# Peering into the hidden world:

# Photosynthetic microbes as functionally-significant indicators of peatland degradation and restoration.

Alexandra Eleanor Burkitt

PhD

University of York

Environment and Geography

January 2024

# Abstract

Peatlands are increasingly being recognised for their importance as global carbon stores, in their role managing water, and as sites of biodiversity value. Microbes are key drivers of peatland functioning, but most research has focused on microbes living within the peat column, which are involved in nutrient cycling and carbon dynamics. However, peatlands, especially ombrotrophic bogs, are also home to large numbers and diverse communities of photosynthetic and mixotrophic microorganisms. Whilst the responses of certain key groups have been previously studied, there is little prior research on the photosynthetic microbial community on peatland surfaces as a whole.

Peatlands in the UK face varied anthropogenic disturbances, such as drainage, nutrient pollution and increased fire risk due to anthropogenic climate change. The responses of the photosynthetic microbial community to these disturbances are not well understood. Until their response patterns have been identified, it is difficult to predict what impacts changes in algal functioning from disturbances (including algal carbon and nitrogen fixation) may have on peatlands. Furthermore, if changes in microalgae community are consistent and measurable, they may have potential for usage as bioindicators. This could allow for better monitoring of peatlands for 'health status', as well as guiding restoration work. Using light microscopy and molecular methods, this thesis investigates changes in the photosynthetic microbial community with short- and long-term environmental perturbations and successional processes, and looks for evidence of shifts in community and functional changes.

The peatland photosynthetic microbial community appears to be resistant to the three main stressors included in this thesis: fire, nutrient addition and the presence of trees. Instead, microtopography, *Sphagnum* spp. cover and moisture appear to be the most important drivers in algal community at this scale. This suggests that, where conditions allow their *Sphagnum* hosts to persist, algae will also be present and maintain their diversity.

# Author’s Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references.

# Acknowledgements

I would like to start by thanking my supervisors, past and present. These are, in no particular order, Rob Mills, Roxane Andersen, Mark Hancock, Andrew Coupar, Claire Hughes, Thorunn Helgason, James Chong and, of course, Richard Payne. I am especially grateful to Rob Mills for agreeing to take over as my primary supervisor at a very difficult time, and since then being far more patient with me than I deserve. This would not have been possible without all of your support even (especially) when it was hardest to ask for.

Of course, this PhD would not have been possible without Richard Payne, who I greatly miss. I don't think I'll ever be as good a scientist as Richard was, but I hope to one day be as good a colleague and friend. I feel so grateful to have had the opportunity to meet and work with Richard at the start of this PhD. I only wish it could have been for longer.

Thank you to NERC (via the ACCE doctoral training programme) and Nature Scot for providing the funding for this PhD. Thank you to RSPB Scotland for supervision, access and support.

This work would not have been possible without fieldwork access and assistance. Starting in chronological order, thanks to Yorkshire Wildlife Trust for access to Askham Bog for preliminary work.

Thank you to the staff and volunteers at RSPB Forsinard for access and assistance with fieldwork and accommodation. Thanks especially to Henk-Pieter Sterk, Roxane Andersen, Mark Hancock, Juulia and Katey for field assistance.

Thanks to everyone at CEH Edinburgh and the Whim Bog site for allowing me access, especially during the pandemic when field access was difficult. Thanks also to Luke Andrews for agreeing to do the trip in a single day and driving to and from Edinburgh as well as assisting with sampling.

Thank you to the RSPB team at Abernethy for access, and to Rob Mills, Josh Thurston and Ash Taylor for excellent field assistance and help with travel.

Thanks to Josh Thurston again and Hannah Walker for assistance and guidance in molecular laboratory work, and for allowing me to keep using the lab space in Biology. Huge thanks also to the Environment and Geography technicians, for help with everything from fieldwork preparations to making up reagents. Additional thanks to Dave Hay for coming to the rescue when computer problems struck in the final stretch.

I am indebted to Will Burn for his assistance with the initial bioinformatics process and the associated data processing and analysis. The revisions to this thesis would not have been possible without Josh Thurston and his assistance and willingness to share ideas and support with Kraken, as well as being an intermediary between me and the research computing system.

Thanks to Annabel Cansdale for assistance with ImageJ, additional proofreading, chocolate and motivational memes.

Thank you to Andreas Heinemeyer and Sylvia Toet for their support on my thesis advisory panel and progression committee. Thanks also to my examiners, Nathalie Fenner and Mark Hodson, for their invaluable suggestions which improved this thesis significantly.

Thank you to Will Barrop for introducing me to the Flow Country and giving me my first taste of peatland fieldwork.

Thank you to all of the Environment and Geography C-CREW (or PEGRG) research group and the Functional Ecology lab for interesting and inspiring conversations and support.

A huge thanks to my friends in the Environment and Geography and ACCE PhD student community, and outside it, for support, encouragement, cheering up and cat videos. There are too many of you to name here but I hope you know who you are.

Finally, to my family and friends, Pete, and especially my Mum and Dad, without whose love and support none of this would have been possible. Thank you for looking after me and looking out for me, cheering me up and tolerating me talking far too much about moss.

*“It was an interesting ditch. There were jiggling corkscrew things in it which were direct descendants of things which had been in the primordial soup of creation. Anyone who thought that ditchwater was dull could have spent an instructive half-hour in that ditch with a powerful microscope.”*

Terry Pratchett, *Wyrd Sisters*

# Table of Contents

Abstract ..... 2

Author's Declaration ..... 3

Acknowledgements ..... 4

List of figures ..... 12

List of tables ..... 18

1 Literature review ..... 19

1.1 What are peatlands? ..... 19

1.1.1 Peatland microtopography ..... 22

1.2 Importance of peatlands ..... 25

1.3 Threats to peatlands ..... 26

1.3.1 Pollution ..... 26

1.3.2 Climate change ..... 27

1.3.3 Fire ..... 27

1.4 Peatlands in the UK ..... 33

1.4.1 Extent and condition ..... 33

1.4.2 UK peatlands and carbon ..... 35

1.4.3 Peatlands in policy ..... 35

1.4.4 Restoration ..... 37

1.5 Photosynthetic microbes in peatlands ..... 38

1.5.1 Key taxa ..... 41

1.6 Photosynthetic microbes and methodology ..... 45

1.7 Environmental variables and changes in microbial communities ..... 45

1.8 Functional roles ..... 48

1.9 Links to thesis ..... 54

2 Photosynthetic microbial community changes following fire ..... 56
2.1 Introduction ..... 56
2.1.1 Aims and hypotheses ..... 57
2.2 Methods ..... 59
2.2.1 Site description ..... 59
2.2.2 Field methods ..... 60
2.2.3 Peatland composite burn index ..... 62
2.2.4 Microscopy ..... 67
2.2.5 Data analysis ..... 69
2.2.6 Microbial community data ..... 69
2.2.7 Indicator species analysis ..... 71
2.3 Results ..... 71
2.3.1 Environmental data ..... 71
2.3.2 Diversity ..... 73
2.3.3 Microbial communities ..... 74
2.3.4 Indicator species analysis ..... 75
2.4 Discussion ..... 77
2.4.1 Importance of microtopography for microbes ..... 78
2.4.2 Microtopographic and hydrological variation in fire ..... 79
2.4.3 Indicator species ..... 79
2.5 Conclusions ..... 80
3 Bioinformatics method development ..... 82
3.1 Introduction ..... 82
3.2 Laboratory work ..... 83
3.2.1 Primer choice ..... 83

