11	0	0	11
	Hygrocybe phaeococcinea					0	2	0	0	2

	Hygrocybe punicea		ONE			6	14	6	0	8
	Hygrocybe quieta		ONE		ONE	13	16	13	0	3
	Hygrocybe reidii				ONE	9	15	9	0	6
	Hygrocybe spadicea		ONE			0	10	0	0	10
NO LSU	Hygrocybe splendidissima					3	0	0	3	0
	Neohygrocybe ingrata					0	1	0	0	1
	Neohygrocybe nitrata			ONE		1	0	0	1	0
	Neohygrocybe ovina					0	2	0	0	2
	Porpolomopsis calyptriformis	BOTH	BOTH	BOTH	BOTH	6	3	2	4	1
	JQ657783.1 Hygrocybe sp. TU112116		ONE		ONE	0	3	0	0	3
						Average				
	Total	11	17	10	19	11.00	18.12	155	35	154
	BOTH	4	7	5	6	4.88	8.65			
	% Both surveys	36.36%	41.18%	50.00%	31.58%	44.09%	46.71%			

3.2 Llanishen (Cardiff)

Fruitbody data for Llanishen (see section 2.4 above) was published in a report for Dŵr Cymru Welsh Water (Sturgess, 2020). The survey was conducted in a continuum around the two reservoirs and hence some of the survey area was outside of the eDNA quadrats. Table 7 shows species present in a series of surveys from 19th September 2019 to 13th January 2020 to capture as many fruitbodies as possible as well as those found from the eDNA studies (ITS and LSU), samples for which were taken in autumn 2018. The combined ITS and LSU species total matches well with the fruitbody survey (24 spp. via eDNA, 25 spp. via FB surveys and 27 spp. across both surveys) but highlights the issue that no single metabarcoding method gives a perfect answer. The four species found only in the LSU data were at low abundance and did not appear in the ITS2 data given the lower relative abundance of longer sequences (see section 2.4). This might be resolved with greater sequencing depth but also at a greater cost. The two species where only ITS2 barcodes are available were *H. splendidissima* and *C. colemannianus*.

Table 7. Llanishen fruitbody surveys autumn to winter 2019/20 compared to eDNA results

Species	eDNA	01-Sep-	08-Oct-19	17-Oct-19	24-Oct-19	01-Nov-	11-Nov-19	18-Nov-	29-Nov-	06-Dec-	13-Dec-	20-Dec-	06-Jan-20	13-Jan-20
Cuphophyllus colemannianus (ITS ONLY)	+			+	+	+	+	+	+	+				
Cuphophyllus flavipes	+	+												
Cuphophyllus fornicatus (LSU ONLY)	+				+	+			+					
Cuphophyllus pratensis	+	+		+	+	+	+	+	+	+	+	+	+	+
Cuphophyllus russocoriaceus						+	+	+	+	+	+			
Cuphophyllus virgineus	+			+	+	+	+	+	+	+	+			
Gliophorus irrigatus	+		+	+	+	+	+		+					
Gliophorus laetus	+													
Gliophorus psittacinus	+	+	+	+	+	+	+	+	+					
Hygrocybe acutoconica	+			+	+									
Hygrocybe aurantiosplendens					+	+		+	+					
Hygrocybe cantharellus (LSU ONLY)	+	+		+		+								
Hygrocybe ceracea				+	+	+		+	+		+		+	
Hygrocybe chlorophana	+	+	+	+	+	+	+	+	+					
Hygrocybe citrinovirens	+	+	+	+		+								
Hygrocybe coccinea	+		+	+	+	+	+	+	+	+	+			+
Hygrocybe conica	+	+	+	+	+	+	+		+					
Hygrocybe glutinipes	+													
Hygrocybe insipida (LSU ONLY)	+		+	+	+	+	+	+	+					
Hygrocybe intermedia	+	+		+	+	+								
Hygrocybe mucronella	+			+		+	+	+	+					
Hygrocybe punicea	+			+	+	+	+	+	+	+	+	+	+	+
Hygrocybe quieta	+	+		+	+	+	+	+	+					
Hygrocybe reidii(LSU ONLY)	+	+		+	+		+	+						
Hygrocybe splendidissima (ITS ONLY)	+				+	+	+		+	+				
Neohygrocybe ingrata	+		+											
Porpolomopsis calyptriformis	+	+	+	+	+	+	+	+	+					
Total waxcap species	24	11	9	19	19	21	17	15	18	7	6	2	3	3

3.3 Lundy Island

A comparison of fruitbody and eDNA data for Lundy Island (Griffith et al., 2020; Hedger et al., 2010) is shown in Table 8. Fruitbody data are available for the airfield quadrat (acid grassland; 51.171N,-4.672W) and the northerly heathland sites (51.199,-4.674W), which were dominated by fruitbodies of *C. lacmus.* This corresponds well to the eDNA data

(using the same sampling protocol as above) from two acid grassland quadrats (the Airfield and Castle Hill) and three heathland sites (John O'Groats House, Squire's View, and Rocket Pole Pond). A fourth heathland site (Old Hospital) was dominated by *Calluna* associated dark septate endophytes (DSE), with few Hygrophoraceae species. Several species were found in the fruitbody survey of the airfield but not found in the eDNA soil survey, these were: *C. colemannianus; G. laetus*; *H. aurantiosplendens*; *H. coccinea*; *H. insipida; H. marchii; H. miniata* and *H. splendidissima.* In addition, *C. fornicatus, C. roseascens, H. contrictospora, H. helobia, H. mucronella and H. pseudoconica* were present in the eDNA but no fruitbodies were found. The differences between fruitbody and eDNA surveys are greater in Lundy than at other sites we have analysed, in particular *G. laetus* and *H. coccinea*, absent from eDNA but commonly found in fruitbody surveys. One possible explanation for this discrepancy is that the quadrat for soil sampling was only a small subsection of the area used for fruitbody surveys and hence some mycelial networks were missed.

