Llywodraeth Cymru
Welsh Government

1872 PRIFYSGOL ABERYSTWYTH UNIVERSITY 1872 150 2022

Higher Plant DNA Sequencing in Soil

November 2022

Final Report

Higher Plant DNA Sequencing in Soil (WG contract C343/2017/2018)

Lina Avila Clasen, Andrew P Detheridge, John Scullion and Gareth Wyn Griffith

IBERS, Aberystwyth University,
April 2022

Executive summary

The purpose of the Environment Impact Assessment (EIA) (Agriculture) (Wales) Regulations 2017 is to ensure that the natural and historic environment found in Wales is not adversely affected by agricultural practices. In practice this involves the on-site assessment of an intact grassland sward prior to works being carried out. The EIA (Agri.) Unit also investigates potential breaches in the Regulations, and can encounter recently cultivated areas, or fields consisting of bare soil.

Utilising the existing fungal eDNA methodologies, this study investigates whether the process could also detect higher plant tissues (roots) and residues in the soil. Thereby having the potential to aid in the evidence gathering of the EIA team. A process of this nature would also have further applications within the team, and could be used for example in out of season (winter months) habitat monitoring of remediation sites.

The effectiveness and accuracy of the eDNA metabarcoding was tested by comparison of DNA extracted from soil cores with standard botanical (% cover) assessments at three distinct grassland habitats. eDNA-based and % cover data were highly congruent and the eDNA method was superior at identification of grasses. However, some species present at very low abundance were missed.

The development of the fungal eDNA methodology has been successfully transposed to include data for higher plants, the primary endeavour of this project. However, it is not a 'Golden Bullet' which can be utilised in isolation, without additional testing across a wide range of 'real-world' samples from EIA investigations.

- The development of the fungal eDNA process has been successfully transposed to include data for higher plants, the primary endeavour of this project. The eDNA process does however have some limitations and cannot at present be utilised in isolation to prove that a breach of the Regulations has occurred, or that a Remediation site is returning back to its pre- intensification event quality.

- The eDNA process does perform comparably to more traditional surveying techniques, classifying plant communities to the same NVC groupings. However, care must be taken with factors such as windblown contamination, failure to detect species at very low abundance, and plant root architecture

potentially skewing results. However, the eDNA process consistently identified more species of grass than field surveyors.

- There is significant degradation of plant residues over time in the soil but some residues were detectable after 12 months. Thus intervention and soil collection should carried out as soon as possible after disturbance to give an accurate as possible representation of the previous sward.

- Specific quantification of 'key indicator' species using the more sensitive PCR method would very likely prove more effective on more historic soil samples, when more than 3-6 months has elapsed since the disturbance event.

- Data derived from fungal residues provide additional and distinctive indication of past management practices (lack of soil disturbance) even after 12 months has elapsed since the disturbance event.

- The current W-transect techniques used by the EIA Unit was found to be suitable for sample collection. The collection of 800 - 1000 g of soil is achievable in 30 mins, and provides a suitable size of sample for analysis.

- Inappropriate storage during transit and associated DNA degradation could alter the plant communities later discovered. Therefore, a range of storage conditions were tested and it was found that refrigerated storage for period of up to a fortnight led to only minimal changes. Samples should not be frozen prior to transport to analytical labs but rather stored refrigerated to preserve the DNA.

- Analysis of metabarcoding data for fungal populations in soils before and after disturbance identified several species that were reliable indicators of semi-natural grassland soil and also other species typical of highly disturbed soil.

INDEX

Introduction

The advent of molecular biology in the 1970s has transformed our understanding of biology. Such developments mainly impacted cell biology and medicine but the influence of DNA technology has also spread to ecology and taxonomy. In particular, the idea that biological species could be defined by DNA sequences in addition to their morphology is changing the way in which the natural world and ecological processes are studied.

The use of DNA sequences to identify species was pioneered by Carl Woese (1977) who used a small portion (5.8S region; 160 basepairs [bp] in length) of the ribosomal RNA (rRNA) locus to identify relationships between various bacterial species, and in the process of doing this to identify the existence of a whole new domain of life, the Archaea. His choice of the rRNA locus was an inspired one, since these genes are present in all living organisms, forming key components of the ribosomes, the protein factories of living cells. The invention of the polymerase chain reaction (PCR) by Kary Mullis (1986) permitted the specific multiplication (amplification) and sequencing of small amounts of DNA from complex mixtures using small synthetic DNA fragments (oligonucleotide primers) which match the DNA sequence at the two ends of the desired fragment. Sequencing of these fragments using the Sanger dideoxy chain termination method of DNA sequencing (Sanger et al., 1973) and comparison of DNA sequencing via various phylogenetic approaches was pioneered by Felsenstein (1988) and others.

White et al. (1990) designed a set of oligonucleotide primers to permit PCR amplification of the more variable internal transcribed spacer (ITS) region of the rRNA gene cluster. These were originally designed to permit identification of fungi from mycelium growing in Petri dishes based on similarity to DNA sequences from mushrooms but subsequently they have been widely used for many other eukaryotes. The White et al. (1990) ITS primers were designed to match conserved regions within rRNA genes which flank the ITS region and were therefore able to amplify the ITS region of a wide range of eukaryotes. Despite being amongst the earliest PCR primer sequences to be published, these same primers are still widely used today, including in the present study. The same approach was later used by fungal taxonomists to obtain DNA sequences from reference herbarium/fungarium samples in order to study the relatedness of different groups of fungi. All the new sequences are made publicly available on DNA databases (required by journals as a condition of publication), notably GenBank (at the National Centre for Biotechnology Information [NCBI], based in Bethesda, Maryland). The accumulation of ITS reference sequences linked to clearly identified herbarium samples has allowed progressively more accurate identification of biological structures of unknown provenance, for example fragments of mycelium/plant roots/meat samples etc.

The term DNA barcoding was introduced much later (Hebert et al., 2003) primarily in a zoological context and relying on PCR amplification of a different locus COI-cytochrome oxidase subunit I) in the mitochondrial genome. These sequences are also published in GenBank and provide the basis of molecular identification of animals. One problem with GenBank is that it is not curated and by now contains many sequences that are assigned incorrectly to particular species, as well as many with

limited metadata. Although GenBank has sought to address this problem by creation of the curated RefSeq database (Pruitt et al., 2005), several other curated databases have been created and are more widely used. For instance, BOLD (Barcode of Life Data System) (Ratnasingham and Hebert, 2007).

Whilst, BOLD provides good coverage for animals and to a lesser extent plants, coverage of other biota is poor. For bacteria/Archaea, databases such as the SILVA ribosomal RNA database (Pruesse et al., 2007). For non-animal eukaryotes there are similar sequence databases of curated ITS2 sequences, for example at Würzburg, Germany (ITS2 Databses V; http://its2.bioapps.biozentrum.uni-wuerzburg.de/ (Ankenbrand et al., 2015; Koetschan et al., 2009). For fungi, the most widely used is the UNITE database ((Koljalg et al., 2005); https://unite.ut.ee/) which has the advantage that us is continually curated and updated by a large cohort of taxonomic experts.

For DNA barcoding and metabarcoding of plants, early attempts to use the ITS 1/2 locus encountered problems with inadvertent PCR amplification of fungal DNA due to the widespread presence of fungal endophytes in plant tissues (Zhang et al., 1997). Due to low rates of sequence evolution in plant mitochondrial DNA genomes, attention focused on the development of DNA barcodes based on various loci within the chloroplast genome (Hollingsworth et al., 2009). Most widely used amongst these has been the *rbcL* locus and global efforts have led to the construction of reference databases with good taxonomic coverage (Bell et al., 2017).

Of particular relevance to the present study is the generation of rbcL barcodes for all the flowering plants of Wales (de Vere et al., 2012). However, the taxonomic resolution of *rbcL* (ability to distinguish closely related species) is poor for some groups of plants and it is now often used in combination with the ITS2 locus, with the problem of PCR amplification of non-plant sequences at the ITS locus being solved by use of more specific PCR primers (Moorhouse-Gann et al., 2018; Sickel et al., 2015).

The existence of reliable DNA barcodes and their deployment in eDNA metabarcoding has permitted advances in several areas of study, including the analysis of airborne pollen (Brennan et al., 2019) and diet reconstruction from animal faeces (Moorhouse-Gann et al., 2018; Reese et al., 2019).

These methods have also been deployed for analysis of plant eDNA in soil, with Matesanz et al. (2019) using synthetic ‘mock’ communities to quantify root biomass in Mediterranean soils and Fahner et al. (2016) using the method to assess plant species richness in boreal wetlands. One of the earliest attempts to deploy eDNA metabarcoding for the identification of plant biomass from soil was by Yoccoz et al. (2012) using a short (ca. 90 bp) amplicon region, the P6 loop of the plastid trnL (UAA) intron within the chloroplast genome. This approach had low taxonomic resolution (due to short target sequence) and poor coverage of this area in DNA databases but the authors did detect DNA from some species that had not been cultivated at the site for many decades. In broader soil ecology, Leff et al. (2018) used an eDNa metabarcoding approach to predict the structure of soil communities based on the plant community taxonomy.

Aims of the project.

For the enforcement of EIA regulations, it is necessary to be able to demonstrate unequivocally that a change of land use has occurred on a particular area. In the case of grassland or heathland habitats, this would involve confirming that the plant communities present (often suspected to have resulted from agricultural improvement) are new and that previously the area was differently vegetated. The evidence gathered must be scientifically sound, as regular legal challenges are made against the EIA team's decisions.

One way in which this could be achieved is to confirm the presence of residues of the previous plant community by detecting the presence in the soil of DNA of the species previously suspected to be present. If a field is subject to agricultural improvement, the residues of the previous plant communities will be present in the soil but expected eventually to decompose. The length of time such residues would persist is difficult to predict and will depend on factors such as the prevailing climatic conditions and the current agricultural regime.

We have previously pioneered the use of soil eDNA analysis for the identification of grasslands with high mycological diversity, mostly focusing on waxcaps and allied fungi that are associated with undisturbed pastures, not previously subject to intensive management (Griffith et al., 2019; Griffith et al., 2020). Our approach has been successfully used to assist in the notification of Sites of Special Scientific Interest (SSSIs), in planning applications and to better understand the ecology of these fungi (Williams, 2020).

Here we aim to deploy a similar approach for the detection of plant eDNA in soils. Whilst the protocols for soil sampling, DNA extraction and subsequent molecular biology and bioinformatics are similar to those we have previously used, different PCR primers and data analysis approaches need to be devised and tested. For fungal eDNA, we aim to detect DNA from living fungal mycelia, whereas in the present study the main aim is to detect DNA within residues of decomposing plant communities and thus it is necessary to determine how long DNA is detectable within these residues.

1) Development and testing of plant-specific PCR primers and plant RDP database

The eDNA metabarcoding method deployed in this project generated large numbers (millions) of short (300-400 bp) DNA sequences from DNA barcode loci. The volume of sequence data generated is huge and bioinformatic analysis is required to link these data to species names and to provide abundance data for each species.

In our previous work, mostly relating to fungi, we developed a bioinformatic pipeline for the analysis of data for fungi using the LSU D1 region and ITS2 region barcodes (Fig. 1.1). For the former locus, we used a database of fungal 28S (LSU) sequences collated as part of the RDP (Ribosomal Database Project) based at Michigan State University (https://rdp.cme.msu.edu/). This fungal RDP database contains 125,525 sequences (RDP Release 11, Update 5 :: September 30, 2016) and we also added a number of DNA sequences from reference samples of grassland fungi from the Aberystwyth University fungarium.

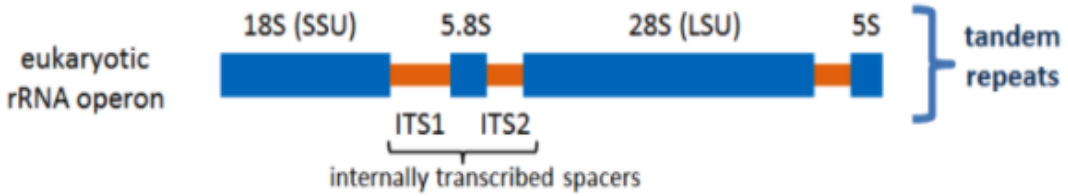


Fig. 1.1. Structure of the eukaryotic ribosomal RNA locus operon. Lengths of different regions vary between organisms (SSU: ca. 1850 bp; LSU: 3-5000 bp; ITS1/2: 200-350 bp each; 5.8S: ca. 160 bp; 5S: ca. 121 bp).

In recent years, we have also focused more on use of the adjacent ITS2 locus, which is now recognised as the primary DNA barcode for fungi (Schoch et al., 2012). Despite being slightly longer (400 bp vs 250 bp for the LSU locus and thus more challenging for PCR amplification and sequencing, the advantage of this locus is that there are many more reference DNA barcodes available and also it allows species-level identification across all fungal taxa. We have also collaborated with mycologist in Estonia who have created the curated UNITE Community database (https://unite.ut.ee/) which now contains >1.75 million sequences (latest release version 8.1; Last updated: 2019-09-21). An important advantage of this database is that it is possible for mycologist globally to participate in the curation of the database, for instance adding new species, revising outdated species concepts etc. (Kõljalg et al., 2013).

1A) Primer choice and optimisation

For higher plants, the primary DNA barcode is the *rbcL* gene located in the chloroplast region, with other chloroplast loci used as secondary barcodes (notably *matK* locus) (Hollingsworth et al., 2009). However, rather like the LSU locus for fungi, these to not always provide species level identification, and additionally both *rbcL* and *matK* amplicons are quite long (ca. 600 bp) which can create problems for PCR amplification and sequencing.

We therefore opted for the ITS2 as the barcoding locus of choice, using the widely used primer pairs ITS3chenS2F (ATGCGATACTTGGTGTGAAT) as the forward primer (in a conserved region of the 5.8S rRNA gene (Chen et al., 2010)) and ITS4 (TCCTCCGCTTATTGATATGC) which sits in a conserved region at the 5' end of the 28S (LSU) rRNA gene. The latter primer was one of the first ITS barcoding primers devised by White et al. (1990) (see above) but is known to be effective in amplification of plant ITS2 regions. Prior to choosing ITS3chenS2F as the forward primer, we also tested a new universal plant primer (TGTGAATTGCARRATYCMG) devised by Moorhouse-Gann et al. (2018) but found this to be less effective (poor amplification success relative to the Chen primers).

The use of the reverse ITS4 primer (which we also use for DNA barcoding of fungi) offered the possibility to simultaneously amplify barcodes for plants and fungi. The advantage of this is that the fungal barcodes provide useful contextual information for the plant species detected (and vice-versa). The potential disadvantage of this is that fewer plant sequences are obtained for each sample. However, the diversity of plants present at any given site is an order or magnitude less than for fungi. For a typical sequencing run, 20,000 sequence reads are obtained but 5000-10000 reads are sufficient to provide good sequence depth for the plant community present.

For the fungi we used an equimolar mix of six forward primers (ITS3NGS1, ITS3NGS2, ITS3NGS3, ITS3NGS4, ITS3NGS5 and ITS3NGS10), as recommended in the most wide-ranging fungal DNA metabarcoding study to date (Tedersoo et al., 2014). Use of this mix has been shown to yield efficient amplification across all fungal phyla and also to amplify the non-fungal oomycetes (*Phytophthora* spp. etc.).

In a series of test experiments, we used the Tedersoo fungal primer mix in various ratios with the ITS3chenS2F primer, finding that a molar ratio of 3:1 (ITS3chenS2F : total fungal primers) typically yielded about twice as many plant sequences as fungal (ca. 65-80%). This mixture was used for all the experiments reported here.

1B) Development of a higher plant RDP database based on ITS DNA barcodes

Since no similarly comprehensive database to UNITE exists for higher plants, we created a new database by downloading ITS2 sequences from GenBank using the search terms "Viridiplantae" [Organism] AND ("UK" [All fields] OR USA [All fields]) AND ("ITS2" [Definition] OR "internal transcribed spacer 2" [Definition]). This search yielded 83,618 sequences from GenBank which were downloaded to Geneious v10.

In order to create a database, it is also necessary to link the DNA sequences, GenBank accession IDs and species names to a taxonomic hierarchy. Within GenBank each species is attributed a taxonomic identifier field (db_xref). We used this to create a taxonomic classification hierarchy for each sequences, as illustrated below (Table 1.1).

Since the main GenBank database is not curated, it contains some misidentified or otherwise erroneous sequences (Fig. 1.2). During our curation 257 such erroneous sequences were removed from our database.

Additionally, for several plant taxa, individual species cannot be identified at species level using ITS2 sequences. Often this is the case due to the high rate of gene flow (hybridisation events) between closely related species. The consequence of this is that some taxa cannot reliably be assigned to single species, as exemplified by *Agrostis capillaris/gigantea* and *A. canina/stolonifera* (Fig. 1.3). For some taxa it is only possible to assign ITS2 sequences to genus level (e.g. *Pinus* spp.).

Table 1.1. Snapshot of the in-house RDP database created in the course of this project showing the structure of our in-house RDP database.

Accession	Kingdo	Phylum	Class	Order	Family	Genus	Species	Sequence	db_xref	Description	Organism
KM999962.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_stol_can	AGTGTGGGT(	taxon:63632	Agrostis stolonifera cultivar 007 clone till	Agrostis stolonifera
KM999963.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_stol_can	AGTGGGAGC	taxon:63632	Agrostis stolonifera cultivar 007 clone till	Agrostis stolonifera
KM999964.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_stol_can	GAGAGGTGA	taxon:63632	Agrostis stolonifera cultivar 007 clone till	Agrostis stolonifera
KM999965.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_stol_can	AGTGAGGTG	taxon:63632	Agrostis stolonifera cultivar 007 clone till	Agrostis stolonifera
FJ042802.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin02 internal trar	Agrostis vinealis
FJ042803.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin11 internal trar	Agrostis vinealis
FJ042804.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin07 internal trar	Agrostis vinealis
FJ042805.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin15 internal trar	Agrostis vinealis
FJ042806.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin18 internal trar	Agrostis vinealis
FJ042807.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin12 internal trar	Agrostis vinealis
FJ042808.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin01 internal trar	Agrostis vinealis
FJ042814.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin17 internal trar	Agrostis vinealis
FJ042815.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin08 internal trar	Agrostis vinealis
FJ042871.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Agrostis	Agrostis_vinealis	TCGTGACCC1	taxon:247443	Agrostis vinealis clone vin04 internal trar	Agrostis vinealis
KX166278.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Aira	Aira_caryophyllea	TGCAGAATC(	taxon:336211	Aira caryophyllea voucher NMW3168 5.8	Aira caryophyllea
EF153021.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Allolepis	Allolepis_texana	TCGTGACCC1	taxon:433842	Allolepis texana internal transcribed spac	Allolepis texana
GU359264.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Allolepis	Allolepis_texana	TGACCTGCG(	taxon:433842	Allolepis texana voucher US:Hitchcock 7	Allolepis texana
GU359265.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Allolepis	Allolepis_texana	TGACCTGCG(	taxon:433842	Allolepis texana voucher US:Le Roy 141	Allolepis texana
KM523749.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Alopecurus	Alopecurus_aequalis	TCGTGACCC1	taxon:114194	Alopecurus aequalis voucher US:Peters(	Alopecurus aequalis
KM523750.1	Viridipl	Strepto	Liliopsi	Poales	Poacea	Alopecurus	Alopecurus_arundinaceus	TCGTGACCC1	taxon:1227975	Alopecurus arundinaceus voucher CAN:(	Alopecurus arundinaceus

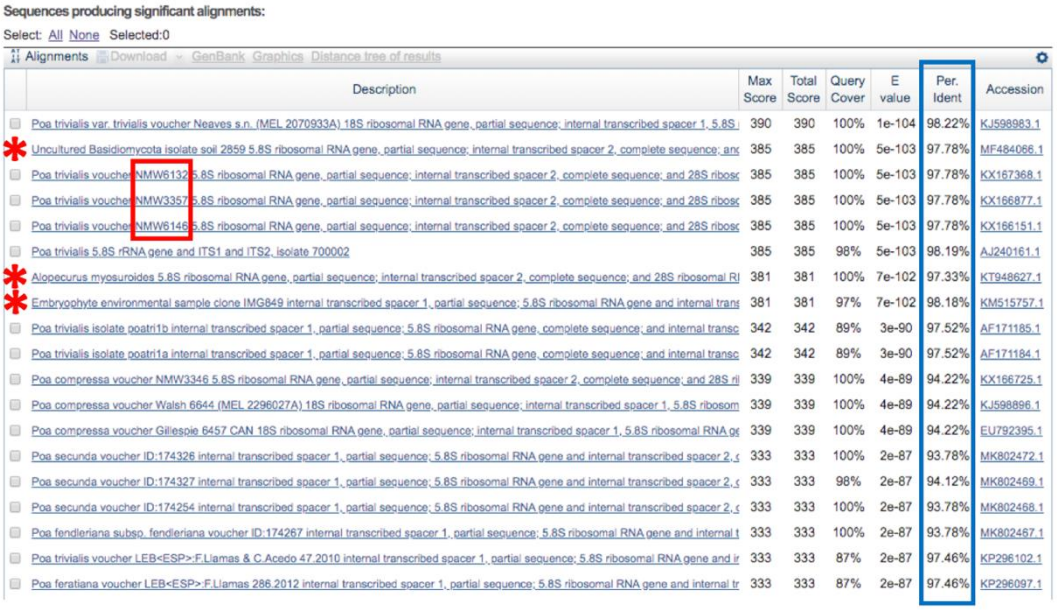


Fig 1.2. Output from a BLAST search of GenBank using a *Poa trivialis* ITS2 sequence from Brignant. Clearly incorrect or poorly identified hits are identified (red asterisks). Note how *P. trivialis* accession are >97% identical to the sample (blue square), whereas other closely related species (e.g. *Poa compressa*) are more dissimilar (ca. 94%). Note the reference samples from Wales (red square; NMW code) published in the landmark study by de Vere et al. (2012)

For initial testing of the RDP database, metabarcoding data was run through the pipeline and then all the taxon assignments were manually checked using BLAST (https://blast.ncbi.nlm.nih.gov). Where there was any uncertainty, closely related GenBank sequences identified via BLAST searches were downloaded to Geneious. These were aligned and used to create phylogenetic trees to confirm which taxa could be reliably separated at species level (as shown in Fig 1.3). Iteration of the process above progressively removed any incorrect or misleading identifications.

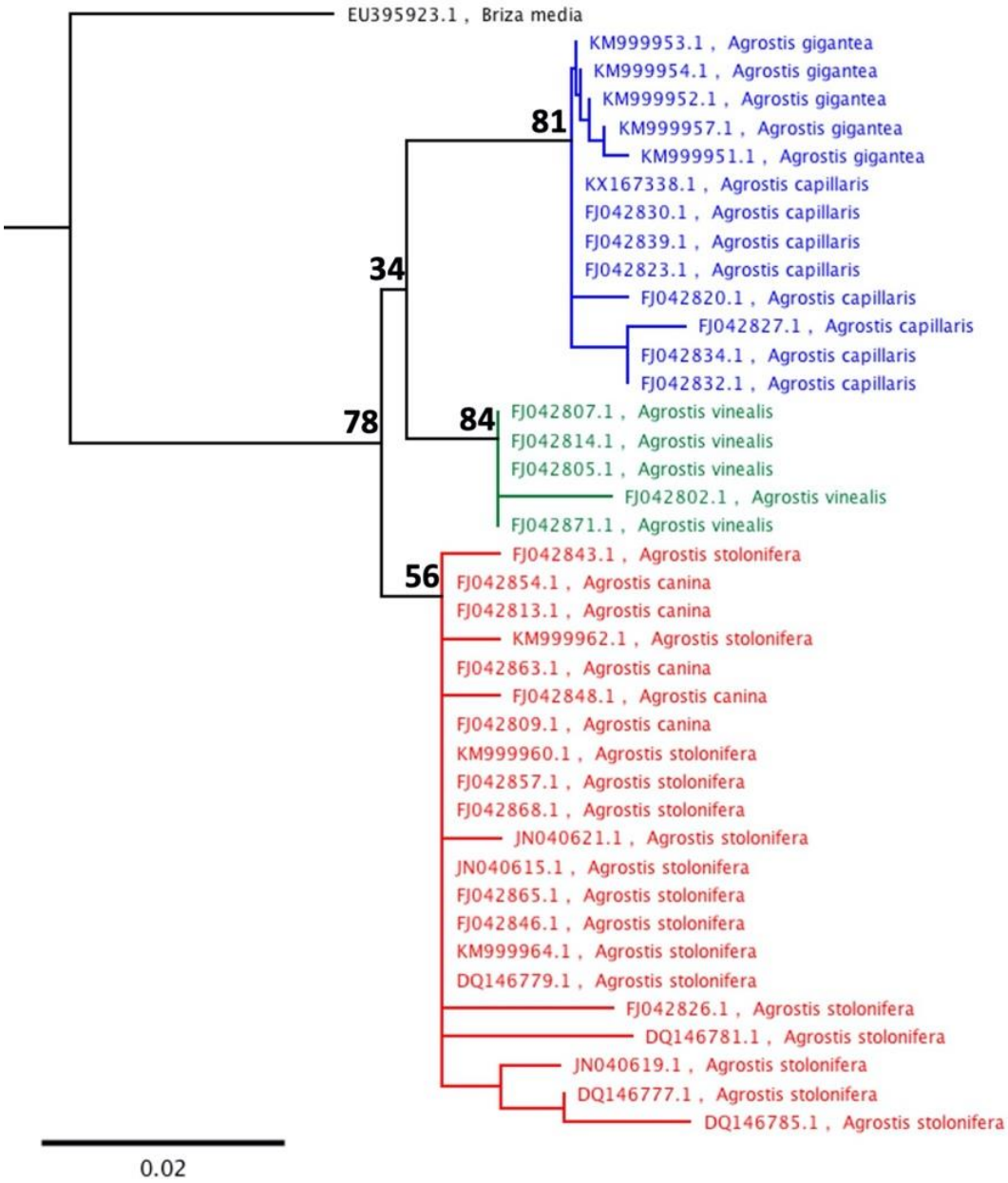


Fig. 1.3. Phylogenetic tree based on ITS2 sequences of UK members of the genus *Agrostis*. Note that *Agrostis capillaris* and *A. gigantea* are not separable from each other, nor are *A. canina* and *A. stolonifera*, most likely due to the high frequency of hybridisation event; whereas *A. vinealis* forms a distinct monospecific clade.