3.2.2 PCR reactions ..... 84
3.2.3 Sequencing ..... 85
3.3 Original bioinformatics method ..... 85
3.3.1 Reference database preparation ..... 86
3.3.2 Data processing ..... 88
3.3.3 Level of taxonomic identification ..... 88
3.4 Benefits and issues with a phylogenetics approach ..... 89
3.5 New bioinformatics method- Kraken2 ..... 90
3.5.1 Database ..... 91
3.5.2 Methodological testing ..... 91
3.6 In silico PCR ..... 103
3.6.1 Introduction ..... 103
3.6.2 Methods and results ..... 103
3.6.3 Implications ..... 104
3.6.4 Prokaryotes ..... 105
4 Photosynthetic microbial communities and nutrient loading – data using a long term experiment ..... 106
4.1 Introduction ..... 106
4.2 Methods ..... 108
4.2.1 Site description and experimental treatments ..... 108
4.2.2 Field sampling ..... 111
4.2.3 Laboratory work ..... 114
4.2.4 Data analysis ..... 118
4.3 Results ..... 125
4.3.1 Algal community changes ..... 125

4.3.2 Block effect ..... 128
4.3.3 Spatial relationships and plot - non plot differences ..... 129
4.3.4 Taxonomic groups ..... 130
4.3.5 Prokaryote to eukaryote ratio ..... 135
4.3.6 Biodiversity indices ..... 141
4.3.7 Plant community changes – Kraken2 bioinformatics ..... 142
4.4 Discussion ..... 152
4.4.1 Algal community ..... 152
4.4.2 Non-plot effect ..... 155
4.4.3 Plant community changes ..... 156
4.5 Conclusions ..... 157
5 The effect of tree cover on photosynthetic microbes in a near natural UK bog woodland system ..... 159
5.1 Introduction ..... 159
5.2 Methods ..... 161
5.2.1 Site description ..... 161
5.2.2 Field methods ..... 163
5.2.3 Sample storage ..... 166
5.2.4 Laboratory analysis ..... 166
5.2.5 Data analysis ..... 168
5.3 Results ..... 187
5.3.1 Trees and canopy cover ..... 187
5.3.2 Environmental variables ..... 192
5.3.3 Transect effects ..... 197
5.3.4 Community composition ..... 199

5.3.5 Correlations between variables ..... 216
5.3.6 Distance based RDA ..... 219
5.3.7 Mantel tests ..... 222
5.4 Discussion ..... 222
5.4.1 Changes in the algal community ..... 222
5.4.2 Vegetation ..... 230
5.5 Conclusions ..... 232
6 Conclusions ..... 234
6.1 Methodology ..... 234
6.2 Algal community responses to stressors ..... 234
6.3 Management ..... 235
7 Supplementary data ..... 237
8 References ..... 244

# List of figures

**1.1.** Part of a permanent peatland pool system at RSPB Forsinard Flows NNR.....23

**1.2.** Distribution of peatland land cover types in the UK.....34

**2.1.** Location of Chapter 2 study site.....59

**2.2.** Location of sampling points and margin between burnt and unburnt areas.....61

**2.3.** Schematic of Chapter 2 field sampling methodology.....62

**2.4.** Sample B1la, with a peatland composite burn index score of 0.5.....65

**2.5.** Sample B2hk, with a peatland composite burn index score of 2.17.....66

**2.6.** Sample B5hk, with a peatland composite burn index score of 2.....67

**2.7.** NMDS stress plot for Chapter 2 microbial community.....70

**2.8.** Box plot of mean soil moisture of different microtopographic landforms.....72

**2.9.** Box plot showing the effect of microtopography on fire damage.....72

**2.10.** Fire severity plotted against mean soil moisture content.....73

**2.11.** Boxplot showing difference in inverse Simpson's biodiversity with fire.....74

**2.12.** NMDS plot showing microbial community from Chapter 2 by microtopography and fire status.....75

**3.1.** Flowchart showing custom reference database construction for analysis of rbcLa amplicons with DADA2.....87

**3.2.** Relative abundance of dominant algae phyla with no filtering for ASVs unclassified at the phylum level (first bioinformatics method).....88

**3.3.** Relative abundance of dominant microbial classes after filtering to remove ASVs unclassified to phylum level (first bioinformatics method).....89

**3.4.** Sankey visualisation of the taxonomy assigned by Kraken (Kraken test 2).....93

**3.5.** Sankey visualisation of the taxonomy assigned by Kraken (Kraken test 3).....95

**3.6.** Sankey visualisation of the taxonomy assigned by Kraken (Kraken test 4).....97

**3.7.** Sankey visualisation of the taxonomy assigned by Kraken (final Kraken run).....99

**3.8.** Percentage of unclassified reads from the final Kraken2 run for all samples.....102

**4.1.** Map showing the location of Whim Bog (Chapter 4 study site).....109

**4.2.** Diagram of the CEH Whim Bog wet deposition experiment.....110

**4.3.** Satellite imagery showing the location of experimental treatment plots and non-plot controls used in Chapter 4.....112

**4.4.** Diagram showing the setup for pore water extraction from *Sphagnum* capitula using a centrifuge.....115

**4.5.** *Sphagnum* pore water nutrient concentrations with different nutrient addition treatments.....116

**4.6.** Chart showing trace from AutoAnalyser3 channel for phosphate.....117

**4.7.** *Sphagnum* pore water nutrients compared to real-world detection limits.....118

**4.8.** Histogram of total read counts for all samples from Chapter 4.....121

**4.9.** Rarefaction curves showing biodiversity by sampling effort for Chapter 4.....122

**4.10.** Correlation plot between observed taxonomic richness and read count (a proxy for sequencing depth) for Chapter 4 data.....123

**4.11.** Location of incorrect coordinate for sample WNP2 (Chapter 4).....125

**4.12.** PCoA plot showing the difference in algal communities between treatments (first bioinformatics method).....126

**4.13.** DCA plot of algal community by treatment and algae classes present (second bioinformatics method).....127

**4.14.** PCoA plot of beta dispersal of algae communities by treatment (second bioinformatics method).....128

**4.15.** PCoA plot showing beta dispersal of algal community by plot/ non plot status (second bioinformatics method).....130

**4.16.** Relative abundance of dominant algal phyla without filtering to remove ASVs unclassified at the phylum level (first bioinformatics method).....131

**4.17.** Relative abundance of dominant microbial phyla after filtering to remove ASVs unclassified to phylum level (first bioinformatics method).....131

**4.18.** Relative abundance of dominant classes after filtering to remove ASVs unclassified to phylum level (first bioinformatics method).....132

**4.19.** Relevant abundance of microbial phyla by nutrient addition treatment (second bioinformatics method).....133

**4.20.** Relative abundance of the algal community by nutrient addition treatment agglomerated at the class level (second bioinformatics method).....134

**4.21.** Relative abundances of the most common orders in the microbial community by nutrient addition treatment (second bioinformatics method).....135

**4.22.** Relative abundance of eukaryote vs. prokaryote ASVs, without filtering to remove ASVs unassigned at any taxonomic level (first bioinformatics method).....136

**4.23.** Relative abundance of eukaryote vs. prokaryote ASVs after removal of ASVs not classified to phylum level (first bioinformatics method).....136

**4.24.** Proportion of prokaryote ASVs of those assigned to empire level (first bioinformatics method).....138

**4.25.** Violin plot showing proportion of microbial reads assigned to prokaryotes by nutrient addition treatment (second bioinformatics method).....139

**4.26.** Relative abundance of microbial reads assigned to Eukaryota and Prokaryota by nutrient treatment (second bioinformatics method).....140

**4.27.** Composition plot of relative abundance of microbial reads agglomerated at the phylum level (second bioinformatics method).....141

**4.28.** Inverse Simpson's biodiversity index by nutrient addition treatment and non-plot controls (first bioinformatics method).....142

**4.29.** Relative abundance of the classes of vascular and non-vascular plants (second bioinformatics method).....144

**4.30.** PCoA plot showing combined plant community beta dispersal by different nutrient addition treatments (second bioinformatics method).....145

**4.31.** Relative abundane of the classes of non-vascular plants by nutrient addition treatment (second bioinformatics method).....146