Table 8. Lundy Island. eDNA (LSU) data from soil samples taken in February 2019 compared to fruitbody presence / absence from surveys from 2013 – 2018 (blue 0.1% to 1%, green 1% to 10%, yellow 10% - 20% and red > 20%)

				Acid Grassland		Heathland				
Species	FB	FB Location	DNA	Airfield	Castle Hill	John O'Groats House	Squire's View	Old Hospital	Rocket Pole Pond	Note
Chrysomphalina sp.			X	0.01%	0.00%	0.02%	0.01%	0.07%	0.02%	
Cuphophyllus colemannianus	X	Airfield								
Cuphophyllus fornicatus			X	0.00%	0.24%	0.07%	0.00%	0.00%	0.00%	
Cuphophyllus lacmus	X	Heath	X	0.00%	0.02%	73.54%	17.34%	0.05%	5.58%	
Cuphophyllus pratensis	X	Airfield	X	3.45%	7.74%	0.00%	0.11%	0.03%	0.13%	
Cuphophyllus roseascens			X	0.00%	0.00%	0.00%	0.00%	0.00%	1.19%	
Cuphophyllus russocoriaceus	X	Airfield	X	0.01%	6.64%	0.00%	0.01%	0.01%	0.04%	
Cuphophyllus virgineus	X	Airfield	X	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	
Gliophorus irrigatus	X	Airfield	X	0.05%	0.01%	0.00%	0.00%	0.01%	0.00%	
Gliophorus laetus	X	Airfield								
Gliophorus psittacinus	X	Airfield	X	0.13%	0.00%	0.00%	0.02%	0.00%	0.03%	
Hygrocybe aurantiosplendens	X	Airfield								No LSU
Hygrocybe cantharellus	X	Airfield	X	1.13%	0.00%	0.00%	0.00%	0.00%	0.01%	
Hygrocybe ceracea	X	Airfield	X	0.01%	0.05%	0.00%	0.00%	0.00%	0.00%	
Hygrocybe chlorophana	X	Airfield	X	0.04%	0.28%	0.00%	0.00%	0.00%	0.00%	
Hygrocybe coccinea	X	Airfield								
Hygrocybe conica	X	Airfield	X	0.90%	0.01%	0.06%	0.11%	0.04%	0.42%	
Hygrocybe contrictospora			X	0.00%	0.03%	0.00%	0.00%	0.00%	0.00%	
Hygrocybe flavipes	X	Airfield								
Hygrocybe glutinipes	X	Airfield	X	0.19%	0.04%	0.00%	0.07%	0.00%	3.28%	
Hygrocybe helobia			X	5.26%	0.00%	0.02%	0.01%	0.00%	0.00%	
Hygrocybe insipida	X	Airfield								
Hygrocybe marchii	X	Airfield								No LSU
Hygrocybe miniata	X	Airfield								

				Acid Grassland		Heathland				
Species	FB	FB Location	DNA	Airfield	Castle Hill	John O'Groats House	Squire's View	Old Hospital	Rocket Pole Pond	Note
Hygrocybe mucronella			X	0.00%	0.29%	0.00%	0.00%	0.00%	0.00%	
Hygrocybe olivaceonigra	X	Airfield								No LSU
Hygrocybe pseudoconica			X	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	
Hygrocybe punicea	X	Airfield	X	46.97%	0.00%	0.03%	0.00%	0.00%	0.00%	
Hygrocybe quieta	X	Airfield	X	0.02%	0.00%	0.00%	0.00%	0.03%	0.02%	
Hygrocybe reidii	X	Airfield	X	9.31%	0.00%	0.00%	0.00%	0.00%	0.00%	
Hygrocybe splendidissima	X	Airfield								No LSU
Lichenomphalia umbellifera			X	0.00%	0.00%	0.02%	0.46%	0.29%	0.00%	

3.4 Hardcastle Crags (Yorkshire)

The aim of this study near Hebden Bridge, Yorkshire (53.76N,-2.03W) was to allow Natural England to evaluate the potential of DNA metabarcoding of soil eDNA as a method for the assessment of biodiversity of fungi (Griffith et al., 2019). Three areas were surveyed. Hollin Hall comprises four sheep-grazed fields owned by the National Trust. Crimsworth Dene includes five ungrazed fields on a steep slope on the opposite side of the valley. The final location were four fields along the Widdop Road, comprising one haymeadow, one sheep-grazed and two sheep-grazed fields at slightly higher elevation. Choice of these areas was guided by previous field surveys in 2015 and 2016 at HH and WID.

A total of 25 species of Hygrophoraceae were found in the eDNA analysis across the quadrats (Table 9). Six had not been observed to fruit at the site during recent autumn surveys. Of these species, the rarest was *Neohygrocybe ingrata*, found in one quadrat at a relatively high abundance. Two Hygrophoraceae species for which fruitbodies had been found in surveys were not detected in soil DNA, these were *C. russocoriaceus,* and *Glioxanthomyces vitellinus.* Based on location, there was good association between eDNA and fruitbody records, (e.g. *C. flavipes*, *C. pratensis*, *H. coccinea*, *H. quieta*). However, for some species, this association was poor (e.g. *H. citrinovirens*, *H. punicea*, *P. calyptriformis*). As noted above if the mycelial systems of these species were present in the quadrat then they were not captured by coring. It would be possible to reduce spacing between cores and so increase the chances of capturing all networks, but this would increase quadrat survey time and analysis costs (see section 2.1).

Table 9. Hardcastle Crags. eDNA (LSU) from soil samples collected October 2017 compared to fruitbody presence / absence

		Crimsworth Dene					Hollin Hall				Widdop Road				
Species	DNA/FB	CD1	CD2	CD3	CD4	CD5	HC1	HC2	HC3	HC4	WID1	WID2	WID3	WID4	WID5
Arrhenia	DNA	0.00%	0.00%	0.00%	0.05%	0.00%	0.00%	0.34%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%
Arrhenia auriscalpia	DNA	0.00%	0.00%	0.03%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%
Cuphophyllus aurantius	DNA	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.28%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Cuphophyllus flavipes	DNA/FB	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3.05%	0.00%	0.00%	0.00%	0.00%	0.00%
Cuphophyllus fornicatus	DNA/FB	0.25%	0.00%	0.00%	0.00%	0.17%	0.00%	0.00%	0.00%	0.00%	0.79%	0.00%	0.00%	0.00%	0.00%
Cuphophyllus fornicatus	DNA/FB	0.00%	0.00%	0.00%	0.00%	0.33%	0.00%	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	1.12%	0.00%
Cuphophyllus pratensis	DNA/FB	10.56%	1.39%	10.33%	3.02%	0.15%	34.12%	3.84%	15.00%	10.56%	1.21%	12.76%	34.17%	12.33%	0.82%
Cuphophyllus roseascens	DNA	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.13%	0.00%
Cuphophyllus russocoriaceus	FB														
Cuphophyllus virgineus	DNA/FB	0.00%	0.00%	0.00%	0.51%	19.98%	0.00%	0.14%	0.44%	1.84%	1.04%	0.07%	0.00%	4.15%	0.87%
Gliophorus irrigatus	DNA/FB	0.00%	1.17%	0.64%	0.80%	0.00%	0.00%	0.00%	1.18%	0.02%	7.15%	0.00%	0.00%	1.00%	2.38%
Gliophorus laetus	DNA/FB	0.00%	0.00%	0.00%	0.00%	2.08%	0.00%	0.00%	3.05%	0.00%	0.00%	0.00%	1.54%	0.00%	0.00%
Gliophorus psittacinus	DNA/FB	3.21%	6.07%	0.76%	4.94%	0.37%	0.03%	3.89%	0.41%	0.36%	0.44%	15.18%	0.00%	0.57%	0.82%
Gloioxanthomyces vitellinus	FB														
Hygroaster albellus	DNA	0.00%	0.02%	0.00%	0.00%	0.00%	0.17%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Hygrocybe aurantiosplendens	DNA/FB	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.31%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Hygrocybe cantharellus	DNA/FB	0.00%	0.00%	0.09%	0.00%	0.04%	0.30%	0.00%	0.01%	0.13%	0.00%	1.64%	0.50%	0.64%	0.00%