2) Comparison of plant eDNA for soils with NVC data

Having developed a protocol for the PCR amplification of plant eDNA from soil, we decided to test the effectiveness of our method in the assessment of plant communities from various grassland sites by comparison with traditional vegetation analysis, using the standard methodology used in NVC (National Vegetation Classification) method devised by Rodwell (1992). Three contrasting grassland sites were chosen: one an upland grassland in Wales (Brignant longterm grassland extensification experiment, Devil's Bridge), a lowland grassland in England (Kirby Muxloe, Leicestershire) and third a calcareous grassland in Ireland (Turlough National School, Co. Clare).

For each site, soil cores were collected in a ziplock bag, kept cool during transit and frozen within a few hours of collection. For consistency any leaves/stems present on the top of the core were cut off at soil level, such that the DNA later detected would be from living roots/tubers or plant necromass.

These bags were transferred to a -80°C freezer at Aberystwyth University and freeze-dried for several days. This process removes water via sublimation (from ice to steam), thus avoiding the formation of any liquid water in the drying process, preventing any biological degradation. It also renders soil very friable, and easy to grind. Grinding was undertaken by rubbing soil through a 2 mm sieve. The soil was thoroughly mixed at this stage and 50 g taken for further grinding through a 1 mm sieve, and then further thorough mixing. Sieves were washed in water and then bleach between samples and disposal latex gloves were worn for the grinding process, in order to minimise the possibility of any sample cross contamination.

A subsample of the finely ground soil (200 mg) was taken for eDNA extraction using a Qiagen PowerSoil Kit according manufactures instructions and subjected to PCR using the primer mix described below. PCR amplification was carried out using the following cycle: denaturing at 95°C for 3 minutes followed by 30 cycles consisting of denaturing 95°C 30 seconds; annealing 55°C 30 seconds; extension 72°C 30 seconds followed by a final extension step of 5 minutes. Success of PCR (judged by the presence of a band of expected size: 400-500 bp) was determined by running a sample on an agarose gel electrophoresis (Fig 2.1).

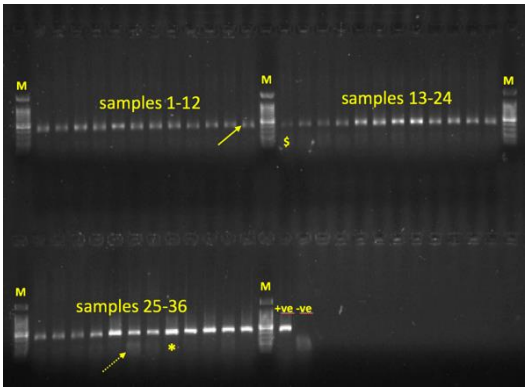


Fig 2.1. An agarose gel showing bands the ca. 450 bp ITS2 amplicon bands (solid arrow) from 36 eDNA PCR products. M indicates 'size marker (PCR Bio Ladder IV; bright band is 500 bp). Some bands were much stronger (*) than others ($). Positive (pure plant DNA) and negative (no DNA) controls are also shown. Smaller PCR products (dotted arrow), likely primer dimers, were removed by later purification.

The region 2 of the internal transcribed spacer (ITS) was the barcode region chosen to identify the plant and fungal species. The primers used for that were the ITS2 S2F ATGCGATACTTGGTGTGAAT (Chen et al. 2010), targeting plants; and for fungi a mix of 6 NGS primers (Tedersoo et al., 2014), as follows:

ITS3NGS1	CTAGACTCGTCATCGATGAAGAACGCAG	(ca. 95% of all fungi)
ITS3NGS2	CTAGACTCGTCAACGATGAAGAACGCAG	(Chytridiomycota)
ITS3NGS3	CTAGACTCGTCACCGATGAAGAACGCAG	(Sebacinales p.parte)
ITS3NGS4	CTAGACTCGTCATCGATGAAGAACGTAG	(Glomeromycota)
ITS3NGS5	CTAGACTCGTCATCGATGAAGAACGTGG	(Sordariales p.parte)
ITS3NGS10	CTAGACTCGTCATCGATGAAGAACGCTG	(Stramenopila [oomycetes])

The reverse primer in all cases was ITS4 (TCCTCCGCTTATTGATATGC). In order to demultiplex sequences after sequencing the ITS4 primer for each sample was synthesised as follows:
CCATCTCATCCCTGCGTGTCTCCGAC*TCAG*CTAAGGTAAC**TCCTCCGCTTATTG ATATGC** (this example is AD1X1ITS4 with the IX1 index).

…with the sequencing primer at the 5' (in red), followed by a 4 bp Key sequencer calibration sequence (italics), a 10-12 bp sample-specific index sequence (underline), followed by the ITS4 primer sequence itself (bold). The index sequences were those recommended by IonTorrent end using the IonExpress (IX) index sequences recommended by Ion Torrent. Thus for each sample destined to be sequenced in the same batch a different IX code was used (from IX1 to IX60), allowing the sequences to be demultiplexed after sequencing.

PCR products of all samples were quantified using a Qubit 2.0® Fluorometer. The samples were then pooled together at the same DNA concentration. An Ampure Clean for pooled libraries for Ion Torrent was used to ensure that any short fragments (usually the result of primer dimers; Fig 1.3, dotted arrow) were removed. The clean pool was then run in the Agilent 2100 Bioanalyzer (Santa Clara, CA) with an Agilent High Sensitivity DNA chip. The results from the Bioanalyzer were then used for the dilutions at a 50 pM target. The Ion Chef System was used to perform the library amplification, ISP recovery and enrichment and chip loading, using a Ion 316 chip for up to 400 bp read lengths. Finally the sequencer used was the Ion Torrent PGM System.

2A) Sequence data processing

Sequence data were quality checked, trimmed to 500 bp and demuliplexed using MOTHUR (v. 1.31.2; (Schloss et al., 2009)). Sequences with mismatching barcode primer sequences and short/poor quality sequences (length less than 100 bp and with an average Phred score less than 20) were discarded. Clustering and chimera removal was performed using the UPARSE pipeline via USEARCH v.9 (Edgar, 2013). Sequence files were dereplicated and singletons discarded as recommended in Tedersoo et al. (2010), and operational taxonomic units (OTUs) at 97% clustering; clusters less containing less than 2 sequences were discarded.

A taxonomy was assigned to each OTU (operational taxonomic unit) using the Naïve Bayesian Classifier (Wang et al., 2007). Where species was not assigned by the classifier (but only assigned to genus, family or order, due to confidence being lower that the threshold), an OTU identifier was assigned to that cluster. Data were then

rendered in Excel and standardized by dividing the number of reads in each taxonomic unit by the total number of fungal reads in each sample to give relative abundances of the assigned taxa for each quadrat; any non-fungal taxa were reported separately.

The R package was used for visualisations of relative abundance matrices using principal coordinate analysis (PCO) ordination on a Bray-Curtis distance matrix to identify patterns in the data. Permutation multivariate analysis of variance (PERMANOVA) was used to determine overall significant differences in community data by soil and grass and was performed in PRIMER 6 & PERMANOVA+ (versions 6.1.12 and 1.0.2 respectively; Primer-E, Ivybridge, UK). Abundance percentage data were subjected to square root transformation and Bray-Curtis distance matrices calculated. PERMANOVA was carried out using default settings with 9999 unrestricted permutations. Analysis of Similarity (ANOSIM) was carried out in PRIMER 6 & PERMANOVA+ using the Bray-Curtis distance matrix calculated above

2B) Brignant NVC vs eDNA data

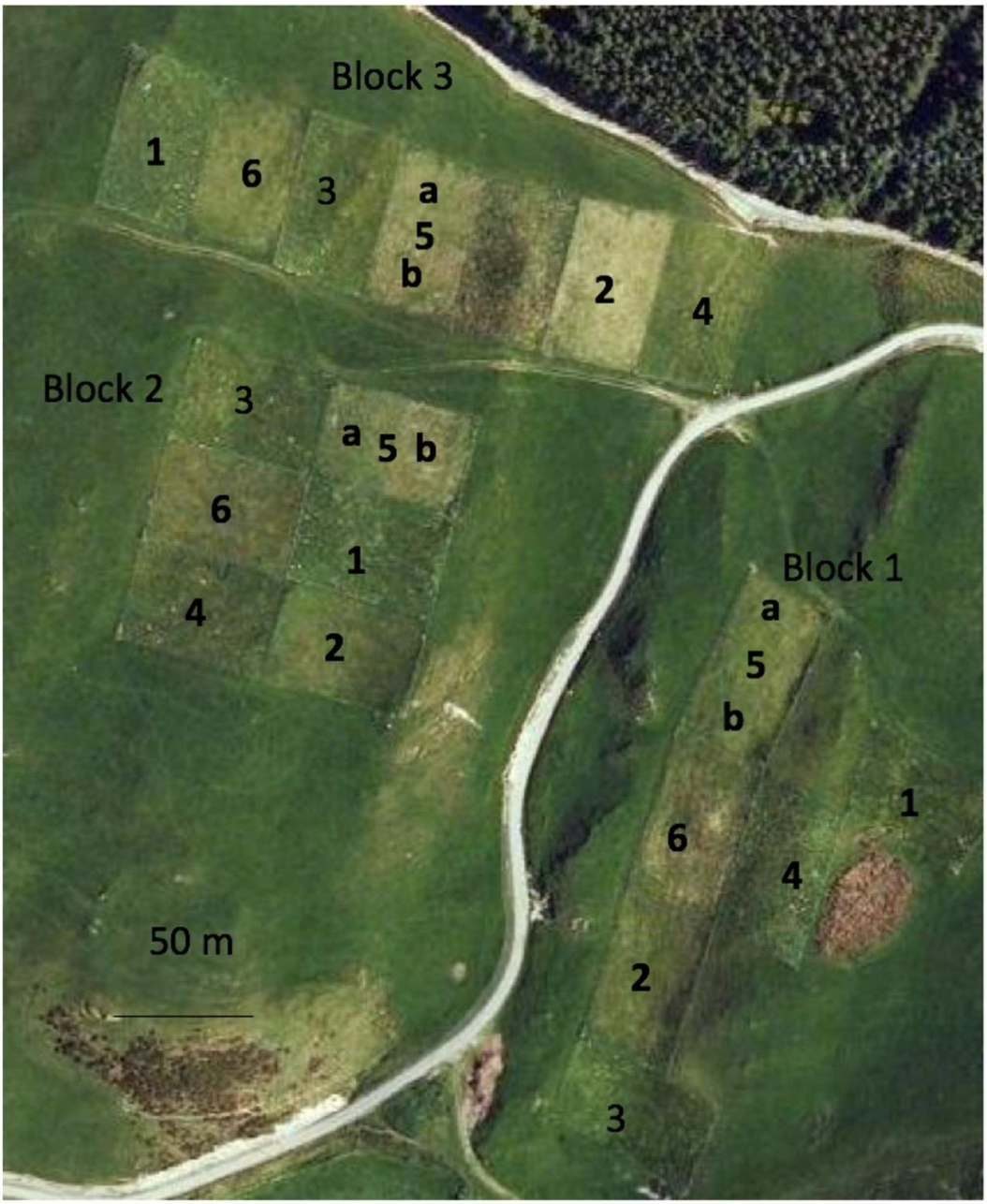
The Brignant longterm grassland extensification experiment near Devil's Bridge, Aberystwyth is part of the Pwllpeiran Upland Research platform managed by IBERS. GWG has had longterm involvement in the management of this experiment which was established by Mike Hayes under DEFRA funding in 1994. The experiment comprises triplicated plots in three blocks subject to seven contrasting agricultural treatments. These involve addition (or not) of fertilisers or lime, and different grazing regimes (summer sheep grazing and haycutting with or without aftermath grazing).

Rather than attempt to undertake vegetation surveys across all treatments, the focus here was on the two most contrasting treatments (P1: fertiliser and lime addition plus summer grazing (the most intensive treatments ["business as usual treatment"] vs P6 (the most extensive treatment; no amendments, haycut with aftermath grazing). To account for the possibility if variation within plots, three 2 x 2m quadrats (Q1-Q3) were set within each plot of each block (B1-B3; Fig. 2.2), a total of nine quadrats per treatment.

The vegetation survey was undertaken on 20th July 2018 by Dr. Rob Rowlands (Director at Ecology Solutions, Manchester) an ecologist with >20 years experience of vegetation surveying for EIA investigations. For each quadrat he identified all the higher plants present and estimated percentage cover.

For each quadrat (on the same day), nine soil cores (15mm diameter and 10 cm depth; ca. 80cm between each core) were taken and pooled. These were frozen at -80ºC and processed as described above.

The weather had been very dry in June/July 2018 so even at Brignant (375m above sea level; annual rainfall 1800 mm/yr), vegetation was very dry and this made species identification more challenging.



Treatment	Fertiliser	Lime (98/16)	Hay Cut	Grazing
1	Y	Y	N	Apr-Nov
2	N	Y	July	Aug-Nov
3	N	N	N	Apr-Nov
4	N	Y	N	Apr-Nov
5a	N	Y	July	N
5b	N	N	July	N
6	N	N	July	Aug-Nov

Fig. 2.2. Map of the Brignant experiment and details of the treatment regimes. Fertiliser 60 kg N ha^{-1} and 30 kg P ha^{-1}, applied each May. Lime was applied in May 1998 to raise pH to 6, and this treatment was repeated in May 2016.

For the eDNA analysis, 79% of reads were from plants, and of these 0.36% (mean; range 0.06%-0.97%) were assigned to green algae (phylum Chlorophyta) and lower plants (Bryopsida [mosses] or Jungermanniopsida [liverworts]). The number of plant sequence reads per sample ranged from 22700 to 57863 (mean 35,628).

In the context of Table 2.1, sequences at very low abundance, for example 0.01%, would be represented by 2 to 5 sequences, with unique sequences occurring only once ["singletons"] being discarded during the quality control processing. Potentially species detected at only very low abundance could have originated from ungerminated seeds or fragments of windblow leaf tissue and would thus not be considered as members of this particular grassland community. It would be possible to set a threshold abundance level, where species are only counted as being present when they exceed this level but it is not clear how such a threshold could be set objectively. Therefore, here all species detected from two or more sequence reads are counted as present.

Data from eDNA and vegetation were merged into a single spreadsheet, showing percentage abundance of each species in each quadrat (Table 2.1). A total of 41 species were identified across all plots using both methods. Of these 35 species were detected by eDNA, 31 spp. by vegetation analysis and 25 spp. by both methods. Thus the eDNA analysis missed 6 species and the vegetation analysis missed 10 spp.

Principal coordinate ordination of the survey datasets (Fig 2.3) showed that the plant communities in the two treatment regimes were clearly separated by both the eDNA (prefix B) and vegetation analysis (prefix Y [Plot1: intensive] or Y [Plot6: extensive]). The ordination derived for the same plots analysed by different methods do not overlay each other but they are adjacent. It is likely that this is due to the relative abundance data obtained from the two methods is often very different. That this is the case is not surprising since one might expect differences in relative abundance of root and leaf/stem biomass to differ between species.

Table 2.1: eDNA and vegetation analysis of two treatments at the Brignant longterm experiment

					Treatment 1 (Summer grazing, fertiliser and lime)																		Treatment 6 (Haycut/ aftermath grazing, no additions)																
					Soil eDNA										Vegetation analysis								Soil eDNA										Vegetation analysis								
					Block1			Block2			Block3				Block1			Block2			Block3			Block1			Block2			Block3			Block1			Block2			Block3		
		No. species detected			Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3		Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3
No.	Species	eDNA	NVC	Both	B1P1Q1	B1P1Q2	B1P1Q3	B2P1Q1	B2P1Q2	B2P1Q3	B3P1Q1	B3P1Q2	B3P1Q3	X111	X112	X113	X211	X212	X213	X311	X312	X313	B1P6Q1	B1P6Q2	B1P6Q3	B2P6Q1	B2P6Q2	B2P6Q3	B3P6Q1	B3P6Q2	B3P6Q3	Y111	Y112	Y113	Y211	Y212	Y213	Y311	Y312	Y313
1	*Achillea millefolium*	1	1	1																				0.19%									1%	1%	1%						
2	*Agrostis capillaris/gigantea*	1	1	1	6.35%	5.20%	12.37%	3.84%	3.93%	5.96%	10.36%	11.12%	10.30%	30%	10%	10%	35%	30%	25%	45%	45%	45%	0.61%	0.54%	1.00%	0.39%	2.96%	2.42%	2.72%	5.05%	0.19%	5%	5%	5%	15%	5%	5%	20%	20%	5%
3	*Agrostis stolonifera/canina*	1	1	1	0.08%	0.02%	0.06%	0.63%	0.01%	0.46%	0.05%	0.07%	0.03%			1%	5%										0.04%	0.01%		0.01%										
4	*Alopecurus pratensis*	1						0.85%			5.73%	23.73%	0.63%										0.31%																	
5	*Anthoxanthum odoratum*	1	1	1		1.27%	1.64%		0.01%	0.01%	5.61%		2.25%		20%	10%							3.01%	0.83%	1.94%	0.22%	8.85%	14.08%	2.61%	4.51%	0.68%	15%	25%	35%	5%	35%	25%	15%	15%	5%
6	*Betonica officinalis*	1	1	1																					0.07%									5%	5%					
7	*Campanula rotundifolia*		1																															1%					
8	*Cardamine pratensis*	1	1	1	0.56%	0.34%	26.51%	62.89%	0.18%		3.71%	0.03%	1.18%	1%		1%	1%						0.01%			0.42%	1.85%	0.98%		0.04%					1%			1%		1%
9	*Centaurea nigra*	1	1	1						0.01%	0.55%												1.07%		12.14%	0.01%					6.07%	10%	15%	15%	1%					25%
10	*Cerastium glomeratum*	1	1	1																			0.20%			0.01%		0.01%				1%			1%	1%	1%	1%	1%	1%
11	*Cirsium arvense*	1	1	1	2.56%									5%	1%				1%																		1%			
12	*Cirsium vulgare*	1	1	1		0.72%																											1%						
13	*Crepis capillaris*	1																					0.02%	0.04%	0.03%		0.03%				0.57%								
14	*Cynosurus cristatus*	1	1	1																			0.85%	0.04%		0.33%		0.04%				15%	15%	15%	20%	5%	5%		5%	1%
15	*Euphrasia officinalis*		1																													5%	1%	1%		5%	5%		1%	1%
16	*Festuca ovina*	1	1	1	0.01%	0.04%	0.01%	0.01%	0.01%	0.01%	0.02%	0.03%	0.51%	35%	35%	35%	55%	60%	60%	45%	45%	45%				0.02%						5%	5%	5%	25%	10%	15%			
17	*Festuca rubra*	1			13.60%	11.81%	11.87%	6.89%	8.71%	13.01%	17.89%	22.37%	22.53%										0.81%	1.38%	0.46%	2.19%	0.46%	2.03%		0.36%	0.22%									
18	*Galium saxatile*		1																															1%					
19	*Holcus lanatus*	1	1	1	15.97%	7.39%	8.94%	1.19%	0.27%	12.20%	2.39%	2.43%	0.38%	35%	35%	35%	10%	10%	15%	10%	10%	10%	2.35%	0.12%	0.25%	0.26%	0.87%	2.53%	0.12%	1.70%	0.30%	5%	5%	5%	5%	5%	10%	5%	5%	1%
20	*Hypochaeris radicata*	1	1	1			0.02%																0.23%	7.33%	9.10%	45.97%	4.06%	0.38%	24.94%	10.37%	11.21%	1%	5%	5%	10%	5%	10%	15%	10%	10%
21	*Leontodon hispidus*	1																													0.44%								
22	*Lolium perenne/multiflorum*	1	1	1	0.04%	0.55%	6.47%				0.05%				1%	10%												0.01%											
23	*Medicago lupulina*		1																													1%						
24	*Myosotis sp.*		1																													1%	1%					
25	*Phleum pratense*	1								0.31%														0.01%															
26	*Plantago lanceolata*	1	1	1				0.02%		0.54%	0.03%	0.02%											61.55%	79.54%	74.36%	40.36%	52.74%	21.48%	20.56%	72.52%	49.30%	25%	20%	20%	15%	15%	15%	20%	20%	10%
27	*Poa pratensis*	1			5.63%	7.51%	3.18%	10.89%	12.74%	26.80%	17.85%	18.55%	30.02%										0.06%	0.13%	0.17%	0.54%	0.54%	1.88%	0.03%	0.02%										
28	*Poa trivialis*	1			49.89%	63.02%	26.94%	12.47%	53.73%	40.10%	35.20%	21.01%	30.47%										0.78%			0.07%		0.17%	0.01%										
29	*Potentilla erecta*		1																															1%					
30	*Prunella vulgaris*	1																								0.52%													
31	*Ranunculus acris*	1	1	1	0.01%		0.01%		16.44%	0.02%				1%		1%			1%				14.02%	6.83%	0.25%	1.61%	13.99%	25.89%	45.83%	5.04%	24.74%	5%		1%		1%	1%	10%	10%	5%
32	*Ranunculus repens/bulbosus*	1			4.55%	0.80%	0.86%																2.00%	0.27%			0.08%	4.29%	0.40%	0.02%	0.10%								
33	*Rhinanthus minor*	1	1	1																			0.05%		0.01%	0.03%		0.04%				5%	5%	5%	5%	10%	15%			
34	*Rumex acetosa*	1	1	1	0.01%																		0.10%		0.01%	6.52%	3.51%	3.57%	0.41%	0.03%			1%	1%						
35	*Scorzoneroides autumnalis*	1	1	1																			0.06%	2.03%	0.02%	0.02%	5.83%	14.13%	2.14%	0.24%	6.02%					5%	5%			
36	*Stellaria alsine*	1	1	1	0.10%		0.72%																										1%						
37	*Taraxacum officinale agg.*	1	1	1									1%										0.02%	0.01%	0.02%		3.75%	0.14%	0.02%	0.03%						1%	1%		
38	*Trifolium dubium*	1																					11.37%	0.02%		0.03%													
39	*Trifolium repens*	1	1	1		1.23%	0.03%							1%	1%			1%					0.32%	0.56%		0.20%	0.14%	5.72%				10%	5%					1%		
40	*Urtica dioica*	1	1	1	0.01%				3.38%				0.71%					1%																					
41	*Veronica serpyllifolia*	1	1	1																				0.05%															1%
	Species count	**35**	**31**	**25**	**15**	**13**	**15**	**10**	**11**	**12**	**13**	**10**	**11**	**7**	**7**	**8**	**5**	**5**	**5**	**3**	**4**	**3**	**22**	**18**	**15**	**20**	**16**	**20**	**12**	**14**	**12**	**15**	**16**	**18**	**12**	**13**	**14**	**8**	**10**	**12**
	Total percentage cover				**99.39%**	**99.91%**	**99.62%**	**99.67%**	**99.39%**	**99.42%**	**99.44%**	**99.34%**	**99.03%**	**108.0%**	**103.0%**	**103.0%**	**106.0%**	**102.0%**	**102.0%**	**100.0%**	**101.0%**	**100.0%**	**99.81%**	**99.92%**	**99.83%**	**99.72%**	**99.71%**	**99.81%**	**99.80%**	**99.94%**	**99.85%**	**109.0%**	**111.0%**	**123.0%**	**108.0%**	**103.0%**	**114.0%**	**87.0%**	**88.0%**	**66.0%**

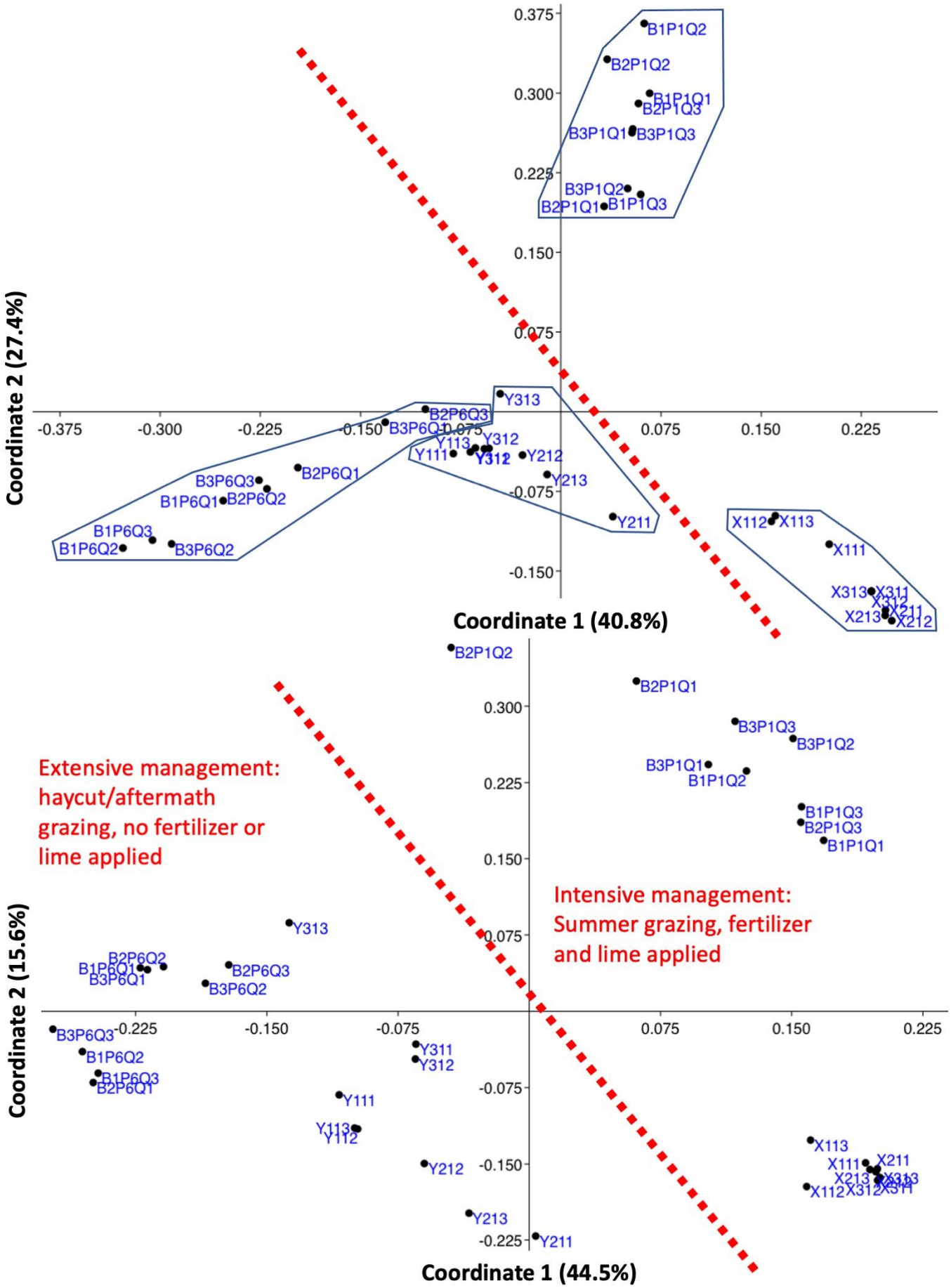


Fig 2.3: Principal Coordinates ordination plots of plant communities in Plots 1 and 6 at the Brignant longterm experiment using eDNA (B.....) or vegetation surveying (X/Y...). A) including all higher plants and B) including only species found detected using both methods.