**4.32.** PCoA plot of nonvascular plant beta dispersal by nutrient addition treatment...147

**4.33.** DCA plot of bryophyte community by treatment.....148

**4.34.** DCA plot of bryophyte community by nutrient addition treatment, zoomed to clearly show plot points.....149

**4.35.** Relative abundance of the classes of vascular plants by nutrient addition treatment.....150

**4.36.** PCoA plot of vascular plant community beta dispersal by nutrient addition treatment.....151

**4.37.** DCA plot of vascular plant community by treatment.....152

**5.1.** The location of Abernethy Forest (Chapter 5 study site).....162

**5.2.** The location of sampling points for Chapter 5 within Abernethy Forest.....163

**5.3.** An illustration of the tree health grading used in Chapter 5.....165

**5.4.** Diagram showing the microbial sample filtering process for Chapter 5.....167

**5.5.** Images showing the setup for hemispherical lens calibration.....169

**5.6.** Merged canopy cover images created using Hugin.....170

**5.7.** Default thresholding of an image with high canopy cover in Fiji.....172

**5.8.** Default thresholding of an image with lower canopy cover in Fiji.....173

**5.9.** An image with low canopy cover thresholded in 'minimum' mode and 'intermodes' mode in Fiji.....174

**5.10.** An image with canopy cover of zero in the original colour and thresholded at different levels.....176

**5.11.** Histogram of total read counts for all samples in Chapter 5.....178

**5.12.** Rarefaction curves showing biodiversity by sampling effort for Chapter 5.....179

**5.13.** Correlation plot between observed taxonomic richness and read count (a proxy for sequencing depth) for Chapter 5 data.....180

**5.14.** Density plot of relative abundance of prokaryote reads (Chapter 5).....182

**5.15.** Frequency of different numbers of tree stems within 2m of the sampling point..188

**5.16.** Number of trees of each health category by plot.....189

**5.17.** Histogram of tree heights sampled in Chapter 5.....190

**5.18.** Histogram of DBH measurements for trees surveyed in Chapter 5.....191

**5.19.** Histogram of calculated percentage canopy cover for plots in Chapter 5.....192

**5.20.** Histogram of bryophyte percentage water for samples in Chapter 5.....193

**5.21.** Histogram of soil moisture probe measurements for samples in Chapter 5.....194

**5.22.** *Sphagnum* pore water nutrient concentrations from Chapter 5 with real-world detection limits.....195

**5.23.** Long-term soil moisture measurements for Abernethy Forest.....196

**5.24.** Long-term surface temperature measurements for Abernethy Forest.....197

**5.25.** Relative abundance of prokaryotes between transects for Chapter 5.....198

**5.26.** Variation in number of tree stems within 2m of sampling point by transect.....199

**5.27.** Relative abundance of the most abundant microbial classes without filtering to remove unclassified ASVs (first bioinformatics method).....200

**5.28.** Relative abundance of the most abundant microbial phyla without filtering to remove unclassified ASVs (first bioinformatics method).....201

**5.29.** Relative abundance of the most dominant microbial phyla after removal of ASVs not identified to phylum level (first bioinformatics method).....202

**5.30.** Relative abundance of the most dominant microbial classes after removal of ASVs not identified to phylum level (first bioinformatics method).....203

**5.31.** Histogram showing frequency of relative abundance of *Sphagnum* spp. reads between Chapter 5 samples (second bioinformatics method).....204

**5.32.** Composition plot showing the relative abundances of Eukarya and Prokarya for Chapter 5 samples (second bioinformatics method).....205

**5.33.** Relative abundance of algal phyla for Chapter 5 samples (second bioinformatics method).....206

**5.34.** Composition plot of algal classes from Chapter 5 samples (second bioinformatics method).....206

**5.35.** Composition plot of algal orders from Chapter 5 samples (second bioinformatics method).....207

**5.36.** Proportional composition plot of all microbial reads at the phylum level (second bioinformatics method).....208

**5.37.** Proportional composition plot of all microbes at the class level (second bioinformatics method).....209

**5.38.** Proportional composition plot of all microbes at the order level (second bioinformatics method).....210

**5.39.** Proportional composition plot of all plants at the class level (second bioinformatics method).....211

**5.40.** Proportional composition plot of all plants at the order level (second bioinformatics method).....212

**5.41.** Proportional composition plot of bryophytes at the order level (second bioinformatics method).....213

**5.42.** Proportional composition plot of vascular plants agglomerated at the class level (second bioinformatics method).....214

**5.43.** Proportional composition plot of vascular plants at the order level (second bioinformatics method).....215

**5.44.** Correlation plot between all numeric variables from Chapter 5.....216

**5.45.** *Sphagnum* percentage moisture by weight for different dominant *Sphagnum* species.....219

**5.46.** Biplot of dbRDA model 3.....220

**5.47.** Biplot for dbRDA model 6.....221

# List of tables

**Table 1.1.** Definitions of peatland microtopographic landforms as used in this thesis and example images.....24

**Table 1.2.** Estimated peatland land cover in the UK and constituent nations.....33

**Table 2.1.** Details of peatland composite burn index scoring, as used in the field to quantify fire damage to *Sphagnum*.....64

**Table 2.2.** Hummock microtopography indicator species.....76

**Table 2.3.** Hollow microtopography indicator species.....76

**Table 3.1.** Primer sequences used in PCR reactions.....83

**Table 3.2.** PCR cycle used for rbcL primers.....84

**Table 3.3.** Percentage of reads unclassified and assigned to different taxonomic groups from Kraken2 methodological testing runs.....100

**Table 3.4.** Taxonomic outputs for the Sphagnaceae from Kraken2 methodological testing runs.....101

**Table 4.1.** Total read counts for the six samples with the lowest number of reads from Chapter 4.....121

**Table 4.2.** Percentage of reads assigned to the genus *Sphagnum* and species within it by sample (data from Chapter 4).....143

**Table 5.1.** Total read counts for the six samples with the lowest number of reads from Chapter 5 .....177

**Table 5.2**. Variance inflation factors for numerical variables for Chapter 5.....218

**Supplementary Table 3.1.** Prokaryote taxa found by *in silico* PCR.....237

**Supplementary Table 5.1.** Thresholding values used during canopy cover calculation for images with visible trees.....238

**Supplementary Table 5.2:** VIFs calculated for dbRDA models.....241

# 1 Literature review

## 1.1 What are peatlands?

Peatlands are areas where the rate of deposition of organic material from net primary productivity (NPP) outstrips the rate of decay. This leads to an accumulation of organic material, which due to environmental conditions including low pH, low nutrient levels and anoxic conditions due to waterlogging, is slow to decay (Clymo, Turunen and Tolonen, 1998). In this way, organic material can be stored for many thousands of years. Carbon has been estimated to accumulate at an average of 18.6 g C $m^{-2}$ $yr^{-1}$ over the course of the Holocene in northern peatlands (Yu *et al*., 2010).

Strict definitions of peatlands vary, with the depth of peat and status of the peatland (how degraded or active it is) amongst the deciding variables (Joint Nature Conservation Committee, 2011). In general, areas with greater than 30-40cm depth of peat are considered to be peatlands (Gorham, 1991). Mire is a similar term used most often in ecological and botanical contexts to describe wetland areas dominated by peat forming vegetation (Rydin and Jeglum, 2006). The two terms are not interchangeable as a mire without peat of an adequate depth would not be classified as a peatland and a peatland without active peat forming vegetation would no longer be a mire. Peat forming vegetation is also difficult to define- while *Sphagnum* mosses are widely described as the most important plants for active peat formation in the policy literature, widely variable plant materials can go on to form peat if conditions are suitable (Rydin and Jeglum, 2006). Globally, it is estimated that peatlands account for 3% of land cover, yet they act as a major global carbon store and have an important role in carbon cycling (Roulet *et al*., 2007). Gorham (1991) estimated that peatlands in the Boreal and Subarctic regions of the northern hemisphere alone account for a third of global soil carbon stocks, totalling 455Gt of carbon. More recent studies have led to even higher estimates; Yu *et al*. (2010) estimated a global peatland carbon store of over 600Gt of carbon, of which 547Gt were found in northern peatlands. Peatlands can also be found in the tropics, with peat swamps an exclusively tropical type of peatland (Posa, Wijedasa and Corlett, 2011).