Hygrocybe ceracea	DNA/FB	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%	0.00%
Hygrocybe chlorophana	DNA/FB	0.00%	0.00%	3.38%	0.00%	0.00%	0.00%	0.00%	0.00%	12.11%	7.10%	0.00%	0.00%	0.10%	0.00%
Hygrocybe citrinovirens	DNA/FB	1.44%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.06%	0.00%	0.00%	0.00%	0.00%
Hygrocybe coccinea	DNA/FB	1.65%	0.00%	12.39%	0.23%	7.24%	1.92%	0.03%	0.00%	0.41%	8.76%	17.71%	2.03%	0.12%	0.00%
Hygrocybe conica	DNA/FB	4.50%	0.02%	0.02%	0.59%	1.94%	0.19%	0.20%	1.48%	4.50%	0.00%	0.47%	1.94%	0.00%	0.90%
Hygrocybe contrictospora	DNA	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.41%	0.14%	0.00%	0.00%	0.00%	0.00%	0.00%
Hygrocybe glutinipes	DNA	3.23%	0.02%	0.02%	0.00%	0.00%	0.19%	0.00%	0.00%	0.00%	0.00%	3.07%	0.00%	10.70%	1.77%
Hygrocybe insipida	DNA/FB	0.00%	0.58%	2.13%	0.00%	0.58%	0.00%	0.00%	0.02%	0.00%	0.03%	0.00%	0.00%	0.00%	0.09%
Hygrocybe mucronella	DNA/FB	0.09%	0.72%	0.06%	0.00%	0.00%	0.05%	0.00%	0.12%	1.88%	0.07%	1.63%	1.82%	0.00%	0.41%
Hygrocybe noninquinans	DNA	3.06%	5.60%	0.05%	2.37%	1.69%	4.37%	6.40%	1.36%	4.63%	0.03%	0.00%	0.00%	0.00%	0.19%
Hygrocybe phaeococcinea	DNA	0.00%	0.13%	0.19%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Hygrocybe punicea	DNA/FB	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	12.65%	0.00%	0.00%	0.00%	0.00%
Hygrocybe quieta	DNA/FB	0.00%	0.00%	0.00%	0.00%	0.00%	5.38%	0.02%	3.58%	3.30%	0.00%	1.09%	22.16%	0.00%	0.00%
Hygrocybe reidii	DNA/FB	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.16%	1.73%	0.55%	12.13%	0.18%	0.00%	0.00%
Hygrocybe sp. TU112116	DNA	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3.00%	0.00%	0.00%	0.00%
Neohygrocybe ingrata	DNA	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	9.13%	0.00%	0.00%	0.00%	0.00%
Porpolomopsis calyptriformis	DNA/FB	0.00%	0.04%	0.00%	0.00%	0.00%	0.05%	0.00%	0.00%	0.00%	2.09%	0.00%	0.00%	2.39%	0.00%

4. Thresholds for identification of sites with diverse populations of grassland fungi via eDNA metabarcoding

eDNA surveys identify more species than do fruitbody surveys. However, since some species previously found fruiting may be missed if the mycelia of those fungi are absent in the soil cores obtained during sampling, the two approaches are to some extent complementary. Additionally, even a brief initial survey in the autumn fruiting period (by a suitably experienced surveyor) can be very helpful to identify the best areas of a site for later positioning of eDNA quadrats.

When a fruitbody is discovered, it will be formed by a mycelium that has reached a particular threshold biomass level of mycelium in the soil, with the threshold being determined in part by the size of the fruitbody. This would correspond to a “mature individual”, as used by IUCN for Red-Listing purposes (IUCN, 2012), and which has been applied in a similar manner for Red-Listing of UK Boletaceae by Ainsworth et al. (2013). Thus, more underlying mycelium would be required to form a basidiocarp of *Hygrocybe punicea* (10-60g fresh weight (FW)) than *Cuphophyllus virgineus* (0.5-3g FW) and it could be estimated very approximately that the total weight of fruitbodies formed in a season is 100 to 1000-fold less that the biomass of the underlying mycelium.

However, for eDNA metabarcoding, assuming 10,000 sequence reads per sample (i.e. quadrat), a species present at much lower biomass levels than sufficient to form a fruitbody could potentially be detected. In our analyses, we (conservatively) estimate that a particular species can be reliably listed as present if >50 sequences are detected (0.5%).

Total fungal biomass in grassland soils (based on our measurements from the Brignant site) can be estimated at ca. 20 mg FW/g soil (i.e. 2%) or 2 mg DW/g soil based on quantification of ergosterol (a fungal-specific sterol which comprises ca. 0.5% dry wt. of fungal mycelia; estimated to be present at 10 µg/g DW soil) and assuming that mycelia/fruitbodies comprise 90% water (Detheridge et al., 2018). Thus a typical pooled-core soil sample for eDNA metabarcoding of a 900 m^2 quadrat (36 cores = 700 g fresh and at ca. 25% water content, ca. 500 g when dried) would contain ca. 1 g DW fungal biomass. In a 900 m^2 quadrat, there is ca. 100,000 kg of ‘topsoil’ (assuming 10-15 cm depth , bulk density of 1 g/cm^3 and moisture content of ca. 25%) of which ca. 0.001% is sampled during coring; thus the whole quadrat would contain ca. 1,000 kg FW (100 kg DW) mycelium.

One particular observation at Parkgrass long term experiment at Rothamsted in November 2008 provides some useful ground-truthing of the calculations above. *H. punicea* was present in profusion on one plot (no fertiliser or lime; 100m^2 area; ca. 10,000 kg ‘topsoil’) and since the annual aftermath cut was about to take place, all the basidiocarps were collected. These ca. 60 basidiocarps weighed >2 kg. Using the estimates above and

generously assuming 1% conversion of mycelial biomass to basidiocarp, the plot contained 200 kg FW *H. punicea* mycelium (= ca. 2 kg FW/m^2). Based on data from a quadrat at The Leasowes (known as '*Punicea* Bank', where this species accounted for 50% of all the fungal DNA present)(Griffith et al., 2018; Griffith et al., 2015), this would give a total fungal biomass of ca. 4 kg FW/m^2I (1m^3 soil area = ca. 100 kg 'topsoil'). This equates to 40 mg FW fungal biomass/g soil, close to the 20 mg FW/g soil stated above for Brignant.