Explanation of ordination plots

This explanation is provided to assist in the interpretation of the various ordination plots presented in this report [e.g. Fig. 2.3] which aim to summarise graphically the plant/fungal relative abundance data presented in the associated tables [e.g. Table 2.1].

In ecology, when collecting data for plant species present and their relative abundances in several quadrats, ordination plots provide an accessible means of presenting an overview of the data, to indicate which quadrats are most similar to each other. Each quadrat in the experiment is represented by a single point, with points plotted closely together having more similar plant/fungal communities.

A range of methods can be used to generate ordination plots. Here we use Principal Coordinates Analysis (PCO), a Multidimensional scaling (MDS) method widely used for presenting abundance data from natural communities of fungi/plants etc.

Ordination diagrams are straightforward to interpret, despite the complex maths that underlies them. Several ordination methods have been developed and are widely used by ecologists/biologists to simplify the interpretation of complex datasets. For such datasets the data are multidimensional and cannot be presented in a simple 2-D graph. So these methods distil the multidimensional data to extract the key differences (variances) between samples to present these in the 2-D ordination plots. Not all the information present in the dataset can be presented in ordination plots, since only the two most important components of variance can be presented. However, it is useful to note from the axes what percentage of the total variance in the data set is presented on each axis of the ordination graph.

Since only 25 of the 41 species found were detected using both methods, a second ordination using only these species was conducted (Fig 2.3B) but the resulting ordination plot was very similar.

As noted above, some plant taxa cannot be identified to species level using eDNA analysis of the ITS2 locus and at Brignant this was the case for *Agrostis capillaris/gigantea*, *Agrostis stolonifera/canina*, *Lolium perenne/multiflorum* and *Ranunculus repens/bulbosus*. Where both were recorded in vegetation analysis the total percentage was summed in Table 2.1.

The following six species were not detected via eDNA but were detected in the vegetation survey. In all cases they represent only a small percentage of total above-ground plant biomass, and with the exception of *Euphrasia*, were only present in a small proportion of the quadrats:-

Campanula rotundifolia (1%, 1 of 9 quadrats in plot6),
Euphrasia officinalis (1-5%, 7 of 9 quadrats in plot6),
Galium saxatile (1%, 1 of 9 quadrats in plot6),
Medicago lupulina (1%, 1 of 9 quadrats in plot6),
Myosotis sp. (1%, 2 of 9 quadrats in plot6),
Potentilla erecta (1%, 1 of 9 quadrats in plot6)

The following 10 species were not detected in the vegetation survey but detected via eDNA. The majority are grasses, especially those not flowering at the time of the

survey (which would likely be the case for all the grasses except *F. rubra*). For the other species, these could have been small seedlings overlooked during the survey; however, the possibility that the detected DNA originated from ungerminated seeds cannot be excluded. An additional (more likely possibility) is that the DNA originated from dead leaves or roots present in the soil:-

Alopecurus pratensis (<24%%, 5/18 quadrat in plots 1/6)
Crepis capillaris (<0.6%, 5/9 quadrat in plot6)
Festuca rubra (8-22%%, 17/18 quadrat in plot 1/6)
Leontodon hispidus (0.44%, 1/9 quadrat in plot6)
Phleum pratense (<0.31%, 2/18 quadrat in plot 1/6)
Poa pratensis (<31%, 17/18 quadrat in plot 1/6)
Poa trivialis (<63%, 13/18 quadrat in plot 1/6)
Prunella vulgaris (0.52%, 1/9 quadrat in plot6)
Ranunculus repens/bulbosus (<4.55%, 10/18 quadrat in plot 1/6)
Trifolium dubium (<11%, 3/9 quadrat in plot 6)

The relative abundance of different plant species is the likely explanation for the slightly different ordination of plots analysed by eDNA vs. vegetation analysis. For grasses, it is likely that estimation of which leaves belong to which species is difficult even when the species are in full flower but even for unambiguously identified species there are still large disparities. For example, *Plantago lanceolata* was detected at an a mean abundance of 52% (range 21-80%) via eDNA but only 18% (range 10-25%) in vegetation analysis. Similarly *Ranunculus* spp. were scored at 5% mean cover (range 1-10%) but 15% (range 0.3%-46%) via eDNA. Such disparities may be due to the fact the root:shoot ratios of different plant species can be divergent and this ratio changes seasonally.

2C) Kirby Muxloe NVC vs eDNA data

A similar comparison of eDNA versus vegetation analysis was undertaken using soil samples originally taken as part of a site assessment by Environmental Development Partnership (EDP; https://www.edp-uk.co.uk/), a company of which Dr. Rob Rowlands was previously a director.

Soil cores samples (5 cores per 2 x 2 m quadrat; using an apple corer) were obtained from four adjacent fields (labelled A to D) near Kirby Muxloe, Leicestershire (52.629,-1.233) on 3rd June 2013 by Mr. William Brown (EDP) and a vegetation survey was undertaken by Dr. Grace O'Donovan (Broadview Ecology Ltd.) at the same time. The fields were reported to be quite similar in vegetation except that field C was slightly wetter, with field A and C managed as haymeadows and field B and D grazed all summer.

A total of 34 spp. were detected during the original vegetation analysis and 35 spp. via eDNA analysis. Of the 43 spp. in total, 26 spp. were detected by both methods (Table 2.2).

Table 2.2: eDNA and vegetation analysis of four fields at Kirby Muxloe.

		eDNA																Vegetation analysis											
		No. species detected			FieldA		Field B			Field C						Field D			FieldA		Field B			Field C						Field D		
	Species	eDNA	NVC	Both	KX18A	KX19A	KX14B	KX15B	KX16B	KX01C	KX02C	KX03C	KX04C	KX05C	KX06C	KX10D	KX11D	KX12D	Q18A	Q19A	Q14B	Q15B	Q16B	Q01C	Q02C	Q03C	Q04C	Q05C	Q06C	Q10D	Q11D	Q12D
1	*Achillea millefolium*		1																											10%	
2	*Agrostis capillaris/gigantea*	1			5.00%	6.86%	3.35%	0.83%	0.23%	39.55%	26.84%	4.77%	7.26%	7.07%	9.89%	7.16%	4.74%	2.33%														
3	*Agrostis stolonifera/canina*	1	1	1	3.68%		0.01%			4.41%	1.32%	0.01%	0.20%	0.05%	0.07%	0.96%		0.57%	40%				10%	50%	40%					60%		60%
4	*Alopecurus pratensis*	1	1	1		0.03%	8.25%	0.70%	0.57%	19.09%	25.68%	1.58%		8.24%	1.50%	2.41%		9.28%			5%	50%	5%	2%	10%	7%	2%	5%		10%		10%
5	*Anthoxanthum odoratum*	1	1	1	10.91%		1.23%	0.52%	0.17%	5.91%	12.49%	23.26%	21.55%	30.12%	5.20%	4.51%		0.60%	5%	20%		15%	15%	20%	15%	20%	60%	60%	50%		30%	
6	*Arrhenatherum elatius*	1						0.10%					8.12%	0.02%	0.88%																
7	*Bellis perennis*		1																											2%	
8	*Bromus hordeaceus*	1							0.09%																						
9	*Centaurea nigra*	1	1	1	0.20%	0.33%									4.51%				7%				3%					5%			
10	*Cerastium fontanum/glomeratum*	1	1	1	2.24%	14.21%	0.34%	0.46%							0.01%	2.10%			3%		3%	1%	1%						1%	2%	
11	*Cirsium arvense*		1																									2%	2%		
12	*Cynosurus cristatus*	1	1	1	0.02%		0.33%	0.01%						0.07%	2.03%	0.13%	0.28%			5%									5%		5%	
13	*Dactylis glomerata*	1	1	1					0.05%						0.16%							1%									
14	*Deschampsia caespitosa*		1																						2%	2%		2%	10%		
15	*Elymus repens*	1			2.98%		0.18%	0.10%	0.13%	0.42%	0.55%	0.02%	3.18%	0.48%	0.90%																
16	*Festuca ovina*	1	1	1	2.24%	5.46%	1.34%	1.59%	0.04%	0.01%	0.02%	0.02%		0.25%	0.03%	1.46%		2.68%	20%			5%								5%		10%
17	*Festuca rubra*		1																	10%	35%			10%	15%	10%		2%	2%			
18	*Heracleum sphondylium*	1	1	1	11.12%			9.77%												5%	5%		1%								
19	*Holcus lanatus*	1	1	1	4.84%	21.83%	1.80%	0.65%	2.23%	5.26%	22.80%	65.65%	18.22%	16.74%	1.17%	0.85%	0.21%		10%			5%	3%	1%		1%	10%	30%	10%	5%		5%
20	*Hypochaeris radicata*	1	1	1	2.16%	14.74%													2%	2%											
21	*Juncus effusus*	1	1	1							0.06%													20%	25%	20%						
22	*Lolium perenne/multiflorum*	1	1	1	5.56%	4.20%	0.04%		3.83%			0.02%		0.02%			2.17%			20%	2%		5%					5%			2%	
23	*Lotus corniculatus*	1	1	1			0.29%			0.19%	0.07%				0.01%	38.23%	44.54%	0.02%											40%	20%	
24	*Luzula campestris*		1																								2%				
25	*Medicago lupulina*		1																1%												
26	*Phleum pratense*	1											20.94%	0.02%	5.39%																
27	*Plantago lanceolata*	1	1	1			1.97%		0.19%					0.81%	0.82%		4.64%	0.14%					3%				10%				30%	1%
28	*Poa pratensis*	1	1	1		0.02%	0.76%									0.05%		0.02%	5%			10%	25%								10%
29	*Poa trivialis*	1			12.02%	0.24%	0.29%		57.19%							0.01%		0.17%													
30	Populus x nigra	1			0.21%																										
31	Potentilla reptans	1	1	1										0.02%				0.36%			5%										25%
32	*Ranunculus acris*	1	1	1	1.24%	2.42%	0.13%	15.98%	24.35%	18.43%		0.19%	0.70%	5.28%	0.03%	5.83%	0.09%	7.27%	15%	10%		5%	20%		2%	1%	5%	5%	10%	5%	2%	
33	*Ranunculus repens/bulbosus*	1	1	1	4.32%	0.78%		0.01%	0.02%	0.05%			6.39%			1.42%	0.04%	0.06%		20%				5%	1%	5%						40%
34	*Rhinanthus minor*	1	1	1	0.41%									7.90%	6.03%					20%									20%			
35	*Rumex acetosa*	1	1	1	0.28%	14.61%	66.35%	0.91%	6.37%	6.02%	8.91%	3.14%	9.73%	16.73%	12.07%	0.77%		16.90%	10%	5%	5%	2%		5%	5%	7%	4%	4%	5%	5%		5%
36	*Schedonorus arundinaceus /pratensis*	1	1	1	0.10%	0.10%	11.91%		0.03%		0.15%	0.80%	3.51%	5.74%	0.08%						15%	10%	5%				7%				
37	*Scorzoneroides autumnalis*	1															3.64%														
38	*Sonchus* sp.		1																				1%								
39	*Stellaria graminea*	1	1	1			0.54%			0.01%	0.01%							54.75%				2%				3%		2%			3%
40	*Taraxacum officinale agg.*	1	1	1	0.27%	4.08%		68.35%	0.03%		0.01%	0.01%			1.83%		0.36%			5%		2%	1%			2%	2%	8%	10%			
41	*Trifolium pratense*	1	1	1					2.35%					0.08%	47.22%		7.21%			10%		1%	10%						5%	5%		
42	*Trifolium repens*	1	1	1		10.00%	0.01%		0.02%	0.09%	0.05%		0.02%			34.07%	31.93%	4.61%					2%							5%	5%	
43	*Vicia hirsuta*	1			29.52%																										
		35	34	26																											

As at Brignant, the grasses are better identified and more species detected via eDNA, In contrast, the vegetation survey identified the presence of several of the low abundance herbs that were missed by the eDNA (e.g. *Achillea millefolium*, *Bellis perennis*, *Medicago lupulina*). One species *Populus x nigra* was detected via eDNA in a single quadrat. This is very likely an example of where windblown leaves from adjacent trees are present in the soil. In contrast, the disparity between the eDNA and the vegetation survey data is less apparent on the ordination (Fig. 2.4) than was the case for Brignant data (Fig. 2.3).

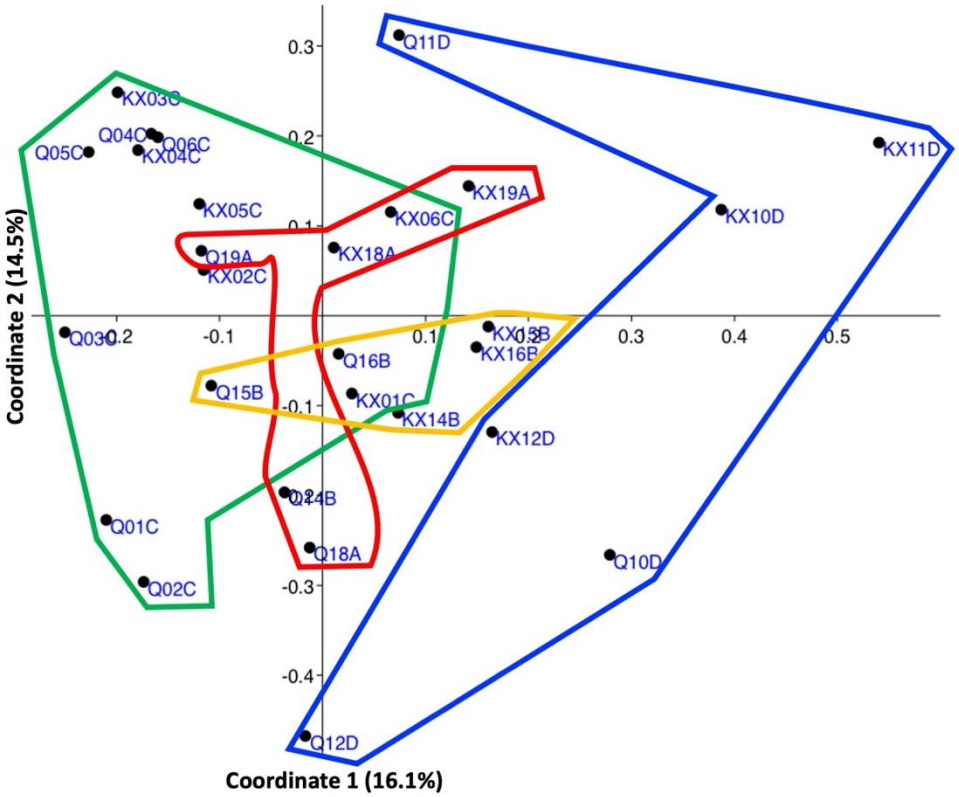


Fig. 2.4. PCO ordination of Kirby Muxloe quadrats. Samples from field A (red), B (orange), C green) and D (blue) are indicated by polygons. Datapoints prefixed by KX are from eDNA data and those from the vegetation survey are indicated by Q.

2D) Turlough

A third comparison of NVC survey data with eDNA was conducted on a species-rich lawn at Turlough, Co. Clare, Ireland (as part of a university fieldtrip class experiment) on 10th September 2018. Seven 50x50 cm quadrats were set across the lawn (within a ca. 3x8 m area) and undergraduate student assessed the vegetation in these quadrats using both percentage cover and point cover methods. After the vegetation survey was complete, soil cores (9 per quadrat) were taken from alternate quadrats (1,3,5,7) and processed using our standard DNA metabarcoding method (see above).

In 2019 (5th Sept) four groups from the next fieldtrip undertook the same experiment. Although the quadrat locations were not identical, they were within the same 3x8 m area studied the previous year. Vegetation survey data from both 2018 and 2019 classes were merged with the eDNA data for comparison (Table 2.3; Fig. 2.5).

A total of 55 species were detected in total, with the eDNA method missing three species, two of which (*Geranium sanguineum*, *Gymnadenia conopsea*), were present only at low abundance (1%). The student vegetation surveys over both years detected 39 species, and a further 14 species were observed to occur on or adjacent to the lawn (but not within any of the quadrats). However, the detection in eDNA of *Quercus robur* and *Rhododendron* sp. in quadrat 3 was less easy to explain since neither species had been observed in any adjacent gardens or fields and would not be expected to occur in a limestone area. Other species which were not present in the lawn but rather in adjacent habitats included *Prunus* sp., *Crataegus monogyna* (trees present ca 10 m away from the survey area) and *Hedera helix* (on adjacent wall 2 m away). As was the case this the *Populus x nigra* detected in grassland soil at Kirby Muxloe, the eDNA linked to these species likely originated from windblown leaves.

In terms of mean eDNA sequence abundance, the commonest species were *Pimpinella saxifraga* (mean 54%), *Pilosella officinarum* (7.6%), *Linum catharticum* (7.0%), *Leontodon saxatilis* (5.8%) and *Lotus corniculatus* (4.3%), whereas for the vegetation surveys, the most abundant species were *Agrostis stolonifera/canina* (mean 55%), *Festuca rubra* (43%), *Danthonia decumbens* (30%), *Trifolium repens* (29%), *Lotus corniculatus* (26%). In the case of *P. saxifraga*, the large disparity in abundance between eDNA and vegetation analysis (54% vs 16%) despite its distinctive leaves is possibly an indication of its extensive root system.

Table 2.3: eDNA and vegetation analysis on the National School Lawn at Turlough, Co. Clare. The vegetation surveys were undertaken by undergraduate students in either 2018 (suffix 18; 7 quadrats) or 2019 (suffix 19; 4 quadrats). Some species were found in both years (2) and others only in one of the survey years (1/0 or 0/1). Several species known to be present at the site but not observed within ny of the quadrats are indicated by *.

	eDNA	NVC	Both	eDNA1	eDNA3	eDNA5	eDNA7	C1-18	C2-18	C3-18	C4-18	C5-18	C6-18	C7-18	C1-19	C2-19	C3-19	C4-19
Achillea millefolium	1	2	1	0.01%	0.20%		0.02%	5%										
Agrostis capillaris/gigantea	1	*		0.15%	0.19%	0.02%	0.08%											
Agrostis stolonifera/canina	1	1/0	1	0.20%	0.52%	0.01%	0.13%		30%		80%							
Anthoxanthum odoratum	1	*		0.11%			0.04%											
Arrhenatherum elatius	1	*		0.06%	0.04%													
Bellis perennis	1	0/1	1	0.01%	0.25%	0.46%	0.01%											1%
Briza media	1	2	1	1.83%	1.85%	1.49%	2.02%	4%	12%	10%	16%	2%				3%		2%
Carex flacca	1	2	1	0.49%	2.07%	2.22%	1.01%		1%	4%	71%		14%	40%			5%	3%
Centaurea nigra	1	2	1	0.16%	2.37%	2.92%	0.08%	10%	14%	10%	19%				1%	1%	2%	1%
Centaurea scabiosa	1	2	1	0.55%		0.01%	0.00%	3%							15%			1%
Crataegus monogyna	1	*			0.04%													
Cynosurus cristatus	1	2	1	0.11%	0.09%		0.14%		2%		9%			1%	25%	6%	2%	
Dactylis glomerata	1	1/0	1	0.07%	0.09%		0.03%			25%		1%		2%				
Danthonia decumbens	1	2	1	0.21%	0.51%	0.55%	0.15%		1%							30%	50%	40%
Daucus carota	1	2	1	0.17%	0.09%	0.11%	0.75%	15%	7%	11%	8%	8%	25%		2%	6%	10%	1%
Euphrasia agg.	1	0/1	1	0.07%	0.26%	0.12%	0.17%									1%	5%	
Festuca rubra	1	2	1	0.90%	0.63%	0.94%	1.14%	70%	12%		33%	48%	51%	55%	30%	6%	80%	40%
Fraxinus excelsior	1	0/1	1			0.04%										1%	1%	
Gentianella amarella	1	1/0	1	0.04%	0.02%	6.50%	0.06%						1%					
Geranium sanguineum		2													1%		1%	1%
Gymnadenia conopsea		2											1%			1%		
Hedera helix	1	*		0.57%	0.58%													
Hieracium castellanum	1	1/0	1	0.74%	6.46%	2.05%	0.91%		1%	3%								
Holcus lanatus	1	1/0	1	0.19%	0.15%		0.02%				3%							
Hypochaeris glabra	1	0/1	1	0.23%	0.49%	3.08%	1.95%										8%	30%
Jacobaea vulgaris	1	1/0	1		0.18%				5%		5%		2%	7%				
Koeleria macrantha	1	*		0.10%		0.02%	0.01%											
Lathyrus pratensis	1	*		0.09%			0.15%											
Leontodon hispidus	1	2	1		0.06%			2%			14%					11%	4%	
Leontodon saxatilis	1	1/0	1	0.23%	1.36%	10.43%	11.30%	2%		8%				6%				
Leucanthemum vulgare	1	2	1		0.48%				1%	6%	15%						1%	
Linum catharticum	1	2	1	0.94%	24.67%	2.21%	0.14%		2%		14%	18%				4%	10%	3%
Lolium perenne/multiflorum	1	*			0.01%		0.01%											
Lotus corniculatus	1	2	1	2.60%	4.92%	3.50%	5.98%	12%	7%		16%	35%	32%	69%	20%	6%	10%	50%
Medicago lupulina	1	0/1	1	0.82%	1.42%	0.07%	0.03%								3%	5%	1%	2%
Pilosella officinarum	1	2	1	1.87%	19.33%	6.91%	2.23%		5%		14%					7%	25%	7%
Pimpinella saxifraga	1	2	1	81.53%	27.86%	50.13%	56.29%	20%	7%	5%		51%	33%	14%	2%	6%	10%	9%
Plantago lanceolata	1	1/0	1	0.26%	0.55%	0.22%	0.46%		1%	6%				35%				
Plantago maritima	1	1/0	1	0.04%			0.04%					9%						
Poa nemoralis/compressa	1	*		0.01%														
Poa pratensis	1	*					0.01%											
Poa trivialis	1	*			0.02%	0.01%	0.01%											
Polygala vulgaris	1	*				0.28%	9.18%											
Potentilla erecta	1	*			0.03%													
Prunella vulgaris	1	2	1	0.31%	0.67%	1.63%	1.36%	2%	5%	5%	1%		3%	6%	30%	1%	2%	1%
Prunus sp.	1	*		0.01%	0.05%	0.01%												
Quercus robur	1				0.04%													
Ranunculus repens/bulbosus	1	2	1	2.10%	0.02%	0.11%	0.06%		3%	7%							1%	3%
Rhinanthus minor	1	0/1	1	0.02%	0.01%		0.01%								2%	1%	9%	2%
Rhododendron sp.	1				0.01%													
Succisa pratensis		1/0							15%		30%	18%						
Taraxacum officinale agg.	1	1/0	1				0.09%						1%	10%				
Trifolium pratense	1	2	1	0.58%	0.54%	2.19%	3.29%			10%	1%		7%	47%	12%	6%	20%	5%
Trifolium repens	1	2	1	1.55%	0.48%	0.07%	0.26%	20%	7%	70%	40%	42%	30%	57%	25%	10%	4%	9%
Vicia cracca	1	2	1			1.30%	0.29%					10%		9%				1%
	52	**39**	**36**															