Terms and definitions for specific types of peatlands vary between scientific disciplines, regions and languages (Rydin and Jeglum, 2006). However, at the coarsest level peatlands can be divided into fens and bogs. Fens are generally considered to be peatlands where water is supplied from overland flow, waterbodies or groundwater. This provides dissolved nutrients, and so they tend to have higher nutrient levels and therefore higher net primary productivity. Bogs, conversely, obtain all of their water from rainfall and atmospheric water, and so tend to be more ombrotrophic (that is, obtaining external nutrients via precipitation and airborne dust). This results in lower nutrient availability and more strongly oligotrophic conditions. The higher mineral level in fens also results in a lower proportion of organic material in fens than in bogs (Rydin and Jeglum, 2006).

In the UK, bogs are broadly divided into raised bogs, the surface of which is raised above the level of the surrounding ground surface, and blanket bogs, which cover the terrain and can include low hills. Blanket bogs tend to be found in humid, oceanic climates (Rydin and Jeglum, 2006). UK bogs are generally considered to be treeless when in good condition, and forestry policy does not support tree planting on deep peat (Forestry Commission and Natural England, 2023; Forestry Commission Scotland, 2015).

Differing nutrient levels lead to differences in the plant and microbial communities between different types of peatland, and different plant communities can result in peat formation as long as net primary productivity (NPP) exceeds the rate of decay (Joint Nature Conservation Committee, 2011). In particular, *Sphagnum* mosses are a key plant in the formation of peat, especially in ombrotrophic conditions (Gorham, 1991). *Sphagnum* is an ecosystem engineer, altering the environment around it. Hollow hyaline cells, which are dead at maturity, soak up water, raising the water table, while high ion exchange capacity lowers the surrounding pH (Bengtsson *et al*., 2020; Clymo, 1963). These combine to make the local environment more suitable for *Sphagnum* and other adapted peatland specialists, while suppressing the growth of other plants which could otherwise outcompete them. The oligotrophic conditions also allow these relatively slow growing plants to remain dominant; when peatlands are polluted with nitrogen,

*Sphagnum* can become less abundant and show reduced growth (Berendse *et al*., 2001).

As well as storing water intracellularly (Bengtsson *et al*., 2020), *Sphagnum* can also store water extracellularly; their long, feathery branches are likely an adaptation for this (van de Koot *et al*., 2024). Both intracellular and extracellular water storage capacity differs between *Sphagnum* species, and extracellular capacity varies with *Sphagnum* colony density (van de Koot *et al*., 2024; Bengtsson *et al*., 2020). Vegetation changes at the species level can therefore affect ecosystem processes at smaller scales.

The microbial community varies by depth in peatland soils, as different layers provide different ecological niches (Andersen, Chapman and Artz, 2013). The surface acrotelm (the active layer, above the lowest water table level) is predominantly oxic, while the catotelm (considered the 'inactive layer', below the water table) is anoxic, and the intermediate mesotelm can experience both oxic and anoxic conditions (Clymo and Bryant, 2008; Ingram, 1978). Note that some studies prefer a diplotelmic model, with just the acrotelm and catotelm (Rydin and Jeglum, 2006), and others consider the acrotelm to be subdivided into a surface oxytelm (the primarily oxic surface layer) and the mesotelm below it (usually anoxic but occasionally oxic), with mean water table as the boundary between the two (Clymo and Williams, 2012). However, regardless of the model used, photosynthetic microbes would be expected to inhabit a narrower depth niche than described in these models, with only the surface subsection of the acrotelm where light is able to penetrate suitable for photosynthesis to take place (Geisen *et al*., 2018). Mitchell and Gilbert (2004) found that the mixotrophic testate amoeba *Hyalosphenia papilio* was found mostly in the top 1cm of *Sphagnum*, though only the top 5cm were tested and differences were not significant due to patchiness of occurrence. Mixotrophic testate amoebae able to colonise moss surfaces may acquire more of an advantage from symbiotic zoochlorellae due to increased exposure to photosynthetically active radiation (PAR) than if they were living deeper within the bryophyte layer, even if still within the photosynthetic zone (Mitchell and Gilbert, 2004). While mixotrophic taxa may be facultative rather than obligate autotrophs, and so may be found deeper within peat, they would not be able to perform photosynthesis

below the level at which light can penetrate. Protists are capable of moving through the soil column at a rate of a few cm per day in the sandy soils where their motility is most rapid (Adl, 2007) and presumably less in dense organic soils such as peat, so any photosynthetic protists are unlikely to be able to move to deeper into the peat profile due to nutritional limitation. While some photosynthetic protists are capable of surviving in deeper soil layers without light if sufficient nutrients are available, they are by definition not photosynthesising while doing so, and survive either by absorbing nutrients dissolved in water or through specialised storage organs (Geisen *et al.*, 2018). It is possible that normally photosynthetic taxa surviving off carbon in storage organs while below the reach of light are another source of labile carbon deeper into the peat profile, although this does not seem to have been investigated (and these organisms may not even be present in peatlands).

### 1.1.1 Peatland microtopography

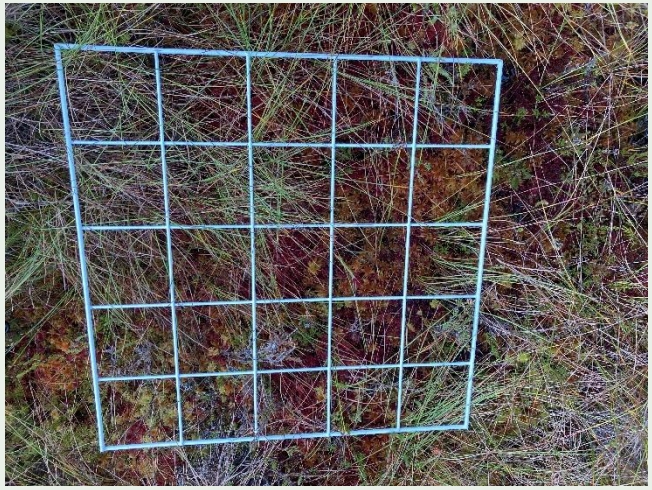
Peatlands show extensive microtopographic variation, often forming a mosaic of hummocks, hollows and lawns. These hummocks often have distinctive vegetation due to environmental differences (especially moisture levels), and are themselves likely formed due to vegetation dynamics (Pouliot *et al.*, 2011). There is no fixed definition for these microforms based on height; the water table in peatlands varies in time, and so the height of any particular microtopographic feature relative to the water table, or if the surface vegetation is saturated or not, may vary (Rydin and Jeglum, 2006). Peatland bacterial communities vary with microtopography within the same peatland, probably reflecting differences in moisture, nutrient levels and vegetation communities (Goodyer, 2014; Andersen *et al.*, 2013). See Table 1.1 for definitions of peatland microtopographies used in this thesis.

Alongside hollows, peatlands also may have more permanent pool systems (Fig. 1.1). These can also have high algal abundance levels (Kreutz and Foissner, 2006), and provide more permanent aquatic conditions for algal species more sensitive to desiccation.



**Figure 1.1.** Part of a permanent peatland pool system at RSPB Forsinard Flows NNR (Chapter 2 field site).

**Table 1.1.** Definitions of peatland microtopographic landforms as used in this thesis and example images from RSPB Forsinard Flows NNR (Chapter 2 field site).

| Hummock | Lawn | Hollow |
|---|---|---|
| A raised mound above the general bog vegetation surface | A flat (or reasonably flat) level of vegetation lower than hummocks, but above hollows. Not continuously below water level. | A depression in the peatland below the level of the lawn. Sometimes small transitional pools. |
|  |  |  |

## 1.2 Importance of peatlands

Peatlands have attracted increased levels of interest in recent years as the important ecosystem services which they provide have become better recognised. While peatlands provide a wide range of ecosystem services, carbon storage has attracted the greatest attention for policy decisions.