Bringing the data above together, we would list a species as present if it comprised >5mg of the 1g fungal biomass present (i.e. >0.5%) in the pooled-core soil sample. This would equate to 10 kg FW mycelium/quadrat for *H. punicea* of the 2000 kg FW total fungal mycelium present in the quadrat. And estimating that 0.1-1% of mycelial biomass is converted to basidiocarps each season, this 10 kg FW would equate to 10-100 g fruitbodies, ca 1-200 fruitbodies depending on the species. These estimates are crude but nonetheless valuable in order to estimate how relative abundance of the DNA of particular species might relate to the number of fruitbodies formed by its mycelia. The greatest uncertainty lies in the estimation of the percentage of mycelial biomass annually converted to fruitbodies (which may vary considerably between species) and the heterogeneous (possibly annular á la fairy ring) distribution of mycelia within soil. Regarding the latter point, it is not known whether the mycelium of a particular species later detected by eDNA metabarcoding originated from a single core (of the 36) where that species was present at a high amount or whether it was present at lower levels in all the cores (the former being more likely that the latter).

To summarise, these calculations suggest that a species present at 0.5% abundance in the metabarcoding data for a 900 m^2 quadrat would represent a well-established mycelial system (or several smaller mycelial systems in different parts of the quadrat), capable of forming several fruitbodies, as opposed to very limited mycelial systems with insufficient biomass to form fruitbodies (i.e. a non-mature individual in IUCN parlance). It is highly unlikely that species present only as ungerminated spores would achieve the 0.5% threshold. Our 50/10,000 sequences threshold (0.5% abundance) could be dropped to 10/10,000 (0.1% abundance) and then more species would be counted as present but the likelihood of counting species which are present but not thriving (and thus not likely to form basidiocarp in the near future) would increase. It should be noted that species detected from <3 identical sequences per sample are removed from all our analyses since these may represent sequencing errors. 'Deeper' sequencing, for example obtaining 100,000 sequences per pooled-core quadrat sample is easily possible, though at increased cost (10/100,000 sequences allows detection of species at 0.01% abundance), but the advantage of this approach is not clear. If the aim is to determine whether any DNA of a particular species is present, then use of a specific RT-qPCR assay (as used for COVID 'PCR tests') would be a cheaper alternative.

Clearly, improved understanding of the spatial distribution of these fungal mycelia in soil may lead to modification or verification of these threshold levels, as will progressive accumulation of datasets where fruitbody and eDNA data can be compared. Evidence that the current thresholds are appropriate comes from The Leasowes where six species, not

previously found fruiting during 20 years of recording by Nick Williams (Williams, 2020) were detected in eDNA analysis. Subsequent “eDNA-guided surveying” in the quadrats, where these undiscovered species had been detected in soil eDNA, led to the discovery of fruitbodies of four of these species (Griffith et al., 2018; Griffith et al., 2015). These discoveries, including *Microglossum olivaceum* agg. (a species of principal importance for biodiversity in Section 41 of the NERC [England] Act 2006) and *Hygrocybe citrinovirens* (listed as Vulnerable in IUCN Red List of Threatened Species; https://www.iucnredlist.org/species/70406652/70406717 consolidated evidence for achieving legal protection as an SSSI for this site (Natural_England, 2019).

A consequence of the future deployment of eDNA analysis in conservation of grassland fungi, would likely be that the ‘bar will be raised’ for sites to achieve legal protection. Currently, following Rald’s (1985) initial thresholds, later modified by others (Griffith et al., 2013), sites with 17-21 waxcap species being of national importance and those with >21 spp. being of international importance. Bosanquet et al. (2018) recommend that sites with ≥19 spp. should be considered for SSSI notification and those with 12-18 spp. further surveyed. We see no reason for any immediate change to this guidance but eDNA-guided fruitbody surveying, as undertaken at The Leasowes, can make an important contribution, not only to the identify the location of additional species at known ‘good’ sites (e.g. The Leasowes) but also as a cost and time-efficient method for rapid assessment (ca. £300 per quadrat, with economies of scale), prior to commissioning more detailed surveys by expert field mycologists. Apart from the rarity of suitably proficient mycologists, their costs may be higher, with several visits needed to each site, and survey timing is seasonally restricted.

As an initial step to identify grassland areas where some waxcaps may be present, the “Novice’s guide to fungal diversity at grassland sites” (Griffith et al., 2004), which requires no mycological expertise would be an efficient and useful means to identify and prioritise sites for follow-up eDNA analysis (similar to a phase 1 habitat survey). This ‘Novice guide’ was recently deployed in Plantlife’s smartphone WaxcApp (www.plantlife.org.uk/uk/ discover-wild-plants-nature/habitats/grassland/ waxcaps-fungi/waxcapp-survey) and with suitable support and publicity could become widely used and an invaluable source of information as to the locations of ‘waxcap’ grasslands. Additionally, Bosanquet et al. (2018) recommended the use of high diversity indicator species to flag up sites meriting further attention.

eDNA metabarcoding is also likely to have an impact at a global level on the IUCN Red-List assessment of the conservation status of species of grassland fungi. The Global Fungal Red List Initiative (iucn.ekoo.se/en/iucn/welcome), initiated in 2013 and led by Dr. Greg Mueller (Chicago Botanic Garden) has led in recent years to the formal assessment of many grassland fungi, with many more proposed for assessment and under review. Of the ca. 40 species of waxcap (Hygrophoraceae) found in the UK, one is now classified as Endangered (*Gloioxanthomyces vitellinus*) at a global level and a further 15 as Vulnerable, including *Hygrocybe citrinovirens*, *H. punicea* and *H. spadicea* (IUCN, 2021). For context the snow leopard [*Panthera uncia*] is also classified as Vulnerable and there are no UK mammals in the VU/EN categories (www.mammal.org.uk/science-research/red-list/).

The more efficient discovery of fungi in new locations may reduce the extinction risk for some species but it will also address the Data-Deficiency (DD) problem that prevents assessment of many species. IUCN Criterion B involves assessment of (i) extent of occurrence, (ii) area of occupancy, (iii) area, extent and/or quality of habitat, (iv) number of locations or subpopulations, (v) number of mature individuals. Categories (ii) and (v) in particular could be efficiently assessed by eDNA metabarcoding data, once quantitative thresholds are agreed. Since eDNA datasets published in peer-reviewed journals are deposited in public repositories, it is possible to search for the presence of particular species. Recently a new website (https://globalfungi.com) has brought together 600 million observations (ITS1 or ITS2 sequences from metabarcoding studies) into a searchable database (Větrovský et al., 2020) and discussions are underway to integrate these datasets into GBIF (Schigel et al., 2019).

5. Conclusions

We have shown that from a 30x30 m quadrat, 36 cores taken in a grid pattern across the quadrat give a good representation of all fungi present, including fungi of conservation interest. If a larger area is required several quadrats may be necessary. Taking larger soil samples makes handling and processing unwieldly. Some species are inevitably missed but comparisons show that this method compares well to traditional surveying with repeated visits. We have also shown that the results are not affected by season, as the mycelial network remains intact and viable, although activity may be lower.