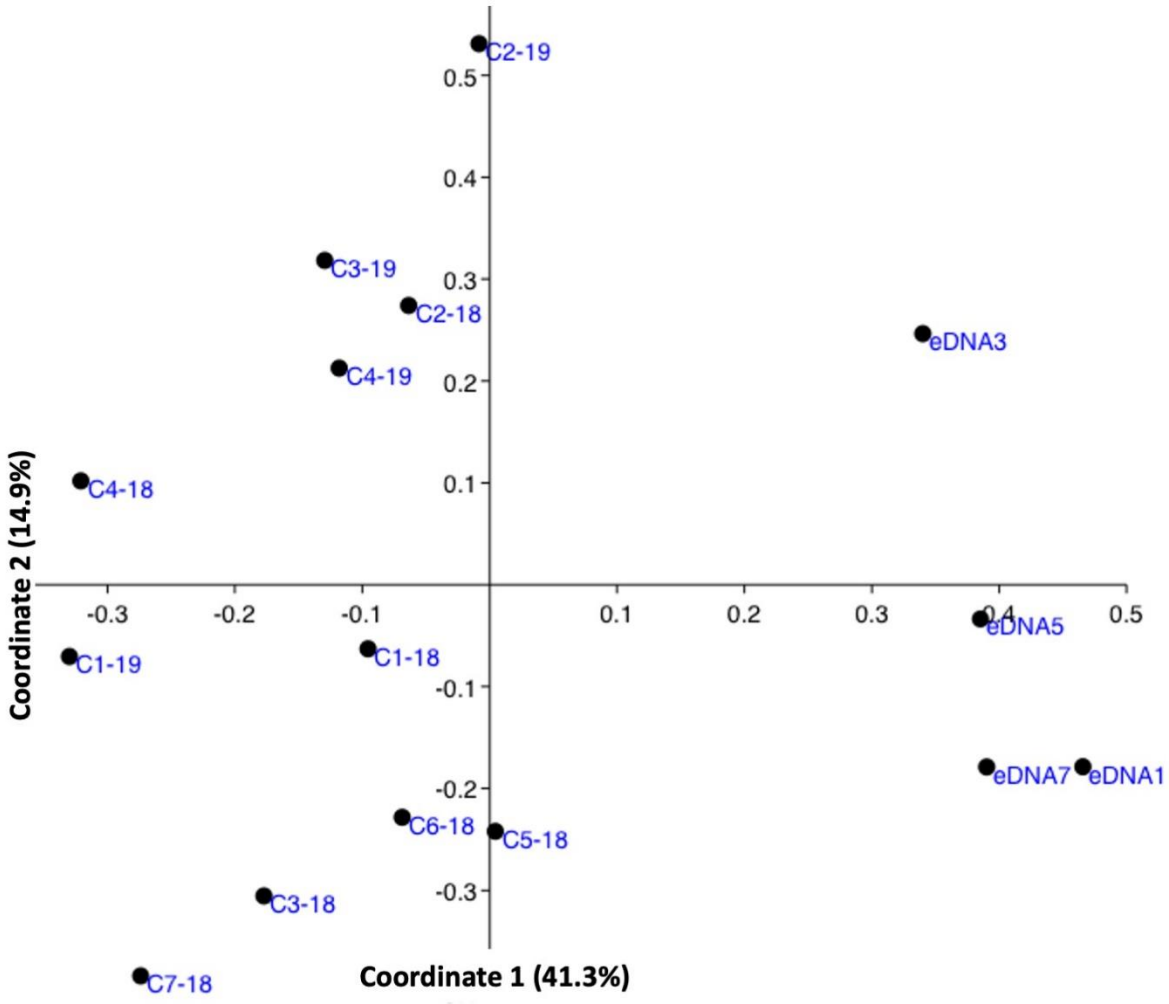


Fig. 2.5. PCO ordination of Turlough quadrats. Only the 33 common to both the eDNA and NVC databasets

3) Quantifying the rate of decay of plant DNA in disturbed soils (Mesocosm [pot] Expt)

3A) Overview

Having established that it is possible to identify eDNA derived from plants (and fungi) from soils, it would be informative to know whether such eDNA is derived from living /recently dead plant biomass, and how long eDNA persists in soil once the living biomass is removed. To explore this, a mesocosm experiment under temperate glasshouse conditions was devised to test the effect of contrasting agricultural practises on the decay of eDNA from a diverse grassland sward.

3B) Method

Topsoil for the experiment was collected from a ca. 30x30 m quadrat at Brignant (Fig. 3.1) on 28/29th June 2018. The area was previously classified as "Fairly homogenous south-facing stand of neutral grassland indicative of past haymeadow management, having previously been managed under the terms of the ESA scheme as a "Tier 2 haymeadow", occasional sheep grazing and a list of plant species was available for this area (Table 3.1; Dave Rogerson, WG, pers. comm.). Turves (30x30cm x ca. 10 cm deep) were cut from across the field and replaced with turf from the edge of the field (total fresh wt ca. 120 kg).

During the soil collection 36 soil cores from across the ca. 30x30 m quadrat were collected and pooled according to our standard protocol.

The turves were broken up using a portable electric soil shredder and further mixed in a cement mixer (Fig 3.2). At the time of collection the sward height on the turves was ca. 6 cm; during turf disruption, vegetation was not removed and any larger clumps were torn manually into smaller pieces before placing in the soil shredder to mix these with soil. We aimed to strike a balance between breaking down soil clods to a small enough size (<3 cm) to avoid excessive pot-to-pot variation but sufficient to realistically mimic the effects of ploughing/rotovation/harrowing.

Once homogenised, soil was dispensed into 24 plastic plant pots (9.5 l capacity), with 3.5 kg soil in each. The initial volume of the soil in the pots was ca. 6 l (bulk density 0.58 g/cm^3). An additional set of four replicate pots was established using three layers of intact but inverted turves (mimicking ploughing without harrowing). The mean weight of the pots containing the inverted turves was higher (4.8 kg/pot) than for the homogenised soils and the initial bulk density also higher (0.69 g/cm^3). The soil was very dry on collection, following a long period without rain. Soil pH was 5.22 with the electric conductivity 29.2 µs/cm. Initial water content was 11% (w/w).

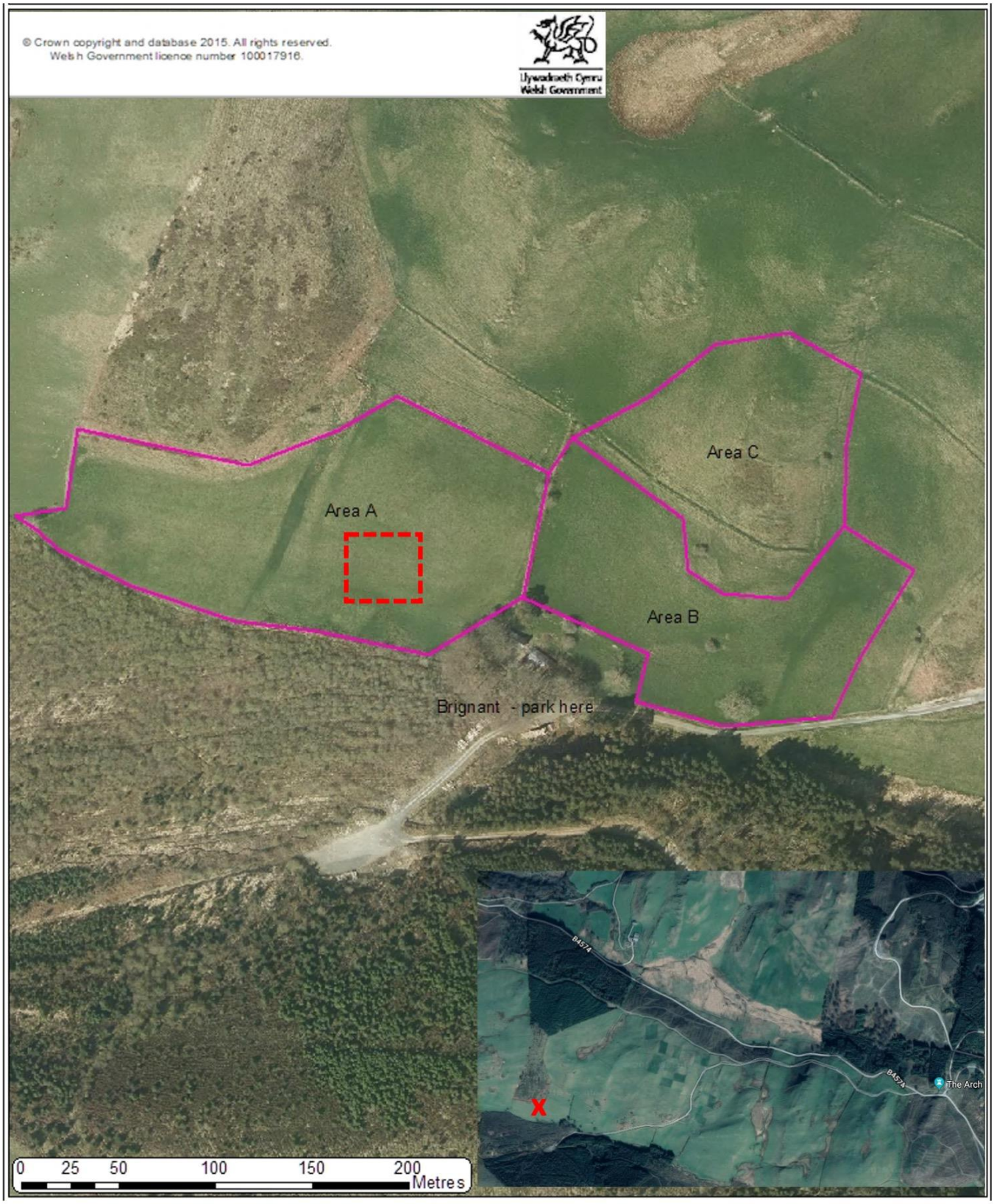


Fig 3.1. Location of turves used for the pot experiment from Area A (52.3636,-3.8381)at Brignant, Pwllpeiran. Inset map shows location of the fields in relation to the Brignant Long Term Experiment (marked x) and Hafod Arch. Turves were collected from a ca. 30x30 m area, as outlined in red font. on 28th June 2018. Collection and preparation of soil from Area A for the pot experiment

Table 3.1: List of plants found in AreaA at the Brignant field site

Habitat descriptions / species lists – Brignant.

Area A: Fairly homogenous south-facing stand of neutral grassland indicative of past haymeadow management. Was in the ESA scheme as Tier 2 haymeadow.

Area B: Semi-improved neutral to acid grassland with high sown species content, between 25-35%. Still species-rich although more grassy and more weed species.

Area C: Bent/fescue acid grassland – partly semi-improved but also species-rich. A good example of this. Includes mountain pansy and devil's bit.

Low and Neutral Grassland, upland haymeadows and species-rich acid grassland are habitats of principal importance as listed in Section 42 of the Natural Environment & Rural Communities (NERC) Act 2006

			A	B	C
	Species found in the grasslands		**Species-rich neutral grassland**	**Semi-improved grassland**	**Acid grassland**
Graminoids	*Agrostis canina*	Velvet bent			X
Graminoids	*Agrostis capillaris*	Common bent	X	X	
Graminoids	*Agrostis vinealis*	Browntop bent			X
Graminoids	*Anthoxanthum odoratum*	Sweet vernal grass	X	X	X
Graminoids	*Bromus hordaceaeus*	Soft brome	X		
Graminoids	*Cynosurus cristatus*	Crested dogstail	X	X	X
Graminoids	*Dactylis glomerata*	Cocksfoot	X	X	
Graminoids	*Danthonia decumbens*	Heath grass			X
Graminoids	*Deschampsia flexuosa*	Wavy-hair grass	X		X
Graminoids	*Festuca ovina*	Sheep's fescue	X		X
Graminoids	*Festuca rubra*	Red fescue	X	X	
Graminoids	*Holcus lanatus*	Yorkshire fog	X	X	X
Graminoids	*Holcus mollis*	Creeping softgrass	X	X	
Graminoids	*Lolium perenne*	Perennial ryegrass	X	X	X
Graminoids	*Molinia caerulea*	Purple moor-grass			X
Graminoids	*Poa annua*	Annual meadowgrass	X	X	
Graminoids	*Poa pratensis*	Smooth-stalked meadowgrass	X	X	
Graminoids	*Carex pilulifera*	Pill sedge			X
Graminoids	*Luzula campestris*	Field woodrush	X	X	
Graminoids	*Juncus effusus*	Soft rush	X	X	
Forbs	*Achillea millefolia*	Yarrow	X	X	
Forbs	*Bellis perennis*	Daisy	X	X	
Forbs	*Campanula rotundifolia*	Harebell			
Forbs	*Cardamine pratensis*	Lady's smock	X	X	
Forbs	*Centaurea nigra*	Black knapweed	Missing		
Forbs	*Cerastium fontanum*	Mouse-ear chickweed	X	X	X
Forbs	*Cirsium vulgare*	Spear thistle		X	
Forbs	*Cirsium palustre*	Marsh thistle	X		
Forbs	*Digitalis purpurea*	foxglove		X	
Forbs	*Galium saxatile*	Heath bedstraw		X	X
Forbs	*Heracleum sphondylium*	Hogweed	X		
Forbs	*Hyascinthoides non-scripta*	Bluebell	X		
Forbs	*Hypochaeris radicans*	Cat's ear	X	X	
Forbs	*Lathyrus pratensis*	Meadow vetchling	X		
Forbs	*Leontodon autumnalis*	Autumn hawkbit	X		
Forbs	*Leucanthemum vulgare*	Oxeye daisy	X		
Forbs	*Lotus corniculatus*	Bird's foot trefoil	X	X	X
Forbs	*Lotus uliginosus*	Greater trefoil	X		
Forbs	*Pilosella officinarum*	Mouse-ear hawkbit			X
Forbs	*Plantago lanceolata*	Ribwort plantain	X	X	X
Forbs	*Potentilla anserina*	Silverweed	X	X	
Forbs	*Potentilla erecta*	Tormentil	X	X	X
Forbs	*Prunella vulgaris*	Selfheal	X	X	X
Forbs	*Ranunculus acris*	Meadow buttercup	X	X	X
Forbs	*Ranunculus ficaria*	Celandine	X		
Forbs	*Ranunculus repens*	Creeping buttercup	X	X	
Forbs	*Rhinanthus minor*	Yellow rattle	X		
Forbs	*Rumex acetosa*	Common sorrel	X	X	X
Forbs	*Rumex acetosella*	Sheep's sorrel			X
Forbs	*Rumex obtusifolius*	Broad-leaved dock	X	X	
Forbs	*Sanguisorba minor*	Lesser burnet			X
Forbs	*Succisa pratensis*	Devil's bit scabious			X
Forbs	*Taraxacum spp.*	Dandelion	X	X	X
Forbs	*Trifolium dubium*	Lesser trefoil	X		
Forbs	*Trifolium repens*	White clover	X	X	X
Forbs	*Trifolium pratense*	Red clover	X		
Forbs	*Urtica dioicia*	Nettles		X	
Forbs	*Vaccinium myrtilus*	Bilberry			X
Forbs	*Viola lutea*	Mountain pansy			X
Forbs	*Viola riviana*	Dogviolet		X	X
Forbs	*Veronica chamaedrys*	Germander speedwell	X	X	
Forbs	*Veronica serpyllifolia*	Thyme-leaved speedwell	X	X	X
Bryophytes	*Pteridium aquilinum*	Bracken	X	X	X
Bryophytes	*Hylocomium splendens*	Glittering wood-moss			X
Bryophytes	*Hypnum cupressiforme*	Hypnum moss			X
Bryophytes	*Pleurozium schreberi*	Feather moss	X	X	X
Bryophytes	*Polytrichum spp.*	Chimmney-sweeper moss			X
Bryophytes	*Rhytidiadelphus squarrosus*	Springy turf moss	X	X	X
		SPECIES TOTALS	47	38	35

Fig 3.2. Collection and preparation of soil from Area A for the pot experiment

The pots were then placed in an unheated and partially shaded glasshouse and temperature was monitored over the course of the experiment with a datalogger (Fig. 3.3). Pots were left unwatered for 7 d and then watered daily for four days prior to imposition of treatment regimes (Table 3.2) on 10th July. The aim of these was to mimic a range of post-ploughing management and climatic scenarios.

Seven treatments were imposed (4 replicates each) on 2nd July 2018.

Table 3.2: Treatment summary

Code	Lolium Sown	Watering	Amendment
WF	NO	WET	FERTILISER
WM	NO	WET	LIME
WI	NO	WET	INVERTED
Wx	NO	WET	NONE
Dx	NO	DRY	NONE
WL	YES	WET	NONE
DL	YES	DRY	NONE

Soil amendments: For the lime treatment, 7.9 g/kg dry wt. soil of calcium carbonate and for the Fertiliser treatment 220 mg/kg (YaraMila Sulfur-cut [NPK:22-4-14 + 7.5% SO_3]) fresh wt. soil was added. These amendments were added to the soil surface and not mixed into the pots (mimicking agricultural practice).

Seeding of pots: For the two pots sown with grass (125 mg seed/pot) of *Lolium perenne* seeds (variety AberWolf) were sprinkled on the pot surface and covered with a thin layer of soil. Additional seed were sown in October 19 because the original seeds had shown a poor germination rate in some pots. Unseeded pots were weeded at fortnightly intervals, aiming to also remove the roots. For *Lolium*-sown pots non-grass seedlings were similarly removed, as were (where possible) any grasses recognised and not being *Lolium*.

Watering regimes: WET treatments were subject to daily watering and initially the DRY treatments watered only twice a week. In the autumn (from 14th November; 5 months after start), watering of the DRY treatments was reduced to fortnightly, to enhance the contrast between the WET and DRY treatments, as the waterholding capacity of the soils looked to be higher as previously expected. Pots were placed on the greenhouse in a random pattern and re-randomised periodically.

The sampling of soils from the pots was undertaken using an apple corer (11 mm diam) to 6 cm depth with 4 cores taken across each pot (ca. 10 g dry wt. per pot [=17-21 fresh wt.]). The samples were weighed, frozen at -80ºC and freeze-dried. Dry weight was also recorded allowing to calculation of water loss. Each set of cores was ground through a <0.5mm sieve), mixed thoroughly and 200 mg taken for DNA extraction using the Qiagen DNase® PowerSoil® Kit. The DNA extract was kept at a -20ºC freezer prior to analyses.

DNA metabarcoding was undertaken to amplify ITS2 region with ChenF primer for plant coverage and NGS Tedersoo (2014) primer mix for fungi together with a barcoded ITS4, the ION Torrent (Ion PGM System for Next-Generation Sequencing – Thermo Fisher Scientific) was used for sequencing the data, as above.

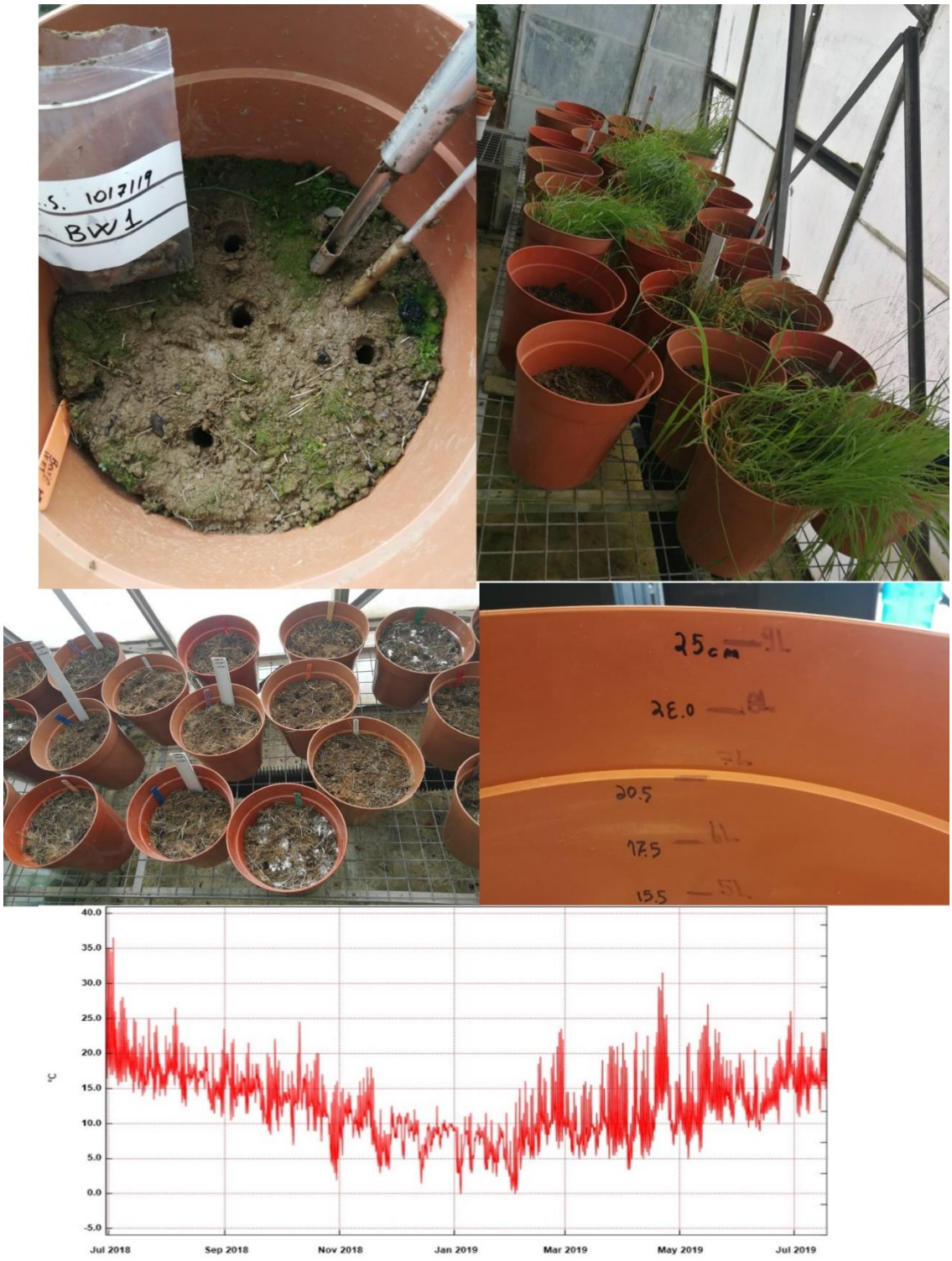


Fig 3.3. Images of the pots *in situ* and temperature profile during the course of the experiment.

The soils in the pots were sampled at four timepoints in order to assess the decay of eDNA under the various conditions:- Baseline (T0; at imposition of treatments; 11/7/18) representing the total presence of initial eDNA; 2nd sampling, one month after T0 (10/8/18); 3rd sampling, three months after T0 (10/10/18); 4th sampling, 5 months after T0 (not sequenced, freeze-dried kept at -80°C); and 5th sampling, after 12 months (11/7/19). Soil samples were also taken in December (11/12/18) but not analysed further.

3C) Results

Pot Experiment Data

Some settlement of the soils in the pots was observed during the course of the experiment, partly as a result of the degradation of the large amount of plant litter that was present in the soils originally. This was more noticeable in the pots composed of the homogenised soil. Soil bulk density was estimated by measuring the level of the soil below the lip of the pot but did not differ significantly between treatments by the end of the experiment (Fig. 3.4F).

The effect of different soil moisture levels on degradation of plant litter was tested by imposing two different watering regimes, either in the presence or absence of plants. The five treatments subject to normal watering (WET; i.e. normal for glasshouse kept plants) were watered daily, whereas the two treatments subject to drier watering regime (DRY) were watered only on alternate days. However, monitoring of pot weights and soil moisture levels suggested that these contrasting watering regimes did not lead to any significant difference in soil moisture level (Fig 3.4A,B). Therefore, the watering regime for the DRY treatments was progressively reduced during August-November 2018. Fortnightly watering was found to provide an optimal contrast with the 'WET' watering regime.