Healthy peatland stores carbon which is fixed from the atmosphere by photosynthesis. Where peat is still growing, new carbon is fixed by photosynthetic organisms and sequestered into the newly forming peat, while historically fixed carbon remains trapped. This occurs when there is an imbalance between the input of carbon from photosynthesis, via deposition of material from photosynthetic organisms, and the loss of carbon to the environment via microbial decay (Clymo *et al.*, 1998). However, when peat becomes degraded and the conditions which slow decay are altered, deposition of carbon may no longer exceed decomposition and peatlands can become a net carbon source. Indeed, much of the peatlands of the east Anglian fens have been so degraded by centuries of drainage and farming that much of the carbon content has been lost (Holman, 2009; Richardson and Smith, 1977).

As well as storing carbon, peatlands also manage water flow and quality. Peatlands are estimated to store 10% of global freshwater (Joosten and Clarke, 2002 via González Garraza, Burdman, and Mataloni, 2019). Filtering through peatlands can also improve water quality of streams and rivers which flow through peat. Conversely, where peatlands are degraded, water can be of lower quality because of high concentrations of dissolved organic carbon (Xu, 2018; Schwalm and Zeitz, 2015). The ability of peatlands to slow water loss and reduce peak flows has been identified by the IUCN (2018) as one of the key ways in which UK peatlands can assist with climate change adaptation. While the generally high water-table of healthy peatlands means that there can be limited capacity for upland peatlands to store additional water, when drainage or erosion channels are present, speed of flow is likely to be higher and the system 'flashier'. Therefore, blocking drainage channels and preventing erosion should reduce peak flows from peatlands (IUCN, 2018; Wilson *et al.*, 2011).

Peatlands in the UK are also a cultural landscape (IUCN, 2018). Peatlands are often iconic landscapes, and dominate all but two UK national parks (National Parks England, 2019; IUCN, 2018). Alongside their use for recreation, peatlands can have important roles in tourism and for local communities. The challenge is therefore balancing this cultural value with managing peatlands for carbon and biodiversity.

## 1.3 Threats to peatlands

Many peatlands globally are degraded. In the UK, which is the focus for this thesis, only 22% (Evans *et al.*, 2017) of peatlands are estimated to be in near natural conditions. This includes peatlands which have been restored and rewetted following historic damage. An estimated 80% of UK peatlands have been affected by historical management practices (IUCN, 2018).

Centuries of cutting of peat for fuel and commercial peat extraction for horticulture have depleted stores of peat. Drainage to convert peatlands to farmland or commercial forestry has resulted in the lowering of the water table. This contributes to peatland degradation both directly and indirectly, by allowing the rate of decomposition to increase, and by allowing plants which grow better under drier conditions to colonise the peatland, changing the biotic community (IUCN, 2018; Lindsay, Birnie and Clough, 2014).

### 1.3.1 Pollution

Peatlands are also threatened by anthropogenic pollution. Low nutrient levels lead to high C:N ratios in healthy peatlands and this nutrient limitation reduces microbial decomposition activity (Rydin and Jeglum, 2006). However, many peatlands have high inputs of anthropogenic nutrients. Nitrogen is gained through wet and dry deposition, in both oxidised and reduced forms (Sheppard *et al.*, 2014; Anderson and Downing, 2006). This leads to changes in both the microbial and plant communities, as peatland specialists are outcompeted by faster growing species which can take advantage of the raised nutrient levels (Payne *et al.*, 2013; Berendse *et al.*, 2001). Historic nutrient deposition, especially of N and S, has been identified as a limiting factor for UK peatland recovery (Artz *et al.*, 2019).

### 1.3.2 Climate change

Rates of microbial degradation are strongly linked to climatic variables, including temperature, and it has been predicted that increases in temperature due to anthropogenic climate change could lead to increased loss of carbon stored in peat (Ofiti *et al.*, 2023). Climate change is also predicted to change rainfall patterns. As the location of peatlands seems to be in part driven by rainfall patterns, these changes could result in areas which are currently home to peatlands being outside of the optimal climate envelope in the future. For example, blanket bogs seem to form only when conditions are consistently wet and cool, with a minimum mean annual precipitation above 1000mm seeming to be required (Lindsay *et al.*, 1988), and UK blanket bog distribution correlating with the number of wet days where precipitation exceeds 1mm in 24 hours (Goode and Ratcliffe, 1977 via Lindsay *et al.*, 1988). As the climate changes, the climate envelope within which UK peatlands are found is predicted to shrink, potentially leading to climate stress (House *et al.*, 2010).

*Sphagnum* hydrological stress is also predicted to increase with climate change, with greater stress likely for hummocks than the wetter hollows. Hollows are projected to have similar hydrological regimes to current hummocks under some climate scenarios (Moore and Waddington, 2015). This will likely cause subsequent shifts in the surface microbial communities, especially algae which are known to have a strong relationship with microtopography (Goodyer, 2014).

### 1.3.3 Fire

Fire seasons are predicted to become longer and more severe with climate change (Flannigan *et al.*, 2013). Climate change is already thought to be increasing burning (and therefore soil carbon loss), with deeper burning in at least some systems pushing soils into net carbon loss (Turetsky *et al.*, 2010). In the UK, wildfires on peatlands appear to be increasing in frequency (Artz *et al.*, 2019), and projections suggest that the UK wildfire risk will increase further with climate change (Shotter and Wentworth, 2019).

Peatland fires result in the loss of stored soil carbon if sub-surface peat is consumed (Artz *et al.*, 2019). They also impact the surface vegetation, and therefore greenhouse

gas exchange and ecosystem functioning (Artz *et al*., 2019). As well as releasing carbon, wildfires also reduce carbon uptake in pristine peatlands by 35% (Wilkinson *et al*., 2023).

Longer term impacts of fire on UK peatlands, which may limit the success of restoration, are largely unknown (Artz *et al*., 2019). More fragmented peatlands, which restored sites may be, are at greater risk from fire, in part due to hydrological changes from increased edge effects (Nelson *et al*., 2021). Soil moisture controls both initial peat ignition, and the severity and depth of burn (ibid). Drained peatlands are therefore at greater risk from fires, which may burn deeper into the peat than in undrained counterparts (Nelson *et al*., 2021). Ombrotrophic bogs, isolated from groundwater sources, may be more vulnerable to water table changes, and therefore reduced moss recovery, following fire (Nelson *et al.,* 2021). Conversely, high moisture content protects most peat in pristine peatlands from burning (Turetsky *et al*., 2015). Pristine peatlands are resilient to fire, and tend to return to net carbon sinks (Wilkinson *et al*., 2023). This highlights the importance of peatland restoration to prevent climate feedback cycles from increased fires on degraded and unrestored peatlands.

Drying from climate change and drainage allows deep peat to ignite more easily, increasing the frequency and severity of peatland fires, leading to positive feedback loops (Turetsky *et al*., 2015). Peatland fire feedback loops are varied and extremely complex, with various positive and negative feedback effects (Nelson *et al*., 2021). However, positive feedback effects seem to dominate, and with climate change positive feedback loops are likely to worsen the risk of peatland fires and carbon loss (Nelson *et al*., 2021). Climate change increases carbon losses through increased burn rate and burn severity, reducing the carbon sink capacity of peatlands (Wilkinson *et al*., 2023).

Fires can have varied physico-chemical impacts, which will in turn effect peatland ecology and functioning. Where fires remove the shading and insulating vegetation, they result in an increase in soil temperature (Mieczan, Bronowicka-Mielniczuk and Rudyk-Leuska, 2022; Holden, Berhe and Treseder, 2015). Soil pH generally increases with burn severity (Brown *et al*., 2019; Ferrenberg *et al*., 2013). Across different ecosystem types, soil pH can affect the biodiversity and phylogenetic diversity of bacterial communities

(Lauber *et al*., 2009). Fire can also change nutrient conditions; burning of a forest soil led to reduced C and N levels but increased $NH_4^+$ and DON (dissolved organic nitrogen) (Ferrenberg *et al*., 2013). Increases in biogenic nutrients e.g. $NH_4$ and $NO_3^-$ etc. can depress populations of nutrient sensitive microbes (Meiczan *et al*., 2022).