Once samples have been taken it is recommended that they are kept cool and shipped in a cool box within a day or so of collection. It is essential that they are not frozen if there is a possibility that the sample will thaw before final frozen storage.

Long term storage at -80°C is recommended before samples are processed. Processing begins with freeze drying which arrests biological processes during the drying and subsequently allows for efficient homogenisation. This is an essential step to ensure the sub sample taken for DNA extraction is representative of the whole. Homogenisation is achieved in two stages, first passing through a 2mm soil sieve and thoroughly mixing by hand and then a sub sample is passed through a 0.5mm soil sieve. The subsample ground through the fine sieve is then used for DNA extraction and stored at -80 for any subsequent analysis. The 2mm ground soil is used for chemical analysis. If homogenisation is performed then only one DNA extract per sample is required.

There is no perfect metabarcoding region for CHEGD grassland fungi. LSU will give a better idea of relative abundance but some (non-CHEGD) species cannot be distinguished. In addition, if rare species have a low abundance they may be missed by ITS2, swamped by shorter sequences. This can be overcome by increasing sequencing depth which increases the chances of picking up rarer species, but of course costs per sample and sample processing time will increase. Nonetheless, both regions perform well compared to fruitbody surveys with equal or greater numbers of Hygrophoraceae taxa found in eDNA analyses. The most abundant tend to compare well between the two

methods but some of the rarer species may be missed by one or the other method. A particular advantage of using ITS2 over LSU, in addition to its greater taxonomic resolution is that in the same sequence run a primer can be added that also amplifies plant DNA.

Currently the most suitable sequencing technology would be Illumina although this should be reviewed regularly as it is a rapidly changing area.

6. Next Steps

- There is still some work required on the molecular taxonomy of grassland fungi. For example, there are still some species of Hygrophoraceae missing from the LSU and ITS2 databases, for example *H. aurantiosplendens* and *H. splendidissima*. This is sometimes due to difficulties in identification, with sequences from fruitbodies identified as *H.splendidissima* grouping with either *H. coccinea* or *H. punicea* and *H. aurontiosplendens* grouping with *H. quieta*. In addition, in the genus *Cuphophyllus* sequences from the species *russocoriacea* splits two groups of *virgineus* making sequence identification difficult. These inconsistencies need to be resolved before accurate identification of all species is possible. This process will involve continual updating of the UNITE ITS2 and RDP LSU databases as the taxonomic concepts of CHEGD species are updated and improved. For Hygrophoraceae, this process is well underway but for the other CHEGD taxa many uncertainties remain.

- Many sites have now been surveyed using these techniques across the UK, building a database of sequence data that could be mined to determine any apparent patterns in the data. For example, co-occurrence of fungal species which may not be apparent from conventional surveys such as the co-occurrence of Hygrophoraceae species with non-fruiting species. Hutchinson's paradox of the plankton (Chesson, 2000; Hutchinson, 1961) suggests that all the different waxcap species occupy different niches and accumulation of accurate distribution and relative biomass data may assist in the elucidation of such niche differences.

- Metabarcoding of ribosomal eRNA could offer a different perspective on grassland fungi as it would indicate activity levels of fungi in soils. Ribosomes are the protein making machines in the cell and are partly formed of ribosomal RNA, more ribosomes and hence more ribosomal eRNA indicates a more active organism and so metabarcoding using eRNA as a template would allow identification of those more active organisms. Levels of activity would vary seasonally, and some organisms could be more active in the spring and some summer or autumn.

- Additionally, if plant data is also obtained through sequencing or conventional surveys then links with putative plant hosts may also become more apparent. As the database is large such analysis would require the development of an automated tool.

- The eDNA methodology benefits from the economies of scale, unlike fruitbody surveying, where there is a limited pool of expert surveyors. As noted above, it would be more cost-efficient (especially of person time) to deploy eDNA analysis at an early

stage in the investigation of fungal biodiversity at a given site, rather than as a confirmatory step. This would efficiently direct the efforts of expert surveyors to the better sites which may qualify for legal protection.

- For SSSI notification there is potential ambiguity, since notification of sites with diverse grassland fungal populations is primarily based on the numbers of “species present”, and thus does not specifically exclude eDNA evidence alone (as opposed to fruitbodies discovered following eDNA-guided surveying). Modification of SSSI guidance (Bosanquet et al., 2018) to make specific mention of ‘mature’ individuals (as in IUCN guidelines) would clarify the guidance since presence of ≥1 fruitbody is by definition a mature individual. Quantification of what amount of biomass would correspond to a mature individual could be undertaken via eDNA metabarcoding and estimation of total fungal biomass via ergosterol as described above, but further research would be required to validate any thresholds.

- The global conservation status of UK grassland and their fungi has been highlighted by recent developments in IUCN Red-listing. There is a need to ensure that policymakers are aware of UK responsibilities in the light of this new information.

List of tables

List of figures

Bibliography

ABARENKOV, K., HENRIK NILSSON, R., LARSSON, K.H., ALEXANDER, I.J., EBERHARDT, U., ERLAND, S., HØILAND, K., KJØLLER, R., LARSSON, E., PENNANEN, T., 2010. The UNITE database for molecular identification of fungi–recent updates and future perspectives. New Phytol. 186, 281-285.

ADAMO, M., VOYRON, S., CHIALVA, M., MARMEISSE, R., GIRLANDA, M., 2021. Metabarcoding on both environmental DNA and RNA highlights differences between fungal communities sampled in different habitats. PLOS ONE 15, e0244682.

AINSWORTH, A.M., SMITH, J.H., BODDY, L., DENTINGER, B.T.M., JORDAN, M., PARFITT, D., ROGERS, H.J., SKEATES, S.J., 2013. Red List of Fungi for Great Britain: Boletaceae. A pilot conservation assessment based on national database records, fruit body morphology and DNA barcoding.

ANSLAN, S., BAHRAM, M., HIIESALU, I., TEDERSOO, L., 2017. PipeCraft: flexible open-source toolkit for bioinformatics analysis of custom high-throughput amplicon sequencing data. Molecular ecology resources, n/a-n/a.

ANSLAN, S., NILSSON, R.H., WURZBACHER, C., BALDRIAN, P., LEHO, T., BAHRAM, M., 2018. Great differences in performance and outcome of high-throughput sequencing data analysis platforms for fungal metabarcoding. MycoKeys, 29-40.

BAZZICALUPO, A.L., BÁLINT, M., SCHMITT, I., 2013. Comparison of ITS1 and ITS2 rDNA in 454 sequencing of hyperdiverse fungal communities. Fungal Ecology 6, 102-109.