Soil pH (initially 5.22) dropped slightly in the initial months of the experiment (mean 4.9 in Nov18) but returned to 5.2 by the end of the experiment. The pH of the lime-treated soils was 1.5 (6.4-6.8) higher than the other treatments (Fig 3.4C,D). Soil conductivity was within the expected range (250-750 µS/cm).

Attempts were made to quantify levels of microbial respiration (soil CO_2 production; measured using an InfraRed Gas Analyser) from the mesocosms, as a means of assessing the rate of decay of the buried vegetation which might show some correlation with the rate of disappearance of plant DNA. However, there was large intra-replicate variation and no clear treatment effect (data not shown), so respirometry measurements were not made beyond the second harvest.

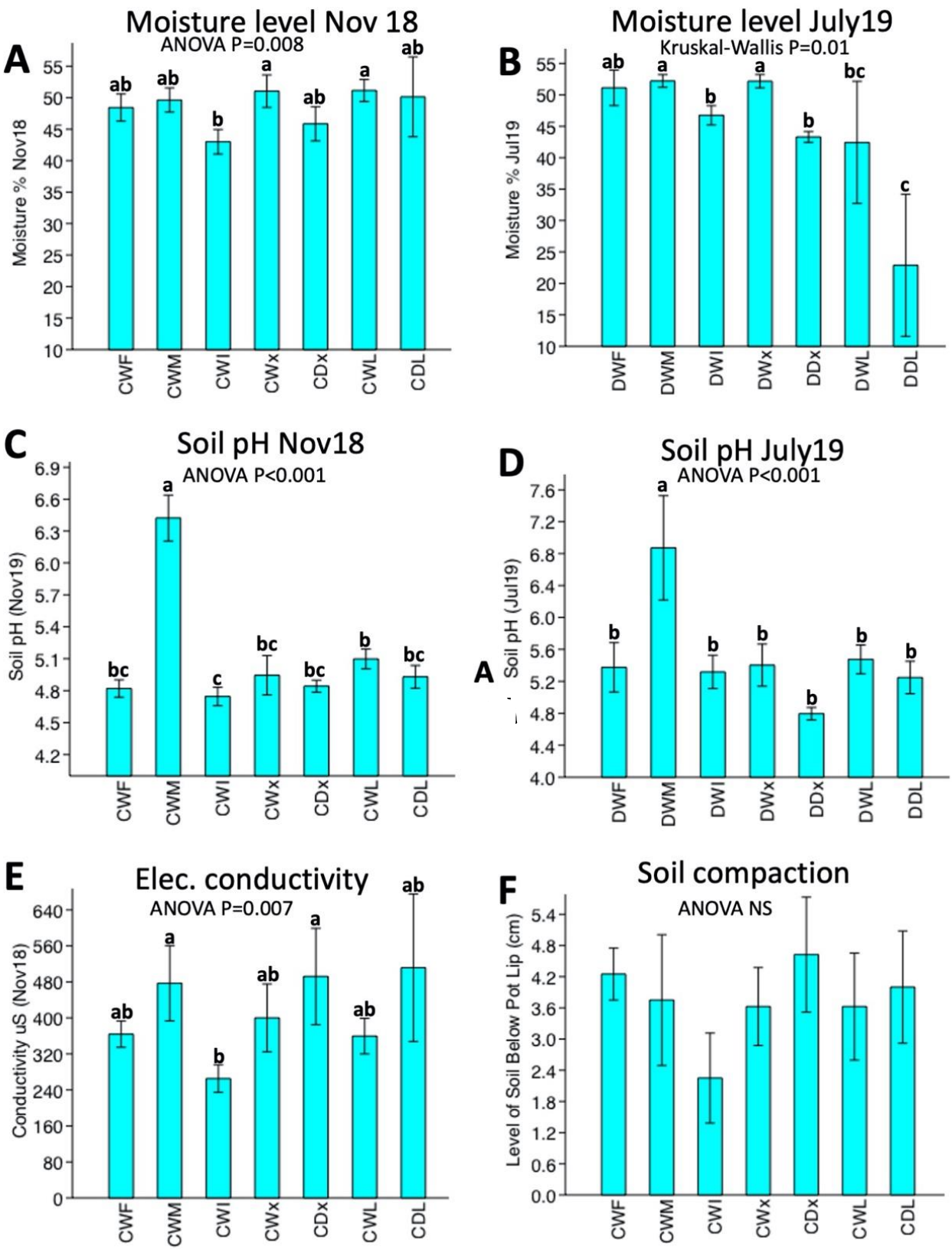


Fig. 3.4: Moisture level (A,B), pH (C,D), Conductivity (E) and soil compaction (F) of pot soils during the 12 month experiment. Error bars indicate standard deviation. Treatments with different superscript lettering are significantly different (Tukey's test). Treatment codes are as follows. 1st letter: C=3rd harvest (Nov18) or D=4th harvest (July19); 2nd letter: W=watered or D=dry; 3rd letter: F=fertilized, M=lime, x=no addition or L=*Lolium* sown.

DNA metabarcoding: Plants

DNA metabarcoding undertaken on the 112 core samples taken from the 28 pots (7 treatments x 4 replicates) and the four sampling timepoints (July 18, August 18, Nov 18, July 19), and additionally the core sample taken from AreaA at the time of turf collection.

Two samples failed at the sequencing stage but the remaining 111 samples worked well. Of the 5.2 million sequences obtained in total, 4.6 million passed the quality control steps (mean 41,524 per sample [range7020-173,968]). Sequences were taxonomically assigned using RDP using both our in-house plant sequence database and the UNITE (v8) fungal sequence databases.

For the PCR amplification of the ITS2 DNA barcode a mixture of primers was used, designed to amplify both plant and fungal DNA barcodes. Overall 33.6% of sequences (range 0.7%-87.9%) were assigned to Kingdom Viridiplantae (Plants), the remainder being fungal and other biota (animals [ca. 0.03%-1.3%], diatoms, protozoa, unassigned etc.). The assignment of sequences to different types of plant (monocots, dicots, mosses/liverworts, algae), fungi and other biota is presented as a stacked barchart (Fig. 3.5).

Over the course of the experiment the proportion of sequences attributed to higher plants (monocots/dicots) decreased from ca. 80% in the original soils to <1% after 12 months, except for the pots seeded with *Lolium* (WL/DL treatments) where 10-50% of sequences were plant (mostly monocot) as expected (Fig. 3.6). In other treatments there was a decline to ca. 50% in the baseline samples (taken 10 d after the turves had been collected/homogenised, 4 d after watering had begun).

Since the initial samples were taken at the same time as application of the treatment regimes, no treatment effect would be expected. Whilst differences in the mean abundance of the various taxa are apparent, there is no significant treatment effect due to the large inter-replicate variations that were observed (Fig. 3.5, 3.6). This inter-replicate variation is likely due to the heterogeneity of the soil, even after soil shredding/mixing etc. and also related to the small size of the samples taken (4 cores, each ca. 5g fresh wt.). It had been expected that the inverted turves would differ at this stage from treatments where soil had been subject to homogenisation but this was not the case (except for mean % dicots being higher). When compared with the single AreaA sample (noting that this was from 36 pooled cores taken in the field), abundance of fungi relative to higher plants was lower in AreaA, suggesting some proliferation of fungi during the time elapsed between soil homogenisation and imposition of treatments. This would be expected since the turf homogenisation created a large amount of dead plant biomass which could be quickly colonised by saprotrophic fungi.

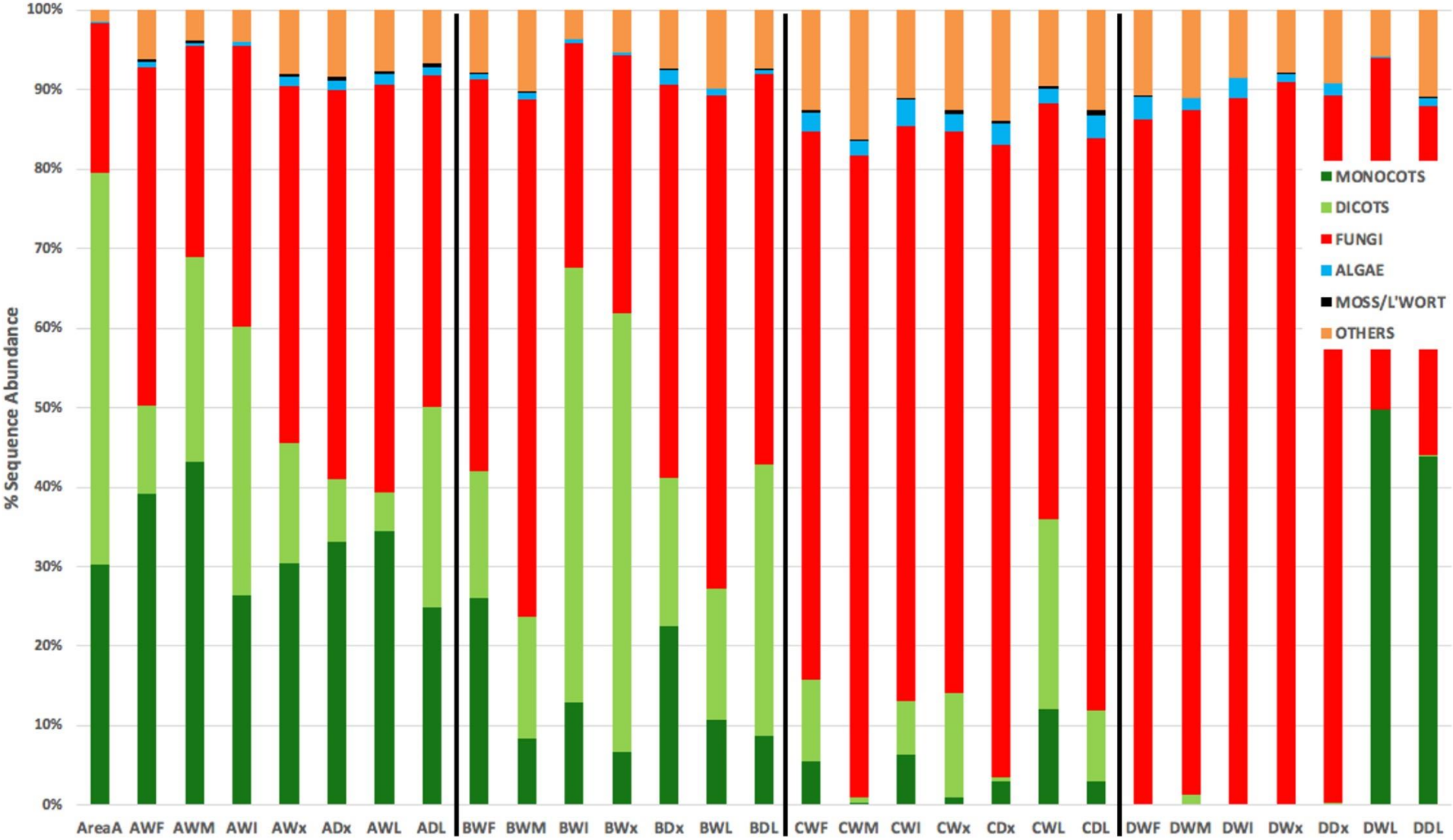


Fig. 3.5. Stacked barchart showing the relative abundance of the main groups of plants and fungi (Including oomycetes) in the seven treatments across the four sampling timepoints.

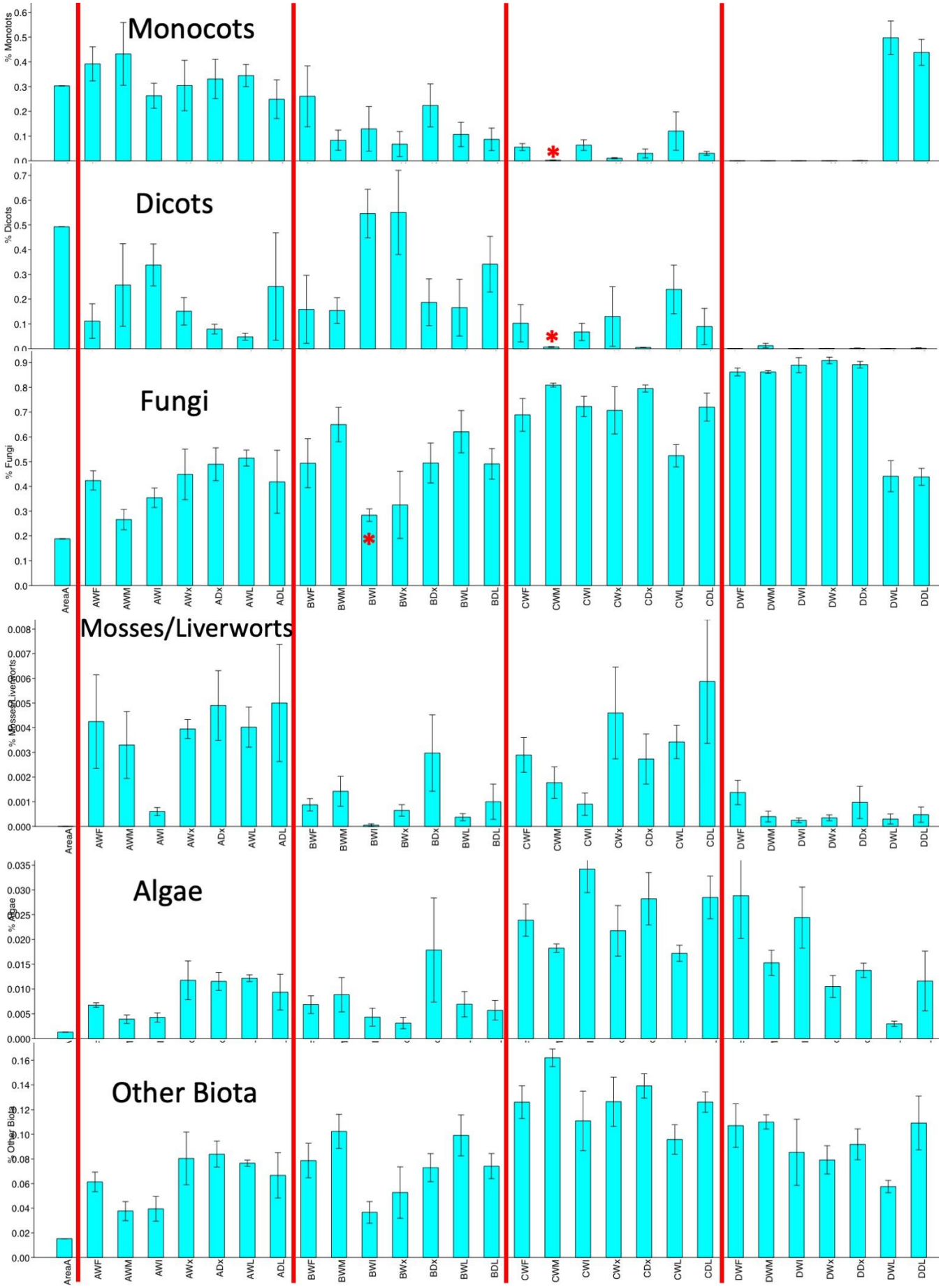


Fig. 3.6: % sequence abundance of different taxa in the soils across different treatments and timepoints. Error bars indicate SE (n=4). Significant treatment effects are indicated by *.

At the second harvest, the only significant treatment effect was the reduced relative abundance of Fungi in the inverted turves (BWI; K-W P=0.027; Fig. 3.6). At third harvest monocots and dicots were at lower abundance in the lime treatment pots (BWM; K-W P=0.023). By the final harvest (July19) abundance of higher plant sequences was very low, except for monocots in the *Lolium* sown pots. Examination of the species present in these pots revealed that whilst *Lolium* was the most abundant species in these treatments (DWL, DDL), other grasses were also present (mostly *Holcus lanatus* but also *Poa trivialis*, *Anthoxanthum odoratum* and *Agrostis capillaris*). Presumably because they had regrown from seed or culms from the original turves, and were possibly more abundant as a result of the poor germination of the initially-sown *Lolium*.

A total of 47 plant species are recorded from the AreaA fieldsite (Table 3.1). Of these, 23 were detected in eDNA from AreaA. Some additional species were detected only in eDNA (three *Poa* spp. [likely difficult to identify under field conditions], *Crepis capillaris* and several mosses)(Table 3.3). However, for the mosses, the status of DNA barcoding lags behind that for higher plants, so these names should be treated with some caution. Three unexpected species (*Castanea sativa*, *Fagus sylvaticus* and *Heliconia bihai*) were detected at low levels in several samples (but not AreaA and the inverted turves). These three species are present near the area where the soil was homogenised (two large trees and *H. bihai* abundant in the adjacent [tropical] glasshouse), so likely to be contamination during the soil homogenisation process.

Principal Coordinates Ordination of the plant communities (including only the higher plant species listed in Table 3.3) did not reveal any clear differentiation of any treatments or timepoints (Fig. 3.7). Some cluster of points for replicate samples was apparent, for example the *Lolium*-seeded pots DWL/DDL in the bottom right quadrant of the ordination, but such patterns would be expected given that the WL/DL treatments contain a high abundance of grasses. Removal of the grasses prior to ordination did not reveal any clearer patterns of ordination (Fig. 3.7B). Statistical analysis of these dataset using PERMANOVA did not reveal any significant treatment differences

Examination of the abundance of four of the 15 non-grass species (Fig. 3.8) confirmed earlier observations that there was a large amount of variation within the treatments and no consistent treatment effects were present across the series of soil samples but that there was a general decline in abundance of all the herb species over time.

Since there was little evidence of any treatment effect (except the detection of the sown grasses in the DL/WL treatments, the relative abundance of the 26 higher plant species detected across all five of the bare soil (non-seeded) treatments was examined (Fig. 3.9). For each species there is a reduction in % sequence abundance in successive samples. In terms of the number of species detected (i.e. presence or absence) there is a progressive reduction in the number of the 26 spp. detected with time. After 12mths, 16 spp. are still detected, albeit at low abundance.

The number of pot from the five bare soil treatments (20 pots in total) in which each of the species was detected was also examined in detail (Fig. 3.10; Fig. 3.11). Here also a progressive reduction is seen and by the final 12mth sampling point, most of the species still detected are found in only 1-3 of the 20 pots.

Table 3.3: Comparison of species detected in AreaA by vegetation surveying compared to those species detected via eDNA

Group	Species	AreaA	eDNA	Notes
Graminoids	*Agrostis canina*		1	*(Agrostis stolonifera/canina)*
Graminoids	*Agrostis capillaris*	1	1	*(Agrostis capillaris/gigantea)*
Graminoids	*Anthoxanthum odoratum*	1	1	
Graminoids	*Bromus hordaceaeus*	1		
Graminoids	*Cynosurus cristatus*	1	1	
Graminoids	*Dactylis glomerata*	1	1	
Graminoids	*Deschampsia flexuosa*	1		
Graminoids	*Festuca ovina*	1	1	
Graminoids	*Festuca rubra*	1	1	
Graminoids	*Holcus lanatus*	1	1	*(Holcus lanatus/mollis)*
Graminoids	*Holcus mollis*	1		******see above
Graminoids	*Juncus effusus*	1		
Graminoids	*Lolium perenne*	1	1	*(Lolium perrenne/multiflorum)*
Graminoids	*Luzula campestris*	1	1	
Graminoids	*Poa annua*	1		
Graminoids	*Poa nemoralis/compressa*		1	
Graminoids	*Poa pratensis*	1		
Graminoids	*Poa trivialis*		1	
Forbs	*Achillea millefolia*	1		
Forbs	*Bellis perennis*	1		
Forbs	*Cardamine pratensis*	1	1	
Forbs	*Cerastium fontanum*	1	1	
Forbs	*Cirsium palustre*	1		
Forbs	*Crepis_capillaris*		1	
Forbs	*Heracleum sphondylium*	1		
Forbs	*Hyacinthoides non-scripta*	1		
Forbs	*Hypochaeris radicans*	1	1	
Forbs	*Lathyrus pratensis*	1		
Forbs	*Leontodon autumnalis*	1		
Forbs	*Leucanthemum vulgare*	1		
Forbs	*Lotus corniculatus*	1		
Forbs	*Lotus uliginosus*	1		
Forbs	*Plantago lanceolata*	1	1	
Forbs	*Potentilla anserina*	1		
Forbs	*Potentilla erecta*	1		
Forbs	*Prunella vulgaris*	1	1	
Forbs	*Ranunculus acris*	1	1	*(Ranunculus acris/occidentalis)*
Forbs	*Ranunculus ficaria*	1		
Forbs	*Ranunculus repens*	1	1	*(Ranunculus bulbosus/repens)*
Forbs	*Rhinanthus minor*	1		
Forbs	*Rumex acetosa*	1	1	
Forbs	*Rumex obtusifolius*	1	1	
Forbs	*Taraxacum spp.*	1	1	
Forbs	*Trifolium dubium*	1		
Forbs	*Trifolium pratense*	1	1	
Forbs	*Trifolium repens*	1	1	*(Trifolium_repens/nigrum)*
Forbs	*Veronica chamaedrys*	1		
Forbs	*Veronica serpyllifolia*	1	1	
Ferns	*Pteridium aquilinum*	1		
Bryophytes	*Pleurozium schreberi*	1		
Bryophytes	*Rhytidiadelphus squarrosus*	1	1	
Bryophytes	*Didymodon sinuosus*		1	Pottiaceae
Bryophytes	*Kindbergia praelonga*		1	Brachytheciaceae
Bryophytes	OTU 1320		1	Pottiales
Bryophytes	*Tortula mucronifolia*		1	Pottiaceae
Bryophytes	*Brachythecium rivulare*		1	Brachytheciaceae
	SPECIES TOTALS	47	32	

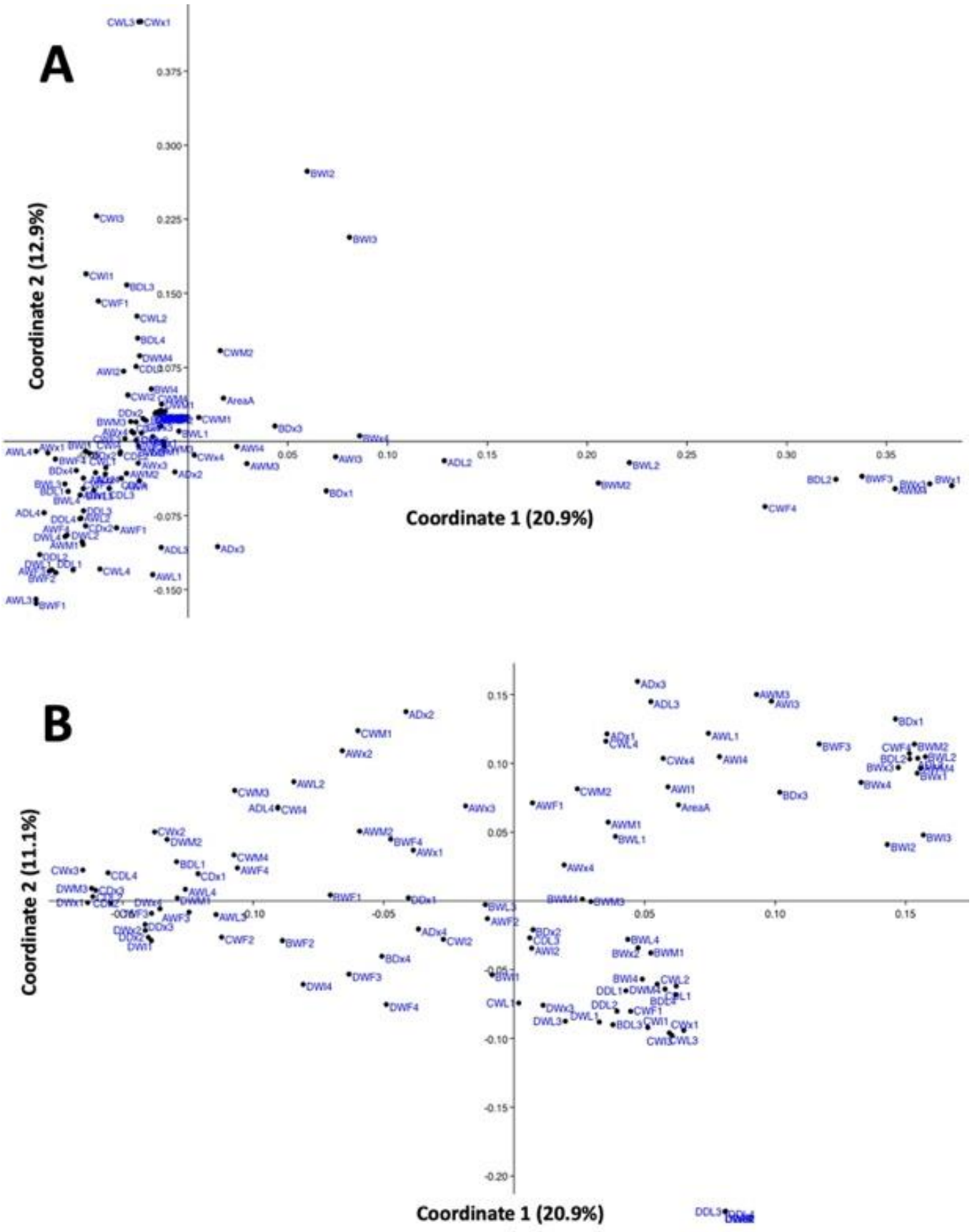


Fig. 3.7: A) Principal coordinates ordination of (A) higher plant populations (grasses + herbs; 26 spp.) and (B) herb populations alone (incl. *Luzula*; 15 spp.)