Fire can also result in reduced oxygen conditions in the soil (Meiczan *et al*., 2022). Testate amoeba are sensitive to low oxygen conditions (Mazei, Tsyganov and Bubnova, 2007), and this is likely the case for many other microeukaryotes, though photosynthetic organisms may be less affected than autotrophs. Low oxygen levels and changes to pH were likely caused by decomposition of plant biomass and the micro- and macro-organisms killed by fire (Mieczan *et al*., 2022).

Fire may also cause increased water levels in peatlands under some circumstances, either via increased bulk density causing surface waterlogging, or removal of local vegetation reducing evapotranspiration rates (Marcisz *et al*., 2015). Clearing of the surrounding landscape can also result in water which had previously been intercepted by vegetation reaching the peatland and contributing to higher water tables (Marcisz *et al.,* 2015).

The interaction between fire and peatland microtopography is variable; with shallow burns, elevation difference between hummocks and hollows can be increased, reinforcing microtopographic variance, whereas with deeper burns, hummocks can be destroyed, becoming hollows, and *Sphagnum* regeneration will be limited (Nelson *et al.,* 2021; Turetsky *et al*., 2015). *Sphagnum* hummocks are often considered more resistant to fire damage than hollows (Nelson *et al*., 2021; Thompson and Waddington, 2013). *Sphagnum* mosses, which store water in their hyaline cells, would be expected to reduce risk of ignition by increasing surface moisture and insulating the deeper peat. Different species of *Sphagnum* in different situations will have different moisture storage capacities; hummock-growing species can increase their hyaline cell volume to increase moisture storage, while hollow-growing species do not (Bu *et al*., 2013). This can result in hollow growing species being more flammable, which has been recorded in Canadian boreal systems (Nelson *et al*., 2021). *Sphagnum* dominated peatlands are likely to be vulnerable to drying following fire, with serious hydrological impacts on the

bryophyte community, reducing recovery, even if burn severity of hummocks does not appear to be extreme (Nelson *et al.,* 2021; Thompson and Waddington, 2013). Peatland microtopography does not improve predictive models of depth of burn, but there are changes in surface burn severity due to microtopography (Kohlenberg *et al*., 2018). Where burns are shallower, the majority of the impact of the fire is on the surface layer of the peatlands, and therefore it makes sense that microtopography can impact burn severity. Where burns are deeper, and go into the deeper peat, microtopography may have less of a protective effect.

Where *Sphagnum* is isolated from the water table, and so cannot take up more water via capillary action, recovery following desiccation can be limited (Schipperges and Rydin, 1998). Therefore, if a hummock becomes hydrologically isolated as a result of fire damaging the lower levels, recovery may not be possible, even if the *Sphagnum* is able to survive the initial fire damage. Different *Sphagnum* species have different resistance to desiccation, notwithstanding relative distance to water table (Nibau *et al.,* 2022). Denser colonies are more resistant to desiccation, and different species have different colony densities (van de Koot *et al*., 2024). The impact of fire on *Sphagnum* mosses may therefore have species specific relationships.

Fire can also reduce the average water level in *Sphagnum* hollows (Mieczan *et al*., 2022), essentially reducing the transient pool microtopographic niche. This effect persisted for two years post-fire, and was typified by extreme change immediately following fire, then slow recovery. These long-term changes in microtopographic niche availability would be expected to have a direct effect on surface communities, especially quasi-freshwater ones such as algae.

The effects of fire on microbial communities are complex and varied (Barreiro and Díaz-Raviña, 2021; Dooley and Treseder, 2012). Most studies of fire and soil ecology focus on prescribed fires and fires within forests, grasslands and shrublands. There is also a taxonomic bias, with most work on the bacterial and fungal community (Pressler, Moore and Francesca Cotrufo, 2019). Fire has been shown to reduce microbial diversity (Mieczan *et al*., 2022). The impact of fire on species richness and evenness is variable by taxa, but most taxonomic groups show significant reductions in diversity following

fire (Pressler *et al*., 2019; Ferrenberg *et al*., 2013). Fire can have strong negative effects on soil microbial biomass, but the negative impacts on biodiversity tend to be stronger (Pressler *et al*., 2019). Conversely, fire can also temporarily increase soil microbial phylogenetic diversity, before it drops down to previous levels as ecosystem recovery allows more competitive taxa to dominate once again (Pérez-Valera *et al.,* 2018).

Whether the same patterns would be reflected by surface microbial communities, which would have greater proximity to heat and more significant loss of substrate than those deeper down the soil column, is uncertain. It seems reasonable to believe that surface and bryosphere microbial communities will be more affected by the immediate damage from fire, e.g. consumption or scorching of bryophyte hosts/ substrate. Wildfires have indeed been shown to have a greater effect on surface microbial communities than those in subsoils (Pressler *et al*., 2019). As well as the direct impacts of fire on microbes, fire can shift peatland vegetation communities (Mieczan *et al*., 2022), which is also likely to result in changes in the surface microbial community. Fire has been shown to affect the testate amoeba community indirectly via changes in hydrology and vegetation, in particular landscape openness (Marcisz *et al*., 2014). Mieczan *et al*. (2022) showed persistent changes in algal abundance following a fire, as measured by chlorophyll *a* concentration, which had not recovered in 12 months. However, whether fire also caused structural shifts in the algal community is unknown.

The period immediately following fire favours taxa which can survive the high heat and take advantage of post-fire nutrient release (Pérez-Valera, Goberna and Verdú, 2019). Cyanobacteria, being able to form akinetes (dormant cells), can cope with the high temperatures of fire and proliferate in the aftermath, and may take advantage of the burst in mineral N which often follows fires (Pérez-Valera *et al*., 2019). Post fire changes in microbial abundance and diversity can affect ecosystem function, for example a reduction in fungal litter decomposition for several years following a fire (Holden, Gutierrez and Treseder, 2013).

Fire severity can change soil bacterial communities in forest soils, with communities subjected to more severe fires being more different to unburnt controls than communities subjected to moderate fires (Brown *et al*., 2019). Microbial response to fire

(biomass and soil respiration rate) is significantly correlated with burn intensity, with greater burn intensity causing more severe effects (Holden *et al.*, 2016). Burn severity has significant but low predictive power of soil microbial community (bacteria and fungi) (Whitman *et al.*, 2019).

Fire can shift soil microbial communities to be more similar than in unburnt soils, causing a loss in β diversity (Weber *et al.,* 2014). In Canadian boreal systems, wildfires cause soil bacterial and fungal communities to become more dissimilar to unburnt controls in upland sites but not wetland sites; for wetland sites in the same study understory vegetation community had a greater relationship with microbial communities (Whitman *et al.,* 2019).

Meta-analysis suggests soil microbial communities are not resistant to fire; this is unsurprising when the major initial impact is death of microbes from heat (Pressler *et al.,* 2019). Soil microbial resilience also appears to be limited, based on poor recovery over time (Pressler *et al.*, 2019). Peatlands, where wetter environments may provide some protection and a refugia for subsequent recovery, may not follow this pattern. In mosaic boreal forest systems, wetlands tend to burn less severely and so may play a role in aiding re-colonisation of microbes into nearby more severely burned areas (Whitman *et al.*, 2019).

Fire responsive bacteria and fungi have been identified in previous studies (e.g. Whitman *et al.*, 2019). Fire and its aftereffects can be clearly observed, and so bioindicators of fire are not needed. However, bioindicators of recovery may be useful to assess the trajectory that the community is on and if it is recovering. If photosynthetic microbes are affected by fire, as is probable when they rely on water conditions in the bryosphere, they may be valuable candidates for this.

## 1.4 Peatlands in the UK

### 1.4.1 Extent and condition

Estimates of UK peatland cover vary between around 3.0 million hectares (12.2 % of the total UK land area) (Evans *et al.*, 2017) and 2.6 million hectares (10% of the UK's land area) (IUCN, 2018). They are "the largest component of the UK's wetland environment" (IUCN, 2018). Table 1.2 shows how this is divided between the different constituent nations of the UK.