BEGEROW, D., NILSSON, H., UNTERSEHER, M., MAIER, W., 2010. Current state and perspectives of fungal DNA barcoding and rapid identification procedures. Applied Microbiology and Biotechnology 87, 99-108.

BENGTSSON-PALME, J., RYBERG, M., HARTMANN, M., BRANCO, S., WANG, Z., GODHE, A., WIT, P., SÁNCHEZ-GARCÍA, M., EBERSBERGER, I., SOUSA, F., 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. Methods in Ecology and Evolution 4, 914-919.

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 Microbiology 10, 1.

BIRKEBAK, J.M., ADAMČÍK, S., LOONEY, B.P., MATHENY, P.B., 2016. Multilocus phylogenetic reconstruction of the Clavariaceae (Agaricales) reveals polyphyly of agaricoid members. Mycologia 108, 860-868.

BLACKSTOCK, RIMES, STEVENS, JEFFERSON, ROBERTSON, MACKINTOSH, HOPKINS, 1999. The extent of semi-natural grassland communities in lowland England

and Wales: a review of conservation surveys 1978–96. Grass and Forage Science 54, 1-18.

BLAZEWICZ, S.J., BARNARD, R.L., DALY, R.A., FIRESTONE, M.K., 2013. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. The ISME journal 7, 2061-2068.

BODDY, L., BÜNTGEN, U., EGLI, S., GANGE, A.C., HEEGAARD, E., KIRK, P.M., MOHAMMAD, A., KAUSERUD, H., 2014. Climate variation effects on fungal fruiting. Fungal Ecology 10, 20-33.

BOSANQUET, S.D.S., AINSWORTH, A.M., COOCH, S.P., GENNEY, D.R., WILKINS, T.C., 2018. 2018. Guidelines for the Selection of Biological SSSIs. Part 2: Detailed Guidelines for Habitats and Species Groups. Chapter 14 Non-lichenised Fungi. . JNCC, Peterborough.

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., SEVINSKY, J.R., TURNBAUGH, P.J., WALTERS, W.A., WIDMANN, J., YATSUNENKO, T., ZANEVELD, J., KNIGHT, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7, 335-336.

CASTAÑO, C., BERLIN, A., BRANDSTRÖM DURLING, M., IHRMARK, K., LINDAHL, B.D., STENLID, J., CLEMMENSEN, K.E., OLSON, Å., 2020. Optimized metabarcoding with Pacific Biosciences enables semi-quantitative analysis of fungal communities. New Phytol. 228.

CLASEN, L.A., DETHERIDGE, A.P., SCULLION, J., GRIFFITH, G.W., 2020. Soil stabilisation for DNA metabarcoding of plants and fungi. Implications for sampling at remote locations or via third-parties. Metabarcoding and Metagenomics 4, e58365.

COLE, J.R., WANG, Q., FISH, J.A., CHAI, B., MCGARRELL, D.M., SUN, Y., BROWN, C.T., PORRAS-ALFARO, A., KUSKE, C.R., TIEDJE, J.M., 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic acids research 42, D633-D642.

CHESSON, P., 2000. Mechanisms of maintenance of species diversity. Annual review of Ecology and Systematics 31, 343-366.

DEINER, K., YAMANAKA, H., BERNATCHEZ, L., 2021. The future of biodiversity monitoring and conservation utilizing environmental DNA. Environmental DNA.

DETHERIDGE, A.P., COMONT, D., CALLAGHAN, T.M., BUSSELL, J., BRAND, G., GWYNN-JONES, D., SCULLION, J., GRIFFITH, G.W., 2018. Vegetation and edaphic factors influence rapid establishment of distinct fungal communities on former coal-spoil sites. Fungal Ecology 33, 92-103. DOI: 110.1016/j.funeco.2018.1002.1002.

DETHERIDGE, A.P., CHERRETT, S., CLASEN, L.A., MEDCALF, K., PIKE, S., GRIFFITH, G.W., SCULLION, J., 2020. Depauperate soil fungal populations from the St. Helena endemic *Commidendrum robustum* are dominated by Capnodiales. Fungal Ecology 45, 100911.

EDGAR, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Meth 10, 996-998.

EDWARDS, P.J., 1988. Effects of the fairy ring fungus *Agaricus arvensis* on nutrient availability in grassland. New Phytol. 110, 377-381.

GANGE, A.C., GANGE, E.G., SPARKS, T.H., BODDY, L., 2007. Rapid and Recent Changes in Fungal Fruiting Patterns. Science 316, 71-71.

GARDES, M., BRUNS, T.D., 1993. ITS primers with enhanced specificity for basidiomycetes-application to the identification of mycorrhizae and rusts. Molecular ecology 2, 113-118.

GEML, J., GRAVENDEEL, B., VAN DER GAAG, K.J., NEILEN, M., LAMMERS, Y., RAES, N., SEMENOVA, T.A., DE KNIJFF, P., NOORDELOOS, M.E., 2014. The Contribution of DNA Metabarcoding to Fungal Conservation: Diversity Assessment, Habitat Partitioning and Mapping Red-Listed Fungi in Protected Coastal Salix repens Communities in the Netherlands. PLOS ONE 9, e99852.

GEORGE, P.B., CREER, S., GRIFFITHS, R.I., EMMETT, B.A., ROBINSON, D.A., JONES, D.L., 2019. Primer and database choice affect fungal functional but not biological diversity findings in a national soil survey. Frontiers in Environmental Science 7, 173.

GRIFFITH, G., CAVALLI, O., DETHERIDGE, A., 2019. An assessment of the fungal conservation value of Hardcastle Crags (Hebden Bridge, West Yorkshire) using NextGen DNA sequencing of soil samples. York, UK: Natural England.

GRIFFITH, G.W., ARON, C., EVANS, D.A., EVANS, S., GRAHAM, A., HOLDEN, L., MITCHEL, D., 2006. Mycological survey of selected semi-natural grasslands in Wales.

GRIFFITH, G.W., BRATTON, J.L., EASTON, G.L., 2004. Charismatic megafungi: the conservation of waxcap grasslands. British Wildlife 15, 31-43.

GRIFFITH, G.W., DETHERIDGE, A.P., CAVALLI, O.T., MOORE, T., 2018. Updated eDNA assessment of the fungal conservation value of The Leasowes (Halesowen, West Midlands) including the golf course areas. Report commissioned by Natural England. 10pp. .

GRIFFITH, G.W., DETHERIDGE, A.P., GEORGE, L., 2015. An Assessment of the fungal conservation value of The Leasowes (Halesowen, West Midlands) using NextGen DNA sequencing of soil samples. Report commissioned by Severn Trent Water Ltd. 10pp. .

GRIFFITH, G.W., EASTON, G.L., JONES, A.W., 2002. Ecology and diversity of waxcap (Hygrocybe spp.) fungi. Transactions and Proceedings of the Botanical Society of Edinburgh and Botanical Society of Edinburgh Transactions 54, 7-22.