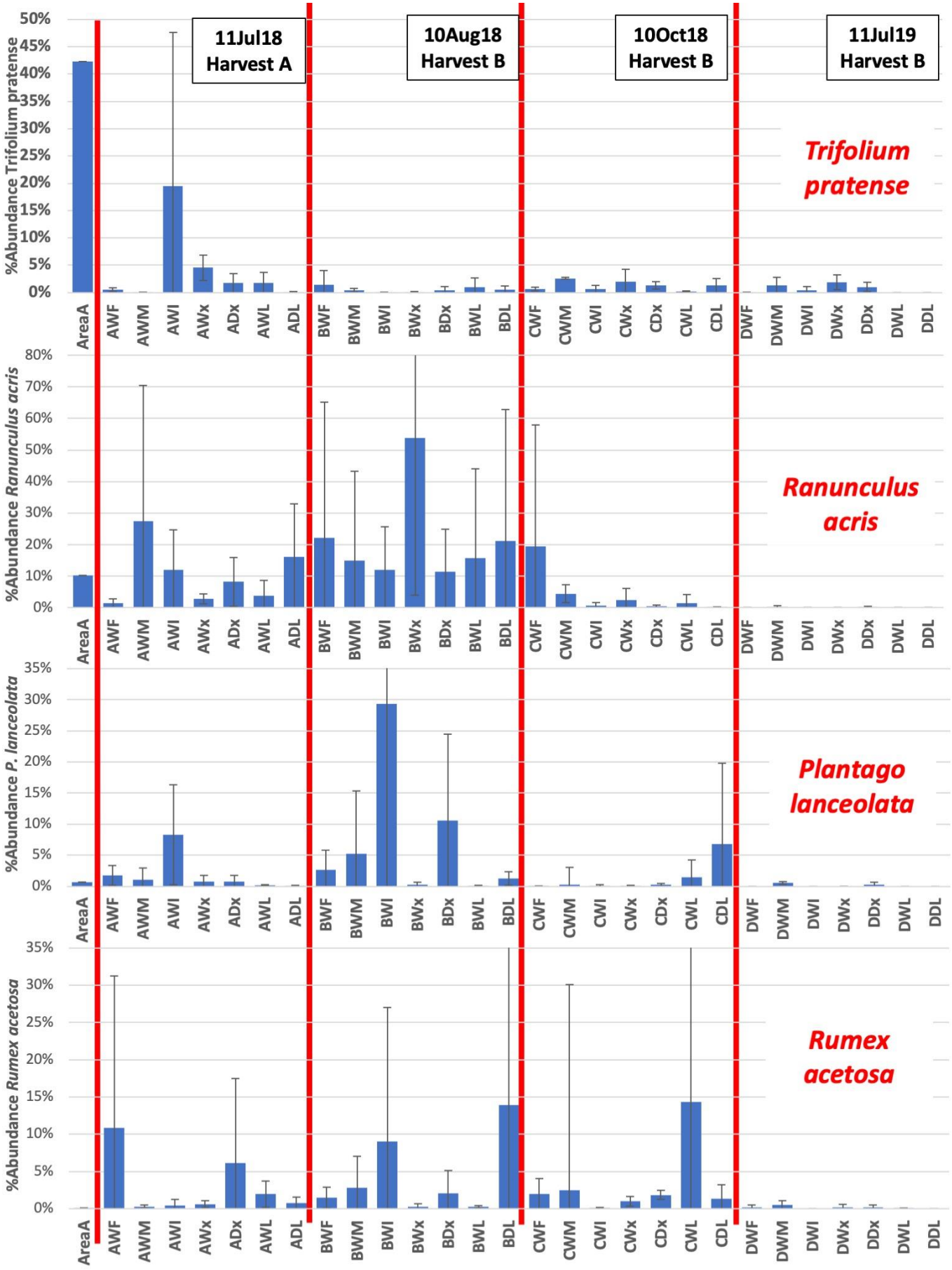


Fig. 3.8: % sequence abundance of four common herb species in the soils. Error bars indicate SD (n=4).

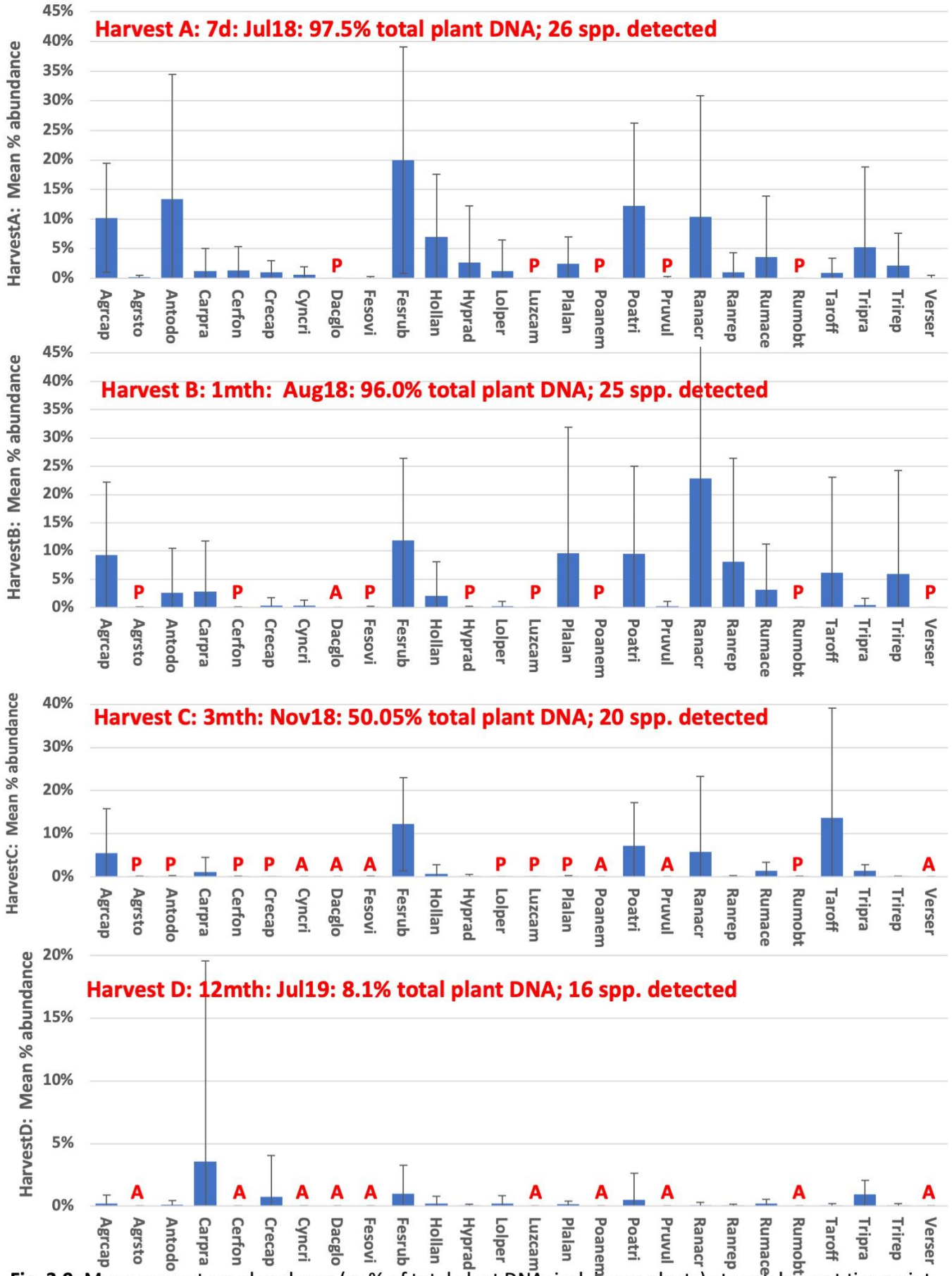


Fig. 3.9: Mean percentage abundance (as % of total plant DNA; incl. lower plants) at each harvest timepoint of each of the 26 higher plant detected as eDNA within the pot experiment soils averaged across **five treatments (excluding those pots sown with *Lolium*)**. Where bars are small 'A' indicates the species was absent (no DNA detected) and 'P' indicates present. Bars indicate standard deviation (n=20)

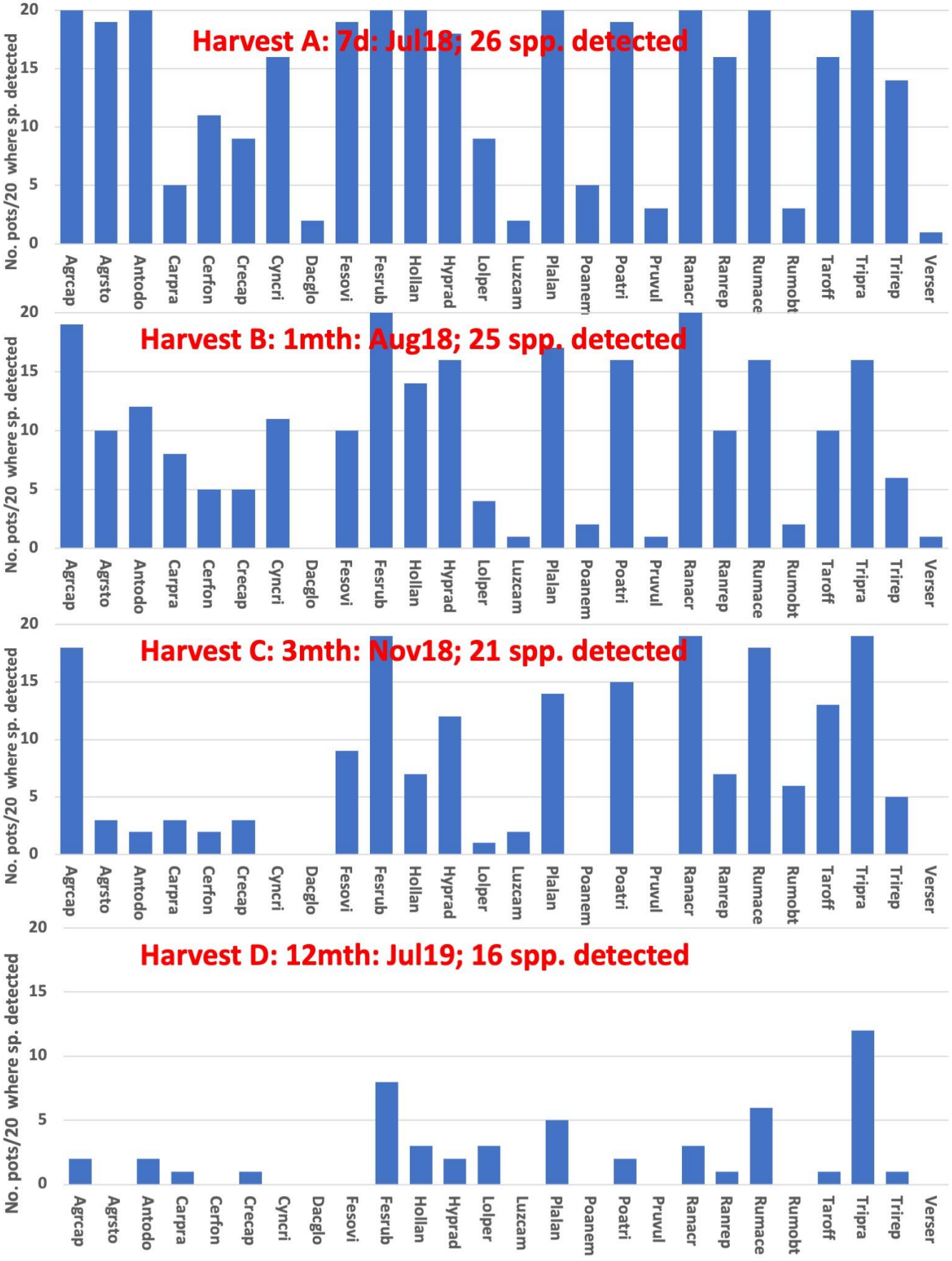


Fig. 3.10: Number of pots (/20) in which each 26 higher plant species was detected at each harvest timepoint. Includes only the five bare soil **treatments (pots sown with *Lolium* were excluded)**.

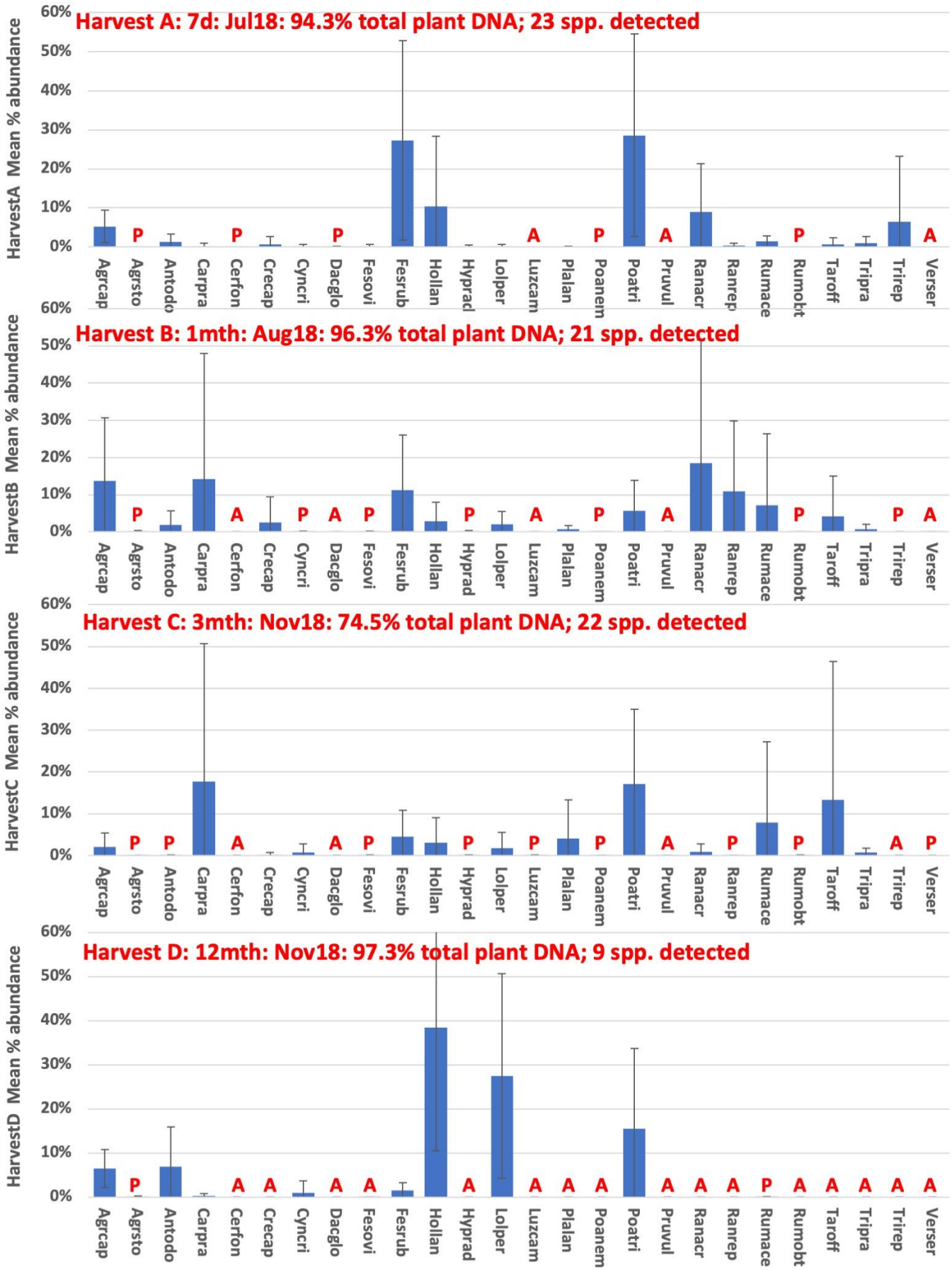


Fig. 3.11: Mean percentage abundance (as % of total plant DNA; incl. lower plants) at each harvest timepoint of each of the 26 higher plant detected as eDNA within the pot experiment soils averaged across **ONLY the two Lolium-seeded treatments**. Where bars are small 'A' indicates the species was absent (no DNA detected) and 'P' indicates present. Bars indicate standard deviation (n=20)

DNA metabarcoding: Fungi

The main focus of this study was on the fate of plant eDNA but we also investigated changes in fungal populations during the course of the pot experiment. PCO ordination of the changes in the fungal communities in the pot soils showed a slightly clearer pattern of change over time (Fig. 3.12) compared to the plant eDNA (Fig. 3.7).

At the level of individual species the four most abundant species detected (across all samples) showed interesting patterns. For two basidiomycete species, levels of eDNA were significantly higher in the inverted (intact) turves (Fig. 3.13) compared to the other treatments, consistent with the fact that these fungi form large (several metre scale), long-lived mycelial networks which would be damaged by soil disruption. For example, *Camarophyllopsis schulzeri*, is a member of the CHEGD group of fungi characteristic of undisturbed, nutrient-poor grassland (waxcaps also belong to this group) which was predominantly observed in the inverted turf treatment but exhibited a pattern of decline over the course of the experiment. CHEGD fungi are suspected to be mycorrhizal with grasses and herbs (Halbwachs et al., 2018), so the absence of any plants with which to associated would be expected to lead to a progressive decline. A very similar pattern was observed for most of the other CHEGD fungi that were detected in these soils (data not shown). *Lepista* sp. (blewit), the most abundant fungus present in the soils across all treatments, is another basidiomycete which forms fairy rings but unlike CHEGD fungi is a terricolous saptrotroph (degrade plant litter in soil). This species also persisted at higher levels in the inverted turves but showed recovery in the disrupted soil treatments by the last harvest.

The distribution of the 2nd and 4th most abundant of the fungi detected in the pots soil (but both barely detected in the original AreaA sample), *Trichoderma hamatum* and *Gibberella fujikuroi* respectively (both ascomycete fungi) was very different with only low abundance in the inverted turves and much higher abundance in the disrupted soils. Both species are fast-growing necrotrophic pathogens of fungi and plants respectively. Presumably *T. hamatum* was attacking the distrupted hyphae of the *Lepista* sp. in the initial harvest (A), declining over time as the abundance of suitable host fungi decreased, whilst *G. fuikuroi* proliferated through attack of weakened/dying plant tissues.

Taken together the distribution of these two pairs of fungi (with similar patterns being exhibited by related species) provided useful insight into the processes which occurred in the different pot treatments, especially whether in relation to soil disruption.

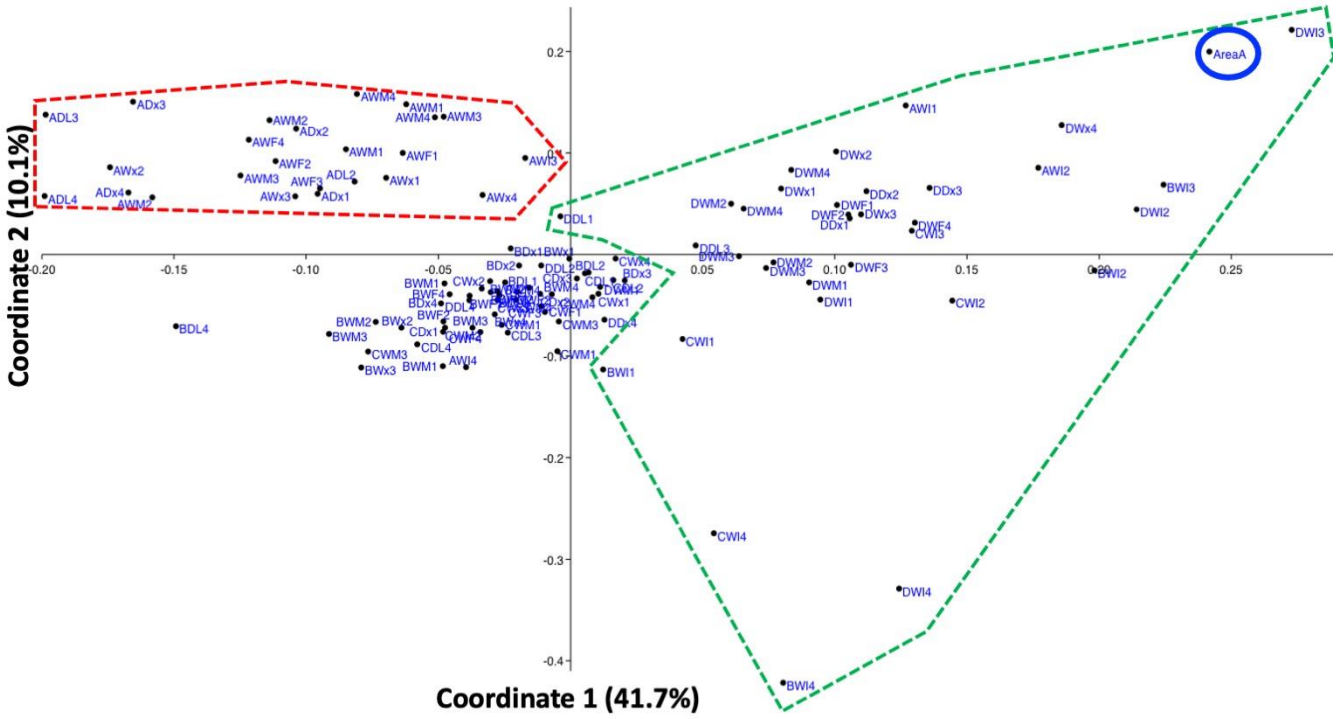


Fig. 3.12: Principal coordinates ordination of fungal populations over the course of the pot experiment. Most of the samples from the baseline harvest (A; red dotted polygon) ordinated together, as do most of the samples from the last harvest (D; green dotted polygon).

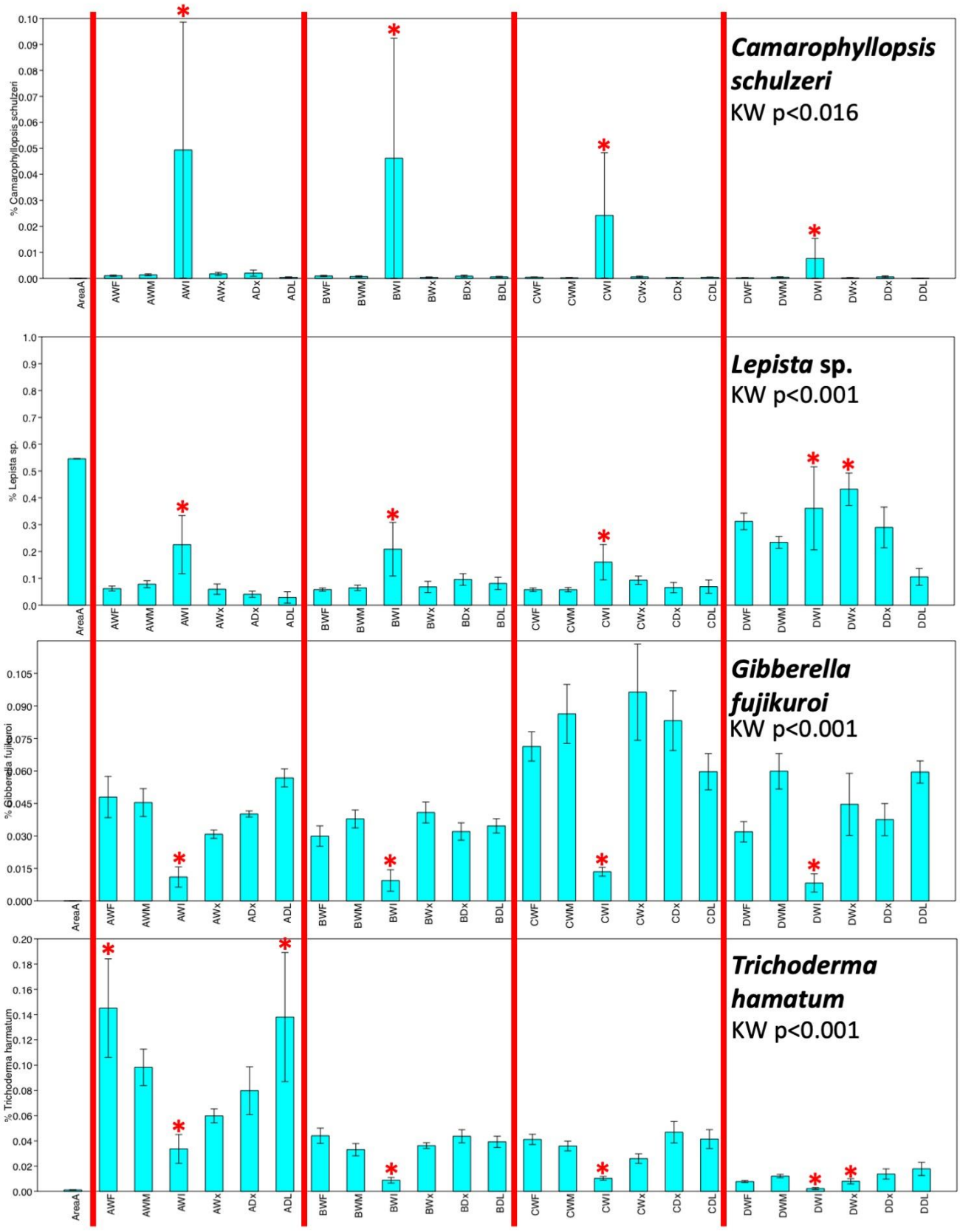


Fig. 3.13: Mean percentage abundance (as % of total fungal DNA) at each harvest timepoint of the four most abundant fungi detected in the pot experiment soils. Error bars indicate SE (n=4). Significant treatment effects are indicated by *.