**Table 1.2.** Estimated peatland land cover in the UK and constituent nations, adapted from Artz et al. (2019). Peat areas reported by the Joint Nature Conservation Committee (JNCC, 2011) and updated area estimates based on mapping from the UK BEIS Inventory project (Evans et al., 2017). These figures include only true peats (peaty soils are excluded). Only England has separate data for deep and wasted peats.

| Country/administration | 2011(ha) | 2017(ha) | Change |
|---|---|---|---|
| Scotland | 1,726,900 | 1,947,750 | +220,850 |
| England (deep) | 495,828 | 495,828 | - |
| England (wasted) | 186,372 | 186,372 | - |
| Wales | 70,600 | 90,050 | +19,450 |
| Northern Ireland | 206,400 | 242,622 | +36,222 |
| UK total | 2,686,100 | 2,962,622 | +276,522 |

Of this, around 23% of UK peatland area is rewetted or remains in a near natural state (Evans *et al.*, 2017) (Fig. 1.2). Only 5.5% of the estimated 70,000ha of UK raised bog habitat remains in near natural condition (Lindsay and Immirzi, 1996). The amount of UK peatland therefore still covered by near-natural mire vegetation is low, though peatlands classed as modified contain semi-natural peatland vegetation (Evans *et al.*, 2017).

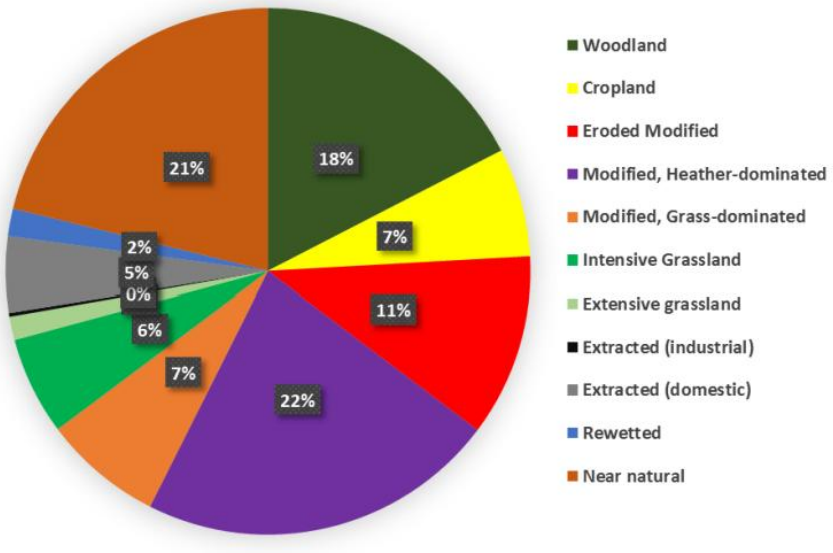



**Figure 1.2.** Distribution of peatland land cover types as detailed in Evans et al. (2017). Rewetting data is prior to 2013, and more peatlands have been rewetted and otherwise restored since that date. Note that these figures include land use on wasted peat, where erosion or decomposition has led to peat depths no longer meeting the criteria for peatlands. Modified peatlands include those with semi-natural peatland vegetation but alteration via drainage, grazing or other management/land use practices. Figure from Artz et al. (2019).

Accessibility plays a role in which peatlands are degraded. Most lowland peatlands, which are generally easier to access than those in uplands, have been drained and are still being intensively managed for agriculture. This in turn impacts the hydrology of any neighbouring near natural peatlands (IUCN, 2018).

These figures do not include shallow peaty soils, which do not have deep enough peat to be considered peatland, but can support priority habitats and other ecosystem services (IUCN, 2018). Where these have vegetation communities similar to peatlands or a similar hydrological regime, especially where *Sphagnum* mosses are present, they may also have similar surface microbial communities to peatlands, if the surface vegetation is the primary driver of the algal community.

### 1.4.2 UK peatlands and carbon

Near-natural peatlands in the UK are a significant net $CO_2$ sink of approximately 1,800 kt $CO_2$ $yr^{-1}$. Emissions of methane counteract this, and make near natural peatlands close to carbon neutral (Evans *et al.*, 2017). This is because of the higher global warming potential of methane compared to $CO_2$ (Boucher, 2012). As $CO_2$ has a longer atmospheric lifetime than methane, over longer time periods natural peatlands are large net carbon sinks (Evans *et al.*, 2017).

In the UK, near-natural bogs are very small net greenhouse gas (GHG) sources, whilst near-natural fens are a very small net GHG sink (Evans *et al.*, 2017). This may be due to greater Net Primary Productivity (NPP) of fens compared to bogs, although there is uncertainty due to the low area of near natural fen in the UK.

Modified peatlands (those with semi-natural vegetation but which have been affected by human activities to varying degrees) also tend to be net GHG sources. Although the emissions are relatively low per unit area, these cover 41% of UK peatlands and so total approximately 15% of total UK peatland GHG emissions (Evans *et al.*, 2017).

Peatlands under forestry (predominantly drained peatlands planted with commercial conifer plantations) account for an estimated 16% of UK peatland area, and are responsible for ~20% of peatland GHG emissions (Evans *et al.*, 2017). Plantation forestry is a major focus for restoration in the Flow Country of Caithness and Sutherland, which is the location of the study site for Chapter 2 of this thesis. Restoration over multiple decades has succeeded in returning some sites which had formerly been under forestry cover to net C sinks (Hambley, 2016).

### 1.4.3 Peatlands in policy

Because of the value of UK peatlands both in terms of carbon stores and other forms of natural capital, and because of the current condition and risk of shifts to C sources, there is a large amount of policy literature surrounding UK peatlands. Of particular interest is the UK Peatland Strategy which was launched in 2018 by the IUCN UK Peatland Programme. It aims to provide a strategic approach, balancing the different

pressures and demands faced by UK peatlands to “secure the best long term solution for society in a way that takes account of these multiple interests” (IUCN, 2018).

The UK Peatland Strategy takes a natural capital approach, viewing peatlands as “stocks of natural assets that provide valuable public benefits or ecosystem services for society” (IUCN, 2018). The central framing is that, if restored, damaged peatlands will increase in value (both environmental and financial), but if no action is taken will continue to lose value as they become further degraded. In this way, they argue for immediate action to avoid greater costs to society in future in case of delay.

The headline target of the UK Peatland Strategy is to have “two million hectares of peatland in good condition, under restoration or being sustainably managed by 2040” (IUCN, 2018). This is ~68% of the total UK peatland coverage, based on the 2017 estimates in Artz *et al*. (2019). There are also a number of additional goals and targets, of which those focused on biodiversity are probably of greatest relevance to this thesis. Most notable of these are points one and two of the IUCN UK Peatland Strategy goals (“Better understanding of the value of UK peatlands,” and “Prioritisation of the restoration and conservation of peatlands and sufficient resources to support this”), and the desired conservation outcomes “Environmental assessment processes are designed to assess full costs to carbon, water and biodiversity” and “Ensure the full long-term costs of potentially damaging activity is properly taken into account during the decision making process” (IUCN, 2018).

These highlight the importance of having a full picture of peatlands and their biodiversity, including cryptic biodiversity. Understanding microbial biodiversity and the things which shape it allows us to assess biodiversity fully for accurate assessment of natural capital costs. Investigating if this cryptic biodiversity changes with degradation and may be suitable for monitoring can in turn inform restoration policy and practice.

Peatlands are also mentioned in a range of international obligations, including United Nations (UN) Convention on Biological Diversity Aichi biodiversity targets, IUCN World Conservation Congress Resolution 043 “Securing the future for global peatlands”

(passed in 2016) and Ramsar Convention on Wetlands. Further details of these can be found in the UK Peatland Strategy annex 2 (IUCN, 2018).