GRIFFITH, G.W., GAMARRA, J.G.-P., HOLDEN, E., MITCHEL, D., GRAHAM, A., EVANS, D., EVANS, S., NOORDELOOS, M., KIRK, P., SMITH, S.L., 2013. The international conservation importance of Welsh ‘waxcap’grasslands. Mycosphere 4, 969-984.

GRIFFITH, G.W., GRAHAM, A., WOODS, R.G., EASTON, G.L., HALBWACHS, H., 2014. Effect of biocides on the fruiting of waxcap fungi. Fungal Ecology 7, 67-69.

GRIFFITH, G.W., HEDGER, J.N., ROWLAND, A., 2020. A comparison of fruitbody and e-DNA survey approaches for assessing the distribution mycelia of macrofungi in the grassland and heathland of lundy. Journal of the Lundy Field Society 7.

GRIFFITH, G.W., RODERICK, K., 2008. Saprotrophic basidiomycetes in grasslands: distribution and function, in: Boddy, L., Frankland, J.C., van West, P. (Eds.), Ecology of Saprotrophic Basidiomycetes. British Mycological Society Symposia Series. Elsevier Ltd., London, pp. 277-299.

GWEON, H.S., OLIVER, A., TAYLOR, J., BOOTH, T., GIBBS, M., READ, D.S., GRIFFITHS, R.I., SCHONROGGE, K., 2015. PIPITS: an automated pipeline for analyses of fungal internal transcribed spacer sequences from the Illumina sequencing platform. Methods in Ecology and Evolution 6, 973-980.

HALBWACHS, H., EASTON, G.L., BOL, R., HOBBIE, E.A., GARNETT, M.H., PERŠOH, D., DIXON, L., OSTLE, N., KARASCH, P., GRIFFITH, G.W., 2018. Isotopic evidence of biotrophy and unusual nitrogen nutrition in soil-dwelling Hygrophoraceae. Environmental microbiology 20, 3573-3588.

HARPER, L.R., LAWSON HANDLEY, L., HAHN, C., BOONHAM, N., REES, H.C., GOUGH, K.C., LEWIS, E., ADAMS, I.P., BROTHERTON, P., PHILLIPS, S., 2018. Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (Triturus cristatus). Ecology and evolution 8, 6330-6341.

HARRINGTON, A.P., O’MEARA, D.B., AUGHNEY, T., MCANEYC, K., SCHOFIELD, H., COLLINS, A., DEENEN, H., O’REILLY, C., 2019. Novel real-time PCR species identification assays for British and Irish bats and their application to a non-invasive survey of bat roosts in Ireland. Mammalian Biology 99, 109-118.

HEBERT, P.D., CYWINSKA, A., BALL, S.L., DEWAARD, J.R., 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences 270, 313-321.

HEDGER, J.N., GEORGE, J.D., GRIFFITH, G.W., PEIRIS, D., 2010. Lundy fungi: Further surveys 2004-2008. Journal of the Lundy Field Society 2, 21-44.

HOUSELEY, J., TOLLERVEY, D., 2009. The Many Pathways of RNA Degradation. Cell 136, 763-776.

HUTCHINSON, G.E., 1961. The paradox of the plankton. The American Naturalist 95, 137-145.

IUCN, 2012. Red List categories and criteria, version 3.1, second edition.

IUCN, 2021. The IUCN Red List of Threatened Species. Version 2021-1. https://www.iucnredlist.org. Downloaded on [24 April 2021].

JONES, E.P., ADAMS, I.P., WALSHAW, J., BENUCCI, M., MACARTHUR, R., BOONHAM, N., BRYCE, S., 2020. Guidance for end users on DNA methods development and project assessment. JNCC Report No. 669a. JNCC, Peterborough, ISSN 0963-8091.

KARNAHL, U., WASTERNACK, C., 1992. Half-life of cytoplasmic rRNA and tRNA, of plastid rRNA and of uridine nucleotides in heterotrophically and photoorganotrophically grown cells of Euglena gracilis and its apoplastic mutant W3BUL. The International journal of biochemistry 24, 493-497.

KÕLJALG, U., NILSSON, R.H., ABARENKOV, K., TEDERSOO, L., TAYLOR, A.F.S., BAHRAM, M., BATES, S.T., BRUNS, T.D., BENGTSSON-PALME, J., CALLAGHAN, T.M., 2013. Towards a unified paradigm for sequence-based identification of fungi. Molecular Ecology 22, 5271-5277.

LARIVIERE, F.J., COLE, S.E., FERULLO, DANIEL J., MOORE, M.J., 2006. A Late-Acting Quality Control Process for Mature Eukaryotic rRNAs. Molecular Cell 24, 619-626.

LAWTON, J.H., MAY, R.M., RAUP, D.M., 1995. Extinction rates. Oxford University Press Oxford.

MEIKLEJOHN, K.A., DAMASO, N., ROBERTSON, J.M., 2019. Assessment of BOLD and GenBank–Their accuracy and reliability for the identification of biological materials. PloS one 14, e0217084.

MILLER, S.L., GONGLOFF, A., 2021. Fairy rings, associated fungi, and assessment of their distribution across environmental variables using GIS. Fungal Ecology 50, 101040.

MONARD, C., GANTNER, S., STENLID, J., 2013. Utilizing ITS1 and ITS2 to study environmental fungal diversity using pyrosequencing. FEMS Microbiology Ecology 84, 165-175.

Natural_England, 2019. The Leasowes SSSI Dudley. Notification under Section 28 of the Wildlife and Countryside Act 1981 (Natural England, 7 February 2021)(https://consult.defra.gov.uk/natural-england/the-leasowes/).

NEWTON, A.C., DAVY, L.M., HOLDEN, E., SILVERSIDE, A., WATLING, R., WARD, S.D., 2003. Status, distribution and definition of mycologically important grasslands in Scotland. Biological Conservation 111, 11-23.

NILSSON, R.H., KRISTIANSSON, E., RYBERG, M., HALLENBERG, N., LARSSON, K.-H., 2008. Intraspecific ITS Variability in the Kingdom Fungi as Expressed in the International Sequence Databases and Its Implications for Molecular Species Identification. Evolutionary Bioinformatics 4, EBO.S653.

NILSSON, R.H., LARSSON, K.-H., TAYLOR, A.F S., BENGTSSON-PALME, J., JEPPESEN, T.S., SCHIGEL, D., KENNEDY, P., PICARD, K., GLÖCKNER, F.O., TEDERSOO, L., SAAR, I., KÕLJALG, U., ABARENKOV, K., 2019. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. Nucleic Acids Research 47, D259-D264.

NILSSON, R.H., RYBERG, M., ABARENKOV, K., SJÖKVIST, E., KRISTIANSSON, E., 2009. The ITS region as a target for characterization of fungal communities using emerging sequencing technologies. FEMS Microbiology Letters 296, 97-101.