4) Testing of different soil sampling and storage methods

Hitherto when deploying eDNA methods for soil analysis, we have either sampled soil ourselves using our standard 6x6 grid soil coring system, and transported the samples back to Aberystwyth on the same day in a coolbox, or we have received frozen samples via overnight couriers. Frozen samples sent wrapped in bubblewrap and/or in insulated coolboxes remain well frozen in transit in our experience.

However, standard soil sampling as undertaken by WG involves sampling in a zigzag manner across a field (W-sampling) which covers a larger area, and data soil samples derived from use of this method may not be directly comparable to samples collected by the grid method.

Additionally, the requirement for the soil collector to freeze samples adds an additional logistical complexity to the process, for example if there is limited availability of storage space or if a series of samples are collected over a period of several days of fieldwork. Therefore, we undertook an experiment to determine the effect of a range of contrasting soil sampling and storage methods on subsequent DNA metabarcoding results.

Soil was collected from Trawsgoed Aberystwyth University Farm from four fields (14th November 2018), aiming to compare differences between two different collectors (GWG and LAC) and by different sampling methods (Grid vs. W-transect method). All four fields were highly improved and dominated by *L. perenne* but three of the fields (labelled Top/Mid/Low) had recently been reseeded (Sept/Oct 2018; Fig. 4.1; 4.2) following ploughing/harrowing earlier that summer. The 4th field (UP) was scheduled for ploughing reseeded in spring 2019.

A 30x30 m quadrat was set in all four fields and soil cores collected according to our standard procedure in a 5x5 m grid (36 cores). In the 4th field the (UP) quadrat was sampled three times, two times using the grid method by different people (GWG/LAC) and then again using the W-transect method. Additionally, the UP quadrat was sampled a fourth time, using a more intensive sampling method (cores taken on ca. 2x2 m grid in order to obtain about 2.5 kg of soil for the storage experiment (this soil had a pH of 4.86 and electric conductivity of 278 µS).

Following collection, all samples except the 3 kg soil for the storage experiment were frozen, freeze-dried and processed as described above (2). For the storage experiment the soil was roughly sieved to 5 mm and thoroughly mixed, immediately after collection. Subsamples of the mixed soil (200g) were stored under a range of conditions prior to freezing at -80°C (Table 4.1; Fig. 4.4 A,B,C). All samples were then freeze-dried, ground finely and subject to DNA extraction as for other samples described above.

Table 4.1: Storage conditions for soil samples

Code	Container	Storage Regime			Final Steps	
FZ-80	Sealed ziplock bag	Freeze at -80°C	_	Continuous -80°C freezing	Freeze at -80°C	Freeze-dry
FT14dFR	Sealed ziplock bag	Freeze -20°C overnight	Thaw	14 days in 4°C fridge	Freeze at -80°C	Freeze-dry
FT5dFR	Sealed ziplock bag	Freeze -20°C overnight	Thaw	5 days in 4°C fridge	Freeze at -80°C	Freeze-dry
FT14dRT	Open tray	Freeze -20°C overnight	Thaw	14 days at room temp	Freeze at -80°C	Freeze-dry
FT5dRT	Open tray	Freeze -20°C overnight	Thaw	5 days at room temp	Freeze at -80°C	Freeze-dry
O5dRT	Open tray	No freezing	_	5 days at room temp	Freeze at -80°C	Freeze-dry
O5d37C	Open tray	No freezing	_	5 days on 37°C air-dryer	Freeze at -80°C	Freeze-dry
S14dFR	Sealed ziplock bag	No freezing	_	14 days in 4°C fridge	Freeze at -80°C	Freeze-dry
S14dRT	Sealed ziplock bag	No freezing	_	14 days at room temp	Freeze at -80°C	Freeze-dry
S5dRT	Sealed ziplock bag	No freezing	_	5 days at room temp	Freeze at -80°C	Freeze-dry

Fig. 4.1. Map of Tawscoed Farm (A) showing the main farm buildings (arrowed) and the study area. The study area (B) comprised four field, three of which had been ploughed and reseeded a few months (Top, Mid, Low) prior to sampling and the fourth scheduled for reseeding in the following spring.

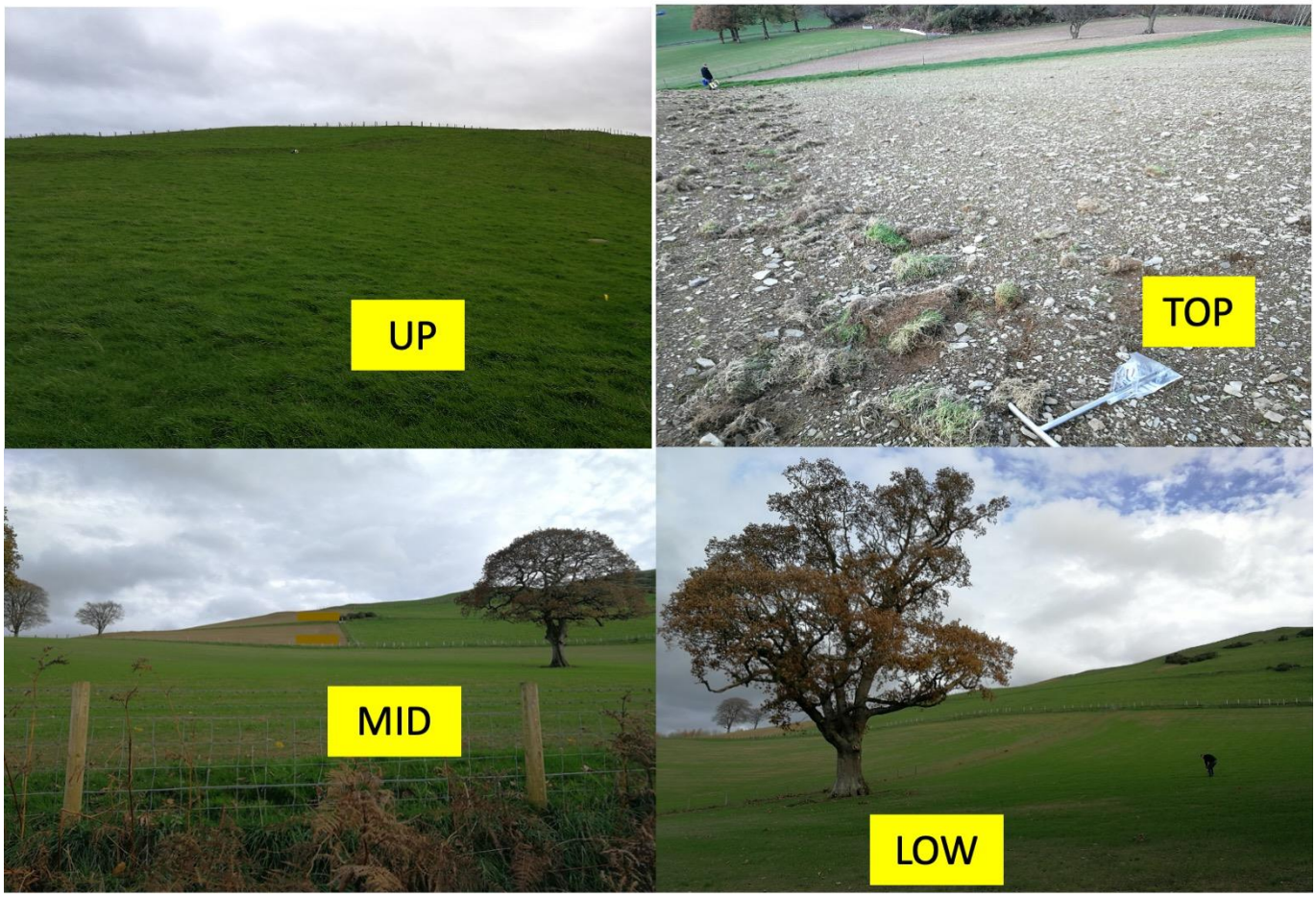


Fig. 4.2. Images of the four field sampled at Tawscoed Farm.

4A) Effect of different soil sampling method

Comparison of the three samples taken from the same quadrat (UP), by two different samplers using the grid method (UP-GWG/UP-LAC) or using the W-transect method (UP-W) showed that the UP-W sample was the outlier in the set when the data were subject to PCO ordination and also had slightly fewer species that the two grid-sampled soils (12 vs. 13) (Table 4.2; Fig. 4.3.). It is likely that this difference is due to the fact that the total amount of soil collected by the W-transect method was lower (25 cores; 496 g) that for the grid method (36 cores; 697 g/860 g).

4B) Comparison of the four fields at Trawsgoed

Trawsgoed farm is intensively managed and most areas have been ploughed and reseeded on a regular 5-10 year cycle. However, it was suspected that one field (TOP), located on a steep slope had not been ploughed until recently (2017). Therefore the four fields were compared to determine whether eDNA analysis could identify signals of the past vegetation. The three lower fields were clearly recently reseeded with *L. perenne*, and patches of bare were soil visible (Fig. 4.3), especially in the TOP field (due to dry weather in late summer). The uppermost field (UP) was known to have been reseeded ca. 5 years ago.

The number of plant species detected in these fields was also low (8/11/12; Table 4.1), including the presence of a some *Quercus robur* in the LOW field where a large tree of this species was growing and the presence of some weed species (*Chenopodium album*/*Stellaria media*). Unexpectedly in the TOP field, evidence of two non-grass crops, *Brassica napus/oleracea/rapa* and *Secale cereale* (rye), was detected. Precise identification of the *Brassica* species was not possible because crop plants (e.g. turnip, fodder rape, kale) tend to be heavily cross-bred and thus their ITS2 sequences are often overlapping. However, later discussion with Stephen Jones, farm manager, revealed that fodder rape and rye had been sown as cover crops to protect the late-sown *L. perenne*. Mr. Jones reported that the field had actually been ploughed in September 2017 but not harrowed and reseeded until 12 months later. Prior to ploughing, he recalled that the vegetation comprised mainly ‘weed grasses’, thistles and nettles. As such 14 months had elapsed since the ploughing event and no sign of these herbs was detected. However, the relative abundance of *Agrostis capillaris/gigantea* was high and some *Cynosurus cristatus* was detected (but these may have regrown in the spring following the ploughing).

Table 4.2. Composition of vegetation as assessed via eDNA in the Trawscoed Farm study area.

	Fresh wt. (g)	860.4	697.3	**496.1**	907.7	1006.4	1001.0
	Moisture content (%)	42.4%	42.3%	42.5%	27.7%	26.2%	27.2%
	Species	**UP-GWG**	**UP-LAC**	**UP-W**	**LOW**	**MID**	**TOP**
1	*Agrostis capillaris/gigantea*	7.57%	6.55%	8.12%	0.07%	0.72%	17.13%
2	*Agrostis stolonifera/canina*	0.18%	0.07%	0.26%	0.04%	0.30%	0.16%
3	*Anthoxanthum odoratum*	0.01%	0.02%	0.01%			
4	*Brassica oleracea/rapa*						11.71%
5	*Cerastium fontanum/glomeratum*	0.92%	0.12%	0.36%		0.07%	
6	*Chenopodium album*					0.10%	
7	*Cirsium vulgare*	1.29%	0.06%	0.03%		0.09%	
8	*Cynosurus cristatus*						0.12%
9	*Festuca rubra*	0.21%	0.94%	0.30%		0.01%	
10	*Holcus lanatus*	2.94%	2.08%	4.06%			0.04%
11	*Lolium perenne/multiflorum*	70.20%	63.71%	36.96%	50.57%	95.00%	57.86%
12	*Phleum pratense*	0.61%	0.77%	1.30%			0.02%
13	*Poa annua*	0.14%			4.33%	0.42%	
14	*Poa pratensis*	0.30%	0.13%	0.17%			0.01%
15	*Poa trivialis*	15.12%	19.09%	19.86%	10.31%	0.28%	4.21%
16	*Quercus petraea*				0.03%		
17	*Secale cereale*						7.46%
18	*Stellaria media*					0.78%	
19	*Taraxacum officinale* agg.		0.10%			0.21%	
20	*Trifolium pratense*				3.56%		
21	*Trifolium repens*	0.51%	6.36%	28.55%	31.09%	2.03%	1.13%
	No. species detected	**13**	**13**	**12**	**8**	**12**	**11**

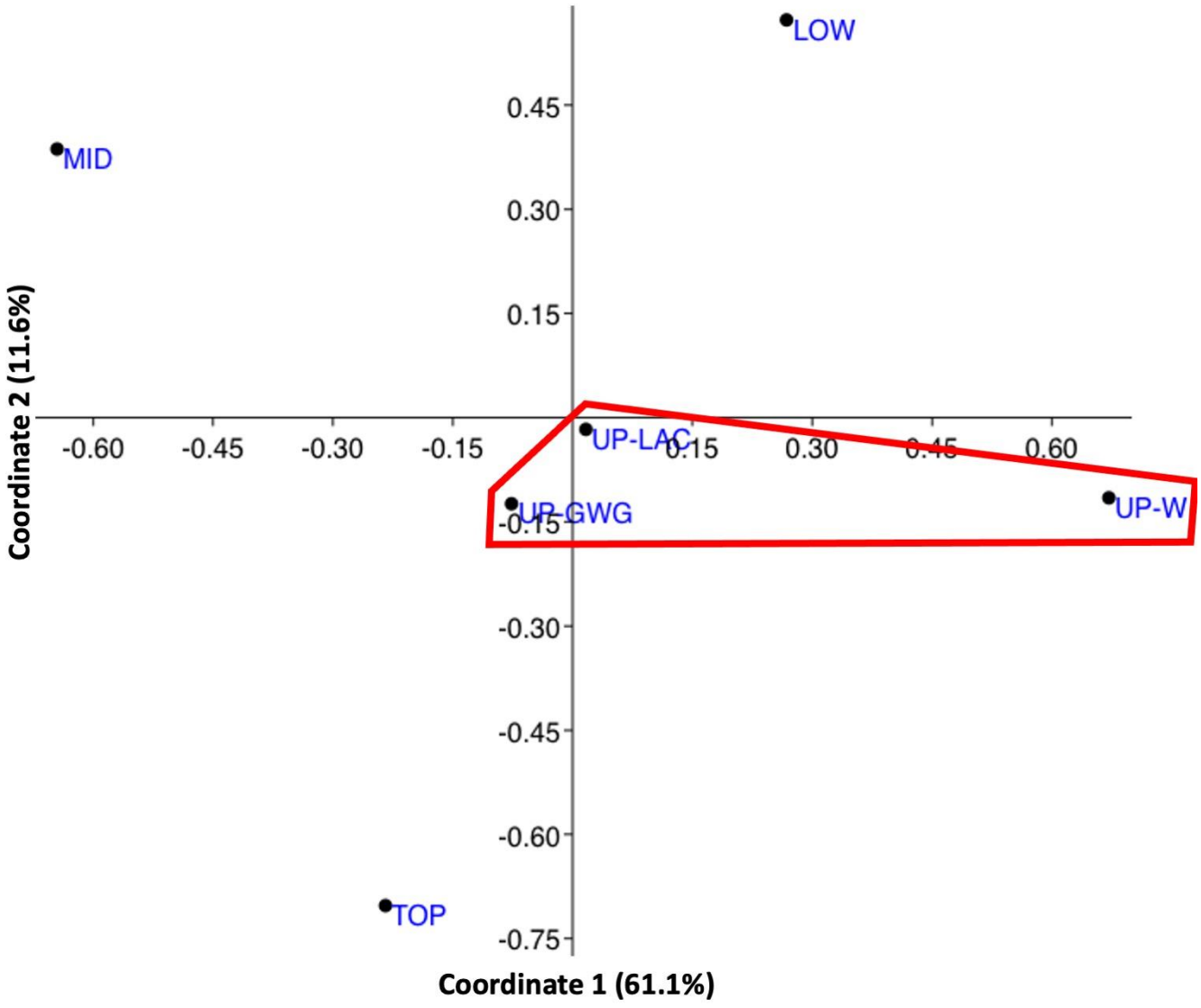


Fig. 4.3. PCO ordination of Trawscoed quadrats. The three samples taken by different methods/surveyors from the same quadrat in the upper non-reseeded field are contained within the red polygon.

4C) Effect of different soil storage conditions

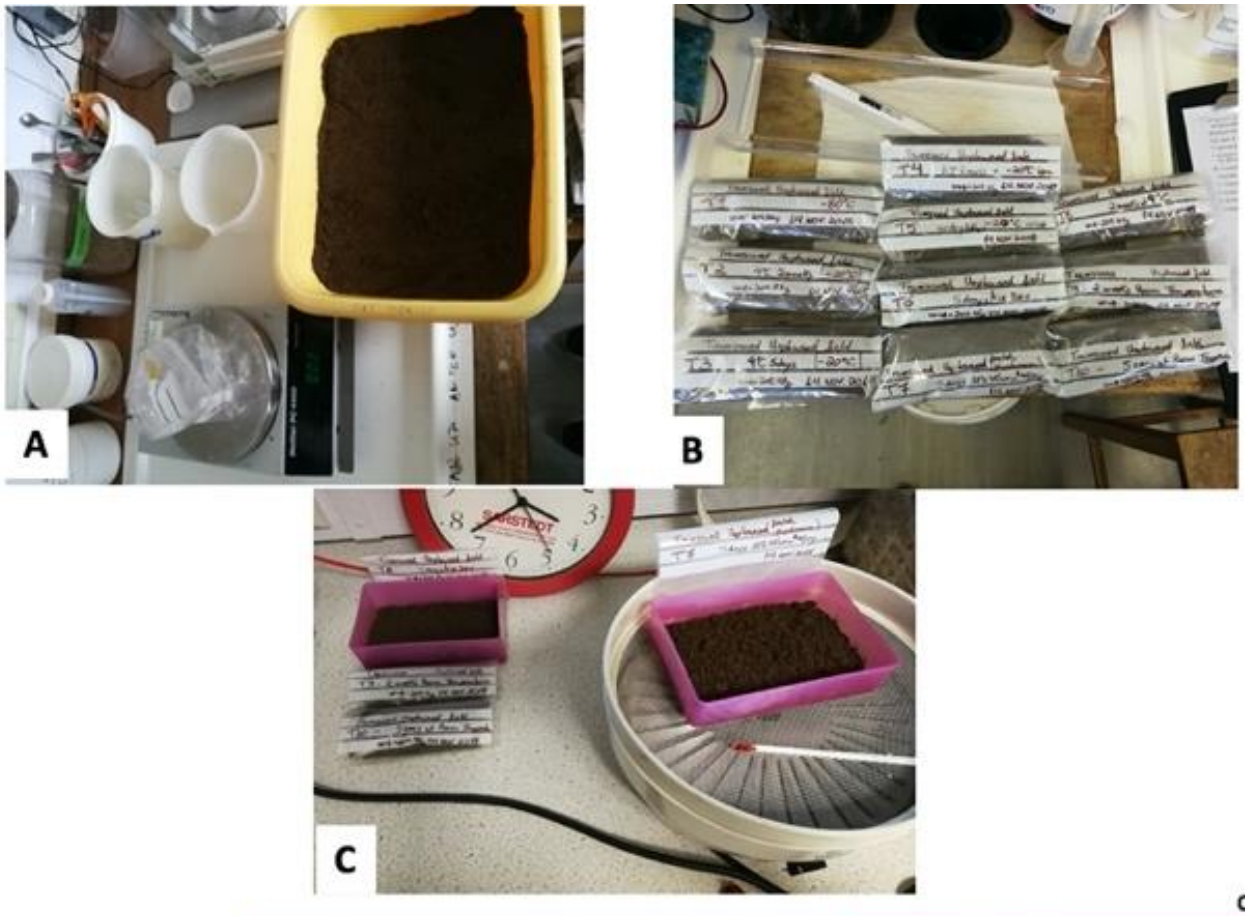
The primary aim of this experiment was to determine what conditions, where immediate freezing until freeze-drying are best for preservation of plant eDNA from soil samples. However, it is of interest also to monitor changes in fungal communities, since it would be anticipated that under suboptimal storage conditions, proliferation of saprotrophic decay fungi would occur. A range of soil storage conditions were tested as shown in Table 4.1 and Fig. 4.4A, B, C.

Our standard mix of plant and fungal primers (see above) was used for the DNA metabarcoding and it was apparent that the relative numbers of plant and fungal sequence reads varied according to treatment. For the sample frozen immediately (T1), higher plants comprise 58% of all the sequences retrieved (Fig 4.4D). However, for several of the treatments where soil was frozen and thawed prior to incubation (prefix FT; FT14dFR, FT14dRT, FD5dRT) and for soil stored at ambient temperature for 14d (S14dRT), this proportion was below 30% suggesting a proliferation of fungi and likely a resultant decay of plant eDNA. Examination of the specific fungi detected in the soil revealed that Mortierellomycota (the phylum to which the two most abundant fungal species belonged) and *Metarhizium carneum* were several fold more abundant in samples where the proportion of plant eDNA was reduced. These species are known to be chitinolytic (Gray and Baxby, 1968; Jackson, 1965), potentially proliferating during storage due to their ability to degrade chitin from fungal hyphae degrading during the storage process.

Excluding algae and bryophytes, 25 species of (higher) plant were detected in the soil samples. Therefore, Principal Coordinates ordination was conducted (Fig 4.5A), alongside the data obtained from adjacent fields at Trawsgoed (see above 3D). The control stored sample ordinated close to samples taken from the same quadrat (UP-GWG, UP-LAC-UP-W) and frozen immediately (prior to any sieving), whereas samples from other field ordinated quite distinctly from the other samples.

For the plant eDNA community composition, the most divergent storage treatments from the ‘control (FX-80) were storage at 4°C in a sealed bag for 2 weeks (S14dRT) and air-drying at room temperature (O5dRT). Also divergent from the control were samples subject to warm air-dying (O5d37C) and freeze-thaw followed by prolonged storage (14d periods; FT14dRT/S14dRT).

The fungal communities present were also subject to PCO ordination, the rationale being that storage conditions where significant levels of decay of the plant biomass had occurred would also be the ones where particular communities of saprotrophic fungi would have developed and this would correlate with a general increase in abundance of fungal sequences relative to plant sequences (Fig. 4.5B).



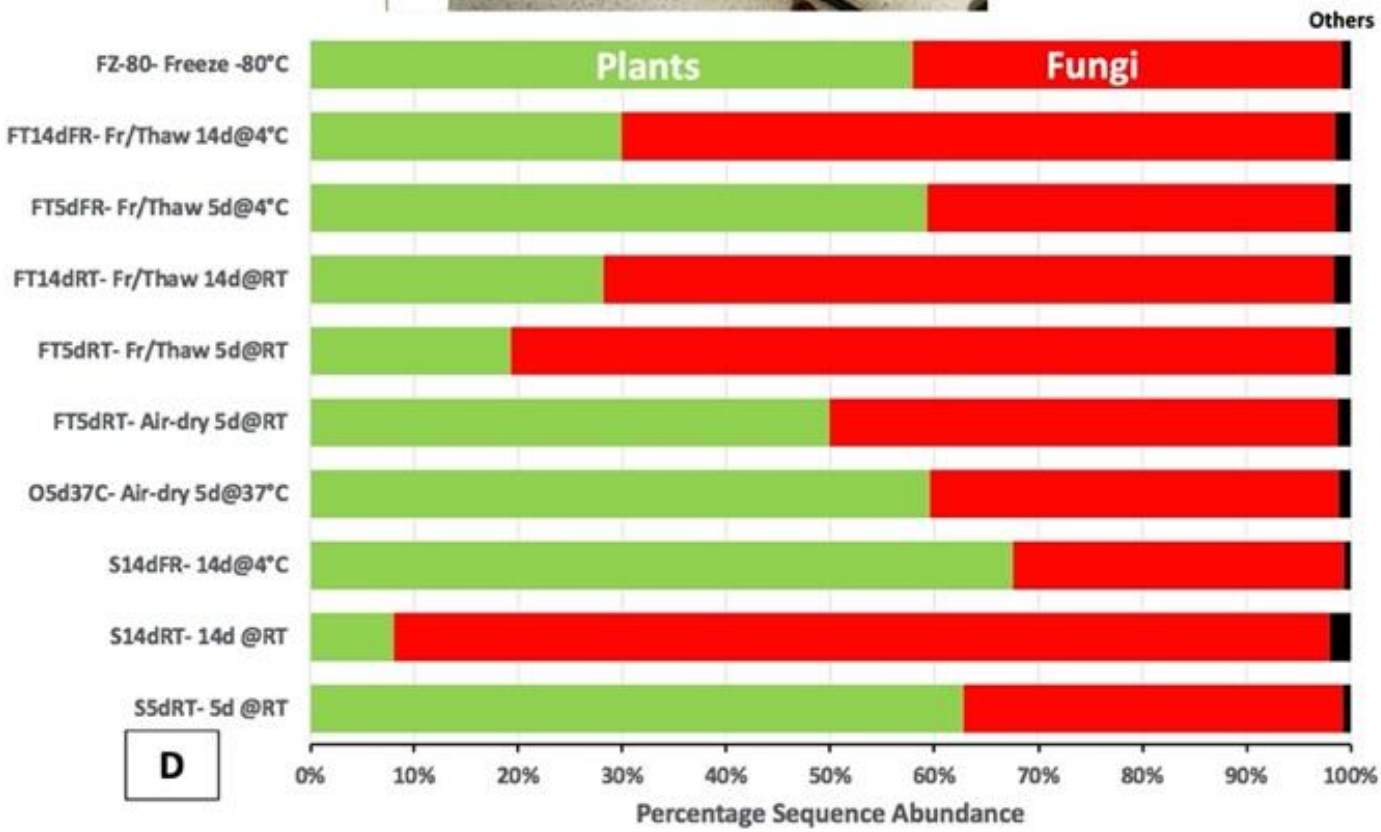


Fig 4.4. Soil from Trawscoed were roughly ground, homogenized and divided into 200g aliquots. The subsamples were subject to a range of storage conditions including freeze-thawing in sealed bags (B), air-drying (C) etc. The relative proportions of plant and fungal DNA also varied across treatments when assessed by via eDNA metabarcoding (D).