### 1.4.4 Restoration

Over 110,000 ha of peatland restoration took place in the UK between 1990 and 2018 (IUCN, 2018). In Scotland, Scottish Natural Heritage (SNH) Peatland Action completed ~250 restoration projects from 2012 to 2018 (Artz *et al*., 2019).

There is no single definition of peatland restoration. Generally, these consider a range of measurements, including a combination of water table depth and stability, peat stability and carbon sequestration, and biodiversity (Artz *et al*., 2019). A local reference system of what constitutes a healthy and functioning peatland may be required, but if this is isolated or fragmentary, it also may not be carrying out ecosystem services to its full potential (ibid). Definitions may also require peatlands to be heading towards a new equilibrium state, preferably with new net peat accumulation (Lindsay *et al*., 2016). The Peatland Action programme uses the phrase 'on the road to recovery' instead of 'restored' (Artz *et al*., 2019), recognising that restoration is not necessarily a process which can be considered complete at the end of a project.

Restoration of UK peatlands is challenging to quantify, due to small scale individual projects having their outcomes reported separately, and often not measured in a standardised way. For example, a project may record number of dykes blocked, but not the area of peatland that consequently has a raised water table (Artz *et al*., 2019).

Goals of restoration vary, and can lead to very different outcomes depending on if the target is to restore the ecosystem to a near-natural state, or to reduce carbon losses or limit erosion (Artz *et al*., 2019; Andersen *et al*., 2017). This also leads to issues around shifting baselines, and questions of what counts as near-natural when all UK habitats have experienced long term anthropogenic influence, which must be addressed as we decide what condition to attempt to restore peatlands back to.

Costs of peatland restoration are also rarely reported consistently, making assessing the cost effectiveness of different restoration types challenging (Artz *et al.,* 2019). Costs

of longer-term monitoring have, in the past, not always been included in budgets for restoration, so this data is often lacking (ibid).

Most peatland restoration monitoring measures effectiveness through either vegetation community composition or hydrological changes e.g. water table depth (Artz *et al.*, 2019). This can lead to a knowledge gap of other peatland processes and functions under restoration. Changes in the microbial community with peatland restoration are not well categorised, in part because of the challenges involved with monitoring microbial life and the expense (and time) required for traditional microscopy methods (which are only appropriate for certain taxa) or molecular methods such as DNA metabarcoding or eDNA.

The more degraded a peatland becomes, the more challenging and expensive it is to restore (IUCN, 2018); therefore, the earlier we can detect problems the better. Microbes with rapid response times may allow opportunities for more rapid monitoring that other variables such as vegetation community, and are more multifactorial than measurements like hydrology which do not take into account nutrient availability and other variables important for biota.

## 1.5 Photosynthetic microbes in peatlands

Green algae are a major source of hidden diversity in various environments, from marine and freshwater systems to deserts. Although often considered freshwater groups, green algae have made the shift to terrestrial environments many times, the most notable being the evolution of land plants (Embryophyta) (Cardon, Gray and Lewis, 2008). Embryophyta are a clade within the Streptophyta, one of the main divisions of green algae. Land plants therefore comprise only a very small subset of green algae diversity (Cardon *et al.*, 2008).

Microbial population sizes in peatlands are high, and their communities are biodiverse (Andersen *et al.*, 2013). Photosynthetic microbes are the most diverse of all microbial functional groups in peatlands. They have been estimated to account for up to 60% of biodiversity in peatland pools, and various studies have found anywhere between 200

and 400 species of photosynthetic microbial eukaryotes and cyanobacteria in peatlands (reviewed in Gilbert and Mitchell, 2006).

Photosynthetic microbes are also known as microalgae and include both eukaryotic protists and prokaryotic cyanobacteria. While the term microalgae is sometimes reserved for eukaryotes, cyanobacteria are often discussed together with microalgae in the ecological literature (Rydin and Jeglum, 2006). For this review, cyanobacteria will be included in the term microalgae. Protists are normally single celled and include all eukaryotes apart from plants, animals and fungi, although autotrophs were historically often viewed separately from other protists as algae more closely allied to plants (Geisen *et al.*, 2018; Lahr, Lara and Mitchell, 2012). The photosynthetic protists are highly polyphyletic, and photosynthetic ability has been gained and lost on multiple separate occasions throughout the protist tree of life (Geisen *et al.*, 2018; Gilbert and Mitchell, 2006). Desmids, diatoms and mixotrophic testate amoebae are commonly found in peatlands as well as other, less abundant genera including *Euglena*, *Cryptomonas* (Cryptophyceae) and *Trachelomonas* (Euglenophycaea) (Gilbert and Mitchell, 2006). Chrysophytes, or golden algae, can also be common in peatlands, and include mixotrophic species (Kapustin and Kulikovskiy, 2023; Ma, Gong and Hu, 2018). Other algal groups found in soils include the Xanthophyceae and the Trebouxiophyceae, the latter of which include the symbiotic algae in lichens (Seppey *et al.*, 2017). Protists which are more commonly thought of as marine such as the dinoflagellates (Bates *et al.*, 2013), and the Foramanifera, some of which are mixotrophic, have also been found in soils (Geisen *et al.*, 2018; Stoecker *et al.*, 2017; Lejzerowicz *et al.*, 2010). The presence and importance of these taxa in peatlands is uncertain to the best of our knowledge. Any microalgae to survive in peatlands need to be adapted to these relatively extreme environments, in particular the low nutrient availability and the low pH (Andersen *et al.*, 2013).

Some algae, including mixotrophic testate amoebae mentioned above, are capable of phagocytosis alongside photosynthesis. These include most autotrophic dinoflagellates and some members of Ochrophyta (Geisen *et al.*, 2018). Therefore, even within food webs, the role of algae is not restricted to that of primary production.

Microalgae are able to dominate microbial communities, even when compared to heterotrophic bacteria (Gilbert *et al.*, 1998b). Desmids are often dominant in peatlands (Gilbert and Mitchell, 2006); however, testate amoebae, both mixotrophic and heterotrophic can be very abundant (Gilbert *et al.*, 1998b). Nevertheless, peatlands tend to have lower ratios of photoautotrophic to heterotrophic microorganisms compared to lake environments (Gilbert *et al.*, 1998b). This may be due to increased competition for light with other vegetation in peatlands than in the open water of lakes. In contrast, eDNA surveys suggest that heterotrophic protists far outnumber autotrophic protists in non- peatland soils, with the lowest proportion of autotrophs in forests where shading was lowest (Seppey *et al.*, 2017). This may suggest any functional significance of soil algae in peatlands is greater than the agricultural and forested soils involved in this study.

Protists, including autotrophic taxa, have generally been less studied in soils than bacteria and fungi (Wilkinson, Creevy and Valentine, 2012). They are comparatively well known and studied in peatlands from a taxonomic perspective (Gilbert and Mitchell, 2006). Despite this, the functional roles of microalgae are still relatively understudied in peatlands, and so knowledge from aquatic systems is sometimes extrapolated to form the basis of hypotheses (e.g. Jassey *et al.*, 2015). Indeed, protists which form the majority of diversity within microalgae can be considered “fundamentally aquatic creatures visiting a terrestrial world” (Rønn, Vestergård and Ekelund, 2012), and were historically generally considered to be a mainly aquatic group (Wilkinson *et al.*, 2012). This makes them well suited to peatlands, which are environments at the interface between terrestrial and aquatic. Photoautotrophic microbes are functionally important for carbon fixation in aquatic systems as primary producers (Ma *et al.*, 2023), and evidence is growing for their importance in peatland carbon fixation (Hamard *et al.*, 2021; Jassey *et al.*, 2015).

It is important that we have an understanding of microbial communities and particularly their functional roles in peatlands in order to accurately understand the impact of peatlands on global carbon balances (Andersen *et al.*, 2013). Wilkinson *et al.* (2012) identified quantifying the importance of endosymbiotic algae as energy sources to

protists as a necessary question for future research. Better knowledge of peatland microbial communities, especially those as diverse as the microalgae, is also important for quantification of overall peatland biodiversity for