O'DONNELL, J.L., KELLY, R.P., LOWELL, N.C., PORT, J.A., 2016. Indexed PCR Primers Induce Template-Specific Bias in Large-Scale DNA Sequencing Studies. PLOS ONE 11, e0148698.

PAUVERT, C., BUÉE, M., LAVAL, V., EDEL-HERMANN, V., FAUCHERY, L., GAUTIER, A., LESUR, I., VALLANCE, J., VACHER, C., 2019. Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. Fungal Ecology 41, 23-33.

RALD, E., 1985. Vokshatte som indikatorarter for mykologisk vaerdifulde overdrevslokaliteter. Svampe 11, 1-9.

RAMSBOTTOM, J., 1953. Mushrooms and toadstools. A study of the activities of fungi. Mushrooms and toadstools. A study of the activities of fungi.

READ, D.J., PEREZ-MORENO, J., 2003. Mycorrhizas and nutrient cycling in ecosystems – a journey towards relevance? New Phytol. 157, 475-492.

RILLIG, M.C., MUMMEY, D.L., 2006. Mycorrhizas and soil structure. New Phytol. 171, 41-53.

SCHIGEL, D.S., JEPPESEN, T.S., FINN, R.D., COCHRANE, G., KÕLJALG, U., QUAST, C., LANFEAR, J., ORRELL, T.M., HOBERN, D., MILLER, J.T., 2019. Going Molecular: Sequence-based spatiotemporal biodiversity evidence in GBIF. Biodiversity Information Science and Standards.

SCHOCH, C.L., SEIFERT, K.A., HUHNDORF, S., ROBERT, V., SPOUGE, J.L., LEVESQUE, C.A., CHEN, W., CONSORTIUM, F.B., 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proceedings of the National Academy of Sciences 109, 6241-6246.

SHANTZ, H.L., PIEMEISEL, R.L., 1917. Fungus fairy rings in Eastern Colorado and their effect on vegetation. Journal of Agricultural Research 11, 191-245.

STURGESS, P., 2020. Llanishen and Lisvane Reservoirs: Survey of grassland fungi 2019, APEM Scientific Report, v 2.0 Final ed. Dŵr Cymru Welsh Water, p. 33.

TEDERSOO, L., ANSLAN, S., BAHRAM, M., PÕLME, S., RIIT, T., LIIV, I., KÕLJALG, U., KISAND, V., NILSSON, H., HILDEBRAND, F., 2015. Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. MycoKeys 10, 1.

TEDERSOO, L., BAHRAM, M., PÕLME, S., KÕLJALG, U., YOROU, N.S., WIJESUNDERA, R., RUIZ, L.V., VASCO-PALACIOS, A.M., THU, P.Q., SUIJA, A., 2014. Global diversity and geography of soil fungi. Science 346, 1256688.

UPTON, R.N., CHECINSKA SIELAFF, A., HOFMOCKEL, K.S., XU, X., POLLEY, H.W., WILSEY, B.J., 2020. Soil depth and grassland origin cooperatively shape microbial community co-occurrence and function. Ecosphere 11, e02973.

VAN DER WAL, A., GEYDAN, T.D., KUYPER, T.W., DE BOER, W., 2013. A thready affair: linking fungal diversity and community dynamics to terrestrial decomposition processes. FEMS Microbiology Reviews 37, 477-494.

VĚTROVSKÝ, T., MORAIS, D., KOHOUT, P., LEPINAY, C., ALGORA, C., HOLLÁ, S.A., BAHNMANN, B.D., BÍLOHNĚDÁ, K., BRABCOVÁ, V., D'ALÒ, F., 2020. GlobalFungi, a global database of fungal occurrences from high-throughput-sequencing metabarcoding studies. Scientific Data 7, 1-14.

WANG, Q., GARRITY, G.M., TIEDJE, J.M., COLE, J.R., 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. Applied and Environmental Microbiology 73, 5261-5267.

WEIßBECKER, C., BUSCOT, F., WUBET, T., 2017. Preservation of nucleic acids by freeze-drying for next generation sequencing analyses of soil microbial communities. Journal of Plant Ecology 10, 81-90.

WHITE, T.J., BRUNS, T., LEE, S., TAYLOR, J., 1990a. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, in: INNIS, M., GELFAND, D., SNINSKY, J., WHITE, T. (Eds.), PCR Protocols: A guide to methods and applications. Academic Press, Orlando FL, pp. 315-322.

WHITE, T.J., BRUNS, T., LEE, S., TAYLOR, J., 1990b. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, in: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR protocols: a guide to methods and applications, pp. 315-322.

WILLIAMS, N., 2020. Fungal DNA and conservation. Conservation Land Management Spring 2020 | Vol. 18 No. 1 9 18, 9-15.

WOESE, C.R., FOX, G.E., 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proceedings of the National Academy of Sciences 74, 5088-5090.

WOESE, C.R., KANDLER, O., WHEELIS, M.L., 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proceedings of the National Academy of Sciences 87, 4576-4579.

WOODCOCK, P., 2020. Opportunities for using DNA in terrestrial monitoring, Report No. 661, ISSN 0963-8091, Peterborough.

ZOTTI, M., DE FILIPPIS, F., CESARANO, G., ERCOLINI, D., TESEI, G., ALLEGREZZA, M., GIANNINO, F., MAZZOLENI, S., BONANOMI, G., 2020. One ring to rule them all: an ecosystem engineer fungus fosters plant and microbial diversity in a Mediterranean grassland. New Phytol. 227, 884-898.

.

List of abbreviations

CHEGD - the group of macrofungi predominantly found in temperate grassland habitats. Acronym is derived from the initial letters of Clavariaceae, Hygrophoraceae, Entolomataceae, Geoglossaceae and *Dermoloma* spp.

eDNA - environmental DNA, obtained from habitats rather than the tissues of a single organisms

ITS - internal transcribed spacer region of the rRNA operon

IUCN - International Union for Conservation of Nature

LSU - Large SubUnit gene of the rRNA operon (often referred to as the 28S gene)

mRNA - messenger RNA (these RNA transcripts are translated by ribosomes to synthesis proteins)

OTU - Operational Taxonomic Unit (DNA equivalent of a species or unit of similar taxonomic rank)

rRNA -ribosomal RNA (these RNA molecules not encode proteins but rather are incorporated into ribosome

Further background information can be found in Jones et al. (2020)

Natural England is here to secure a healthy natural environment for people to enjoy, where wildlife is protected and England's traditional landscapes are safeguarded for future generations.

Natural England publications are available as accessible pdfs from www.gov.uk/natural-england.

Should an alternative format of this publication be required, please contact our enquiries line for more information: 0300 060 3900 or email enquiries@naturalengland.org.uk.

ISBN 978-1-78354-776-0

Catalogue code: NECR374

NATURAL ENGLAND