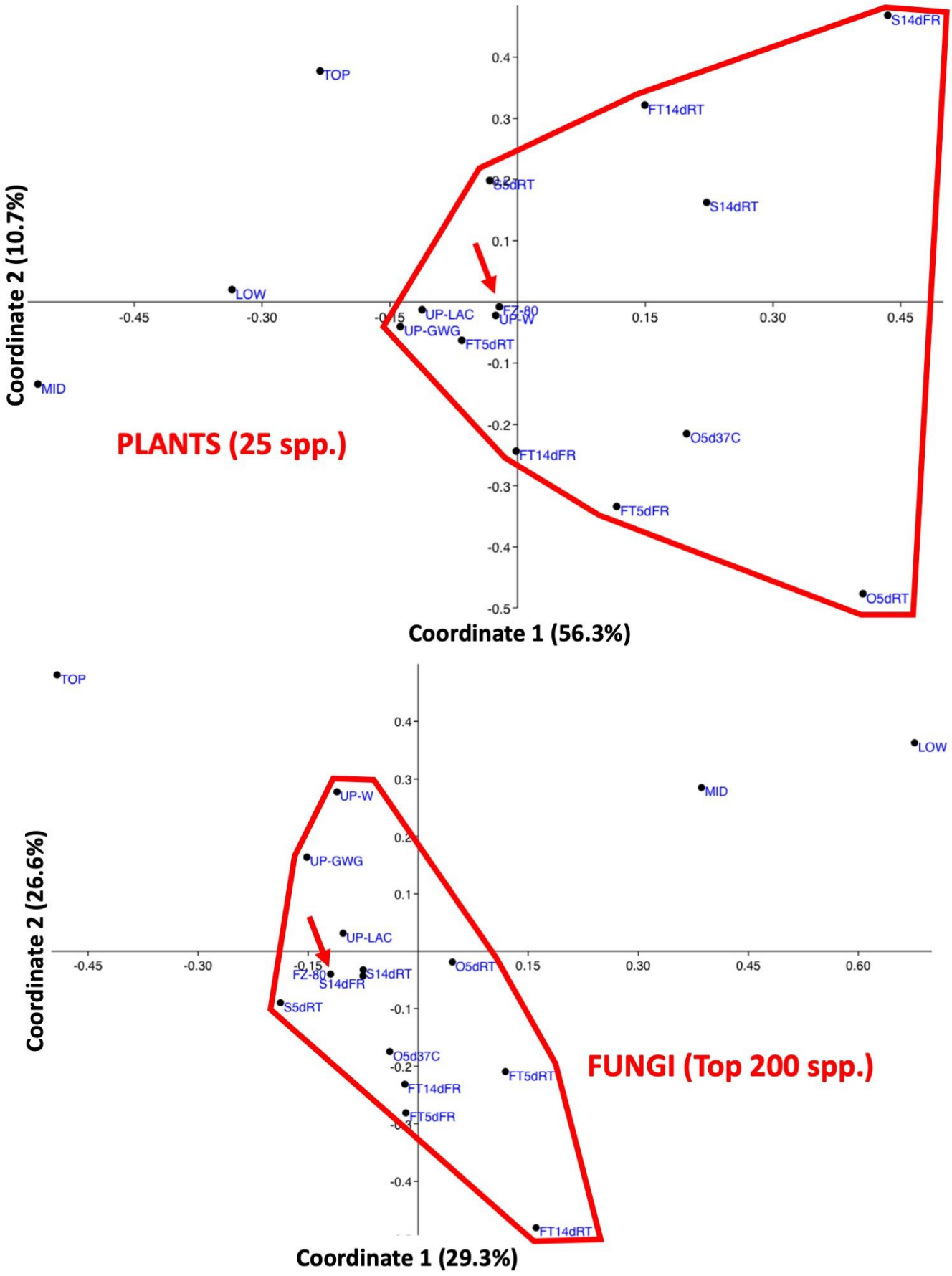


Fig 4.5. Principal Coordinates ordination of the plant (A) and fungal (B) communities subsequently detected via eDNA metabarcoding shows how the various treatments differ from the original (instant -80°C freezing; red ARROW). All the samples derived from the upper field (UP) are enclosed in the red polygon.

The treatments most divergent from the control were also the ones with the lowest plant eDNA relative abundance (Fig 4.4D). Similar ordination of fungal communities (Fig. 4.5B) also found that the same treatments (T4/T5/T9/T10) were divergent from the control. Without more detailed study using full replication (such an experiment is planned for the future), it is difficult to determine clearly which of the other treatments would be recommended but cold storage without freezing for a up to 5 days would not lead to substantial loss of plant eDNA; such treatment would be compatible with short term storage of samples in a refrigerator prior to postage via overnight delivery.

NOTE: The interesting data obtained during the experiment outlined above led us to undertake a more detailed replicated experiment. The conclusion of this later study was consistent, namely that cold refrigerated storage for up to a week led to only minimal changes in the plant/fungal eDNA later detected. This element of the work has recently been published in a peer-reviewed journal:-
Clasen, LA, AP Detheridge, J Scullion, GW Griffith (2020). Soil stabilisation for DNA metabarcoding of plants and fungi. Implications for sampling at remote locations or via third-parties. *Metabarcoding and Metagenomics*, DOI 10.3897/mbmg.@.58365

Discussion

In the course of this project, we have developed a method for the eDNA detection of both plants and fungi from soil samples. This process involved not only the development of methodologies for field sampling, extraction of DNA and amplification of the ITS2 target locus but also the associated bioinformatic pathways that allow reliable linkage of DNA sequences to species names. This linkage was successfully achieved although in a few cases the ITS2 barcodes did not allow differentiation of pairs/clusters of closely related species, usually because these species are known to form hybrids in nature.

We are not aware of the deployment of eDNA approaches in a regulatory/site monitoring context. It is possible that such material published in the 'grey' literature (government reports etc.) are not straightforward to locate compared to articles in the peer-reviewed literature. The discussion below will focus primarily on matters relating to the use of this methodology in a regulatory/site monitoring context.

♦Assessment of current plant communities via soil eDNA.

In order to determine the potential of metabarcoding for analysis of plant-derived eDNA from soil, we first deployed the method in a semi-natural grassland to examine how well it could analyse existing plant communities in undisturbed habitats. The method was compared to the well-established percentage cover method widely used in vegetational analysis. Comparisons were made at three distinct fieldsites (Brignant, Kirby Muxloe and Turlough). The congruence between the two methods was good but in the context of the interpretation of eDNA data in a regulatory context, the following points are of relevance:

Windblow of vegetation:

We identified several occasions where dead leaves blown into a field and had entered the soil system, subsequently being detected in our soil eDNA metabarcoding datasets. The amounts were very low and in most cases were consistent with plants present nearby (typically trees on adjacent field boundary). It is possible that other more exotic plants could be detected, for example if for example garden compost had been spread on field. However, if in total such exotic species comprise <1%, they do not confound attempts to reconstruct the indigenous plant community.

Rarer plants:

Semi-natural habitats are often valued due to the presence of particular rare species (e.g. orchids), which are indicators of past lack of disturbance. Such plants may only present as a few tens or hundreds of individuals and thus comprise only a small proportion of the total plant community even in the areas where they are most dense. For soil sampling to detect these species, it would be necessary for part of the root system of these plants to be cored and if the chance of this happening is low then it is likely that the species will not be represented in the DNA present in the soil core, even though to a botanist observing the sward, the presence of the plant is obvious. For example, at Brignant, the distinctively blue-flowered *Campanula rotundifolia* was

recorded at 1% of total cover in one quadrat, its DNA was not detected in the soil of that quadrat.

c) Effects of root architecture, seed banks and eDNA abundance vs % cover.

Plant species vary considerably in the architecture of their root systems (e.g. grasses with adventitious roots systems vs. many Asteraceae with taproots). These differences will influence the way in which root biomass is 'captured' by soil coring, and also the rate at which dead roots decay (e.g. taproot might be expected to persist longer following plant death). Similarly, perennial plants invest more heavily in root tissues whereas annual plans invest more heavily in seed production and root:shoot ratios will likely vary seasonally as will the seed bank.

Thus in order to calibrate eDNA abundance vs. % cover based on vegetational analysis, investigation of individual species growing in monoculture would be necessary in order to obtain accurate correlation. However, the data presented above for eDNA vs % cover data do identify some common features. For example, that grass abundance is consistently higher in eDNA than botanical surveys; this is more likely due to the subjectivity of % cover assessments, with the human eye being drawn more to distinctive leaf shaped and flowers than grass leaves. It is also apparent that botanical surveyors may rely heavily on the presence of flowering heads to detect the presence of a species in a sward; at both Brignant and Kirby Muxloe, the late-flowering *Phleum pratensis* was detected via eDNA but not recorded in vegetational analysis.

The unusually high abundance, in eDNA surveys, of certain taproot-forming species (eg, *Taraxacum officinale* [68% at quadrat KX15B vs 2% for % cover in that quadrat; Table 2.1]) suggests that the taproot was directly cored in this quadrat. At the Turlough site or *Pimpinella saxifraga* comprised 27-81% of total eDNA across the four quadrats suggesting that taproots were sampled in all quadrats.

Relative abundance of key indicator species.

For EIA assessment, the key concern relates to the total abundance of sown species (mainly ryegrass [*Lolium* spp.] and white clover [*Trifolium repens*]), with 25% being the threshold level. Thus examination of the recently reseeded fields at the Trawsgoed fieldsite (Table 4.2), *Lolium* spp. comprised 37-95% of all the quadrats, with *T. repens* being present at more variable levels. Reliance of *T. repens* as an indicator of recent reseeding is less reliable, as indicated at Kirby Muxloe, where two quadrats (KX10D and KX11D) contained >30% *T. repens* according to eDNA (5% according to vegetation analysis) but the same quadrats also contained indicators of semi-natural vegetation (*L. corniculatus*).

♦Degradation of residual plant DNA in soil.

The ultimate aim of this project was to establish whether it was possible to identify plant residues sampled via soil sampling and thus to reconstruct the pre-existing plant communities from recently disturbed agricultural sites. The rate at which plant tissues (including their DNA) degrades in soil depends primarily on climatic factors (rainfall, temperature) but can be influenced by edaphic factors (e.g. nutrient levels) and also subsequent land management (e.g. cropping regime). Other factors such as whether

weedkiller is applied prior to ploughing, the timing and nature of the ploughing / harrowing that is undertaken and the delay between ploughing and reseeding may also vary considerably.

Therefore, we attempted to constrain these variables by undertaking a mesocosm experiment using soil from a diverse semi-natural grassland and applying various treatments to the soil once it had been disrupted and sieved. Once established, soil cores were taken from the mesocosm pots at 7d, 1 mth, 3mths and 12mths after initiation of the treatment regimes. The source habitat was reported to contain 47 spp. of higher plants/mosses. eDNA analysis revealed the presence of 26 of these species (plus additionally 5 mosses and three grasses not detected by NVC). Over the course of the experiment there was a progressive decline in the amount of plant biomass detected (Figs. 3.5/3.6) and an increase in fungal biomass, due to degradation of plant litter and consequent proliferation of fungal decomposer biomass.

A month after soil disturbance and establishment of the pots, 25 spp. were still detectable in the pots and after 3 months, 21 spp. (Figs. 3.9/3.10/3.11). However, after 12 months, only 16 spp. were detected and in pots which had been reseeded with ryegrass (as opposed to being kept bare), only 9 spp. were detected. In the bare soil pots, the herbs which were sill detected after 12 months were *Cardamine pratense*, *Crepis capillaris, Hypochaeris radicata*, *Plantago lanceolata, Ranunculus acris, Ranunculus repens, Rumex acetosa, Taraxacum officinale and Trifolium repens*). Plants which were present at higher abundance initially were more likely to be detected after 12 months and it is likely that some plants had regrown to some extent from seeds/rhizomes (despite regular 'weeding'). For ryegrass seeded pots, only non-grasses were removed by weeding but residual DNA of some forbs was detected after 12 months (*R. acetosa*, *C. pratense*).

There was large variation between replicate pots in the experiment, due to difficulty in ensuring soil homogeneity following disruption and as a consequence no significant differences between the various treatments were found. Nonetheless, residues of most of the herb species detected initially in the soil were still detectable after 12 mths. The extent to which this experiment mimics what would happen under field conditions is difficult to assess. In this experiment the sward and soil was broken up much more finely than would occur during ploughing/harrowing (with smaller fragments of vegetation likely to decompose more rapidly as a consequence) and regrowth was prevented by weeding, so it might be expected that under standard agricultural conditions the degradation of residues from pre-existing plant communities would be slower (NB. the effects of pre-ploughing glyphosate application in this context are not known).

Species present initially at higher abundance persisted for longer, so it may be the case that more sensitive methods of detection would have detected more species after 12 months. Thus methods such as qPCR targeting particular marker species (e.g. *Lotus corniculatus*/*Luzula campestris* etc.) would be useful, in addition to metabarcoding (see below).

♦Utility of soil fungal community analysis in assessment of site disturbance.

Though less intuitive to the layperson, the fungal communities in soil can be highly informative about the level of past disturbance and of the current plant community. Many fungi have specific association with particular plants, for instance as pathogens, endophytes or mycorrhizal partners, so detection of particular fungi is a reliable indicator of the presence of certain plants or vegetation types. Additionally, many larger (mushroom-forming) fungi are long-lived and very slow growing, so their presence indicates that a site has long been undisturbed (i.e. on decade timescales in the case of waxcap fungi). Evidence from the present study, showed how following disturbance, particular changes occur within the fungal community which can be monitored over time (in this case up to 12 months) but still retain information about the communities which were present prior to the disturbance event. Because the species richness of fungal communities in soils is greater than for higher plants and these fungi are less heterogeneously distributed, some assessment of fungal communities could prove very useful in providing evidence of the nature of pre-disturbance plant communities.

♦Sampling and storage.

For practical use, soil sampling methodologies need to be clear, simple and time-efficient. Comparison of grid sampling with the W-transect sampling method found little difference. Of greater significance, is the weight of soil collected from a given area, with more intensive sampling leading to collection of more cores; whilst this would increase the number of plant species recovered in soil cores but would take more time and increase time taken for downstream processing. Typically, we have opted for collection of ca. 40 cores weighing ca. 800-1000g which can be completed within 30 min. We also established that soil samples do not degrade discernibly if stored refrigerated (NOT frozen) and that door-to-door next-day courier delivery to the processing laboratory is economical (cost is £1-4 per kg depending on total consignment weight [https://www.parcelhero.com/]).

♦Future developments.

Improvement of quantification of plant biomass by species

We have discussed above the relationship between percentage cover of (above-ground) plant communities and eDNA of these species identified by soil coring. The nature of this correlation depends on various factors including rooting structure of the plants and the growth phenology. More detailed investigation of ‘mock’ (synthetic plant communities, comprising root tissues from plants grown in monoculture could be used to calibrate such comparisons, for example for species of particular interest including ryegrass/clover or key indicators of semi-natural grasslands (*Lotus corniculatus* etc.).

Improvement of sensitivity

In addition to eDNA metabarcoding, where the whole plant/fungal community is assessed other DNA-based methods (e.g. quantitative PCR) can provide more rapid and sensitive quantification for specific target species and these assays are amenable to multiplexing (simultaneous detection of up to 6 target species) and to some extent to field deployment (LAMP assays). Such assays can be useful for detection of specific plant pathogens (Khaliq et al., 2018) or invasive plants/animals.

Additional application of eDNA metabarcoding technology

DNA-based methods of detection hold huge potential in many fields of human endeavour and this is particularly true in the context of agriculture nature conservation. For example, we already undertake surveys for waxcap populations in grasslands on behalf of NRW/NE which have greatly facilitate the process of SSSI notification, site monitoring and initial site screening. For botanical analyses, the standard methods of vegetational analysis are better developed and more efficient that for fungi but as is apparent from this study, the detection and identification of grasses/mosses via eDNA is highly effective and superior to standard botanical surveying. eDNA metabarcoding is also more flexible since survey periods are not restricted to high summer and accurate assessment of plant species richness and NVC classification can be obtained from sampling soil during the winter period.

Conclusion

The development of the fungal eDNA process has been successfully transposed to include data for higher plants, the primary endeavour of this project. However, without further testing of 'real world' samples, the method is not a 'Golden Bullet' which can be utilised in isolation to prove that a breach of the EIA (Agri) Regulations has occurred, or that a Remediation site is reverting to its pre-agricultural intensification quality. This project found that:

- Data from eDNA analysis is consistent with more traditional surveying techniques. However, care must be taken with factors such as windblown contamination, missing of rarer species, and plant root architecture possibly skewing results. Although the accuracy of percentage cover of grass species does appear to be more accurate in the eDNA process.

- The project also found that the amount of plant material does degrade over time in the soil. This shows that intervention and soil collection must be carried out as soon as possible to give an accurate as possible picture of the previous sward.

- Specific searches of 'key indicator' species may prove more successful on more historic soil sample, when more than 3-6 months has elapsed since the event.

- Data derived from fungal residue can be highly indicative of original sward constituents (or at least past management practices) even after 12 months has elapsed since the event.

- Inappropriate storage during transit and associated DNA degradation could alter the plant communities later discovered. Therefore, a range of storage conditions were tested and it was found that refrigerated storage for period of up to a fortnight led to only minimal changes. Freezing of soil immediately after sampling is not necessary to preserve the DNA if cooling facilities are available.

- The current W-transect techniques used by the EIA Unit was found to be suitable for sample collection. That the collection of 800 - 1000g of soil is achievable in 30mins, and provides suitable size of sample for analysis.

References

Ankenbrand, M.J., Keller, A., Wolf, M., Schultz, J., Förster, F., 2015. ITS2 database V: Twice as much. Molecular Biology and Evolution 32, 3030-3032.

Bell, K.L., Loeffler, V.M., Brosi, B.J., 2017. An rbcL reference library to aid in the identification of plant species mixtures by DNA metabarcoding. Applications in plant sciences 5, 1600110.

Brennan, G.L., Potter, C., De Vere, N., Griffith, G.W., Skjøth, C.A., Osborne, N.J., Wheeler, B.W., McInnes, R.N., Clewlow, Y., Barber, A., 2019. Temperate airborne grass pollen defined by spatio-temporal shifts in community composition. Nature Ecology & Evolution 3, 750-754.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PloS one 5, e8613.

de Vere, N., Rich, T.C., Ford, C.R., Trinder, S.A., Long, C., Moore, C.W., Satterthwaite, D., Davies, H., Allainguillaume, J., Ronca, S., 2012. DNA barcoding the native flowering plants and conifers of Wales. PloS one 7, e37945.

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

Fahner, N.A., Shokralla, S., Baird, D.J., Hajibabaei, M., 2016. Large-scale monitoring of plants through environmental DNA metabarcoding of soil: recovery, resolution, and annotation of four DNA markers. PLoS One 11, e0157505.

Felsenstein, J., 1988. Phylogenies from molecular sequences: inference and reliability. Annual review of genetics 22, 521-565.

Gray, T., Baxby, P., 1968. Chitin decomposition in soil: II. The ecology of chitinoclastic micro-organisms in forest soil. Transactions of the British Mycological Society 51, 293-309.

Griffith, G.W., Cavalli, O., Detheridge, A.P., 2019. REPORT-An assessment of the fungal conservation value of Hardcastle Crags using NextGen DNA sequencing (NECR258). Natural England Commissioned Report NECR258 (http://publications.naturalengland.org.uk/publication/4543317115404288).

Griffith, G.W., Hedger, J.N., Rowland, A., 2020. A comparison of fruitbody and eDNA survey approaches for assessing the distribution of mycelia of macrofungi in the grassland and heathland of Lundy. Journal of the Lundy Field Society 7, 87-106.

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.

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.

Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M.W., Cowan, R.S., Erickson, D.L., Group, C.P.W., 2009. A DNA barcode for land plants. Proceedings of the National Academy of Sciences 106, 12794-12797.

Jackson, R., 1965. Studies of fungi in pasture soils: III. Physiological studies on some fungal isolates from the root surface and from organic debris. New Zealand Journal of Agricultural Research 8, 878-888.

Khaliq, I., Hardy, G.E.S.J., White, D., Burgess, T.I., 2018. eDNA from roots: a robust tool for determining Phytophthora communities in natural ecosystems. FEMS microbiology ecology 94, fiy048.

Koetschan, C., Förster, F., Keller, A., Schleicher, T., Ruderisch, B., Schwarz, R., Müller, T., Wolf, M., Schultz, J., 2009. The ITS2 Database III—sequences and structures for phylogeny. Nucleic acids research 38, D275-D279.

Koljalg, U., Larsson, K.H., Abarenkov, K., Nilsson, R.H., Alexander, I.J., Eberhardt, U., Erland, S., Hoiland, K., Kjoller, R., Larsson, E., Pennanen, T., Sen, R., Taylor, A.F.S., Tedersoo, L., Vralstad, T., Ursing, B.M., 2005. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. New Phytologist 166, 1063-1068.

Kõljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F., 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.
Leff, J.W., Bardgett, R.D., Wilkinson, A., Jackson, B.G., Pritchard, W.J., Long, J.R., Oakley, S., Mason, K.E., Ostle, N.J., Johnson, D., 2018. Predicting the structure of soil communities from plant community taxonomy, phylogeny, and traits. The ISME journal, 1.
Matesanz, S., Pescador, D.S., Pías, B., Sánchez, A.M., Chacón-Labella, J., Illuminati, A., de la Cruz, M., López-Angulo, J., Marí-Mena, N., Vizcaíno, A., 2019. Estimating belowground plant abundance with DNA metabarcoding. Molecular ecology resources.
Moorhouse-Gann, R.J., Dunn, J.C., De Vere, N., Goder, M., Cole, N., Hipperson, H., Symondson, W.O., 2018. New universal ITS2 primers for high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. Scientific reports 8, 8542.
Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., Erlich, H., 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction, Cold Spring Harbor symposia on quantitative biology. Cold Spring Harbor Laboratory Press, pp. 263-273.
Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., Glöckner, F.O., 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic acids research 35, 7188-7196.
Pruitt, K.D., Tatusova, T., Maglott, D.R., 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic acids research 33, D501-D504.
Ratnasingham, S., Hebert, P.D., 2007. BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). Molecular ecology notes 7, 355-364.
Reese, A.T., Kartzinel, T.R., Petrone, B.L., Turnbaugh, P.J., Pringle, R.M., David, L.A., 2019. Using DNA Metabarcoding To Evaluate the Plant Component of Human Diets: a Proof of Concept. MSystems 4, e00458-00419.
Rodwell, J.S., 1992. British Plant Communities, Volume 3: Grasslands and montane Communities. Cambridge University Press., Cambridge.
Sanger, F., Donelson, J., Coulson, A., Kössel, H., Fischer, D., 1973. Use of DNA polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in phage f1 DNA. Proceedings of the National Academy of Sciences 70, 1209-1213.
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.
Sickel, W., Ankenbrand, M.J., Grimmer, G., Holzschuh, A., Härtel, S., Lanzen, J., Steffan-Dewenter, I., Keller, A., 2015. Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach. BMC ecology 15, 1-9.
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.
Tedersoo, L., Nilsson, R.H., Abarenkov, K., Jairus, T., Sadam, A., Saar, I., Bahram, M., Bechem, E., Chuyong, G., Koljalg, U., 2010. 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. New Phytologist 188, 291-301.
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.
White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications 18, 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.
Yoccoz, N.G., Bråthen, K.A., Gielly, L., Haile, J., Edwards, M.E., Goslar, T., von Stedingk, H., Brysting, A., Coissac, E., Pompanon, F., 2012. DNA from soil mirrors plant taxonomic and growth form diversity. Molecular Ecology 21, 3647-3655.
Zhang, W., Wendel, J.F., Clark, L.G., 1997. Bamboozled again! Inadvertent isolation of fungal rDNA sequences from bamboos (Poaceae: Bambusoideae). Molecular Phylogenetics and Evolution 8, 205-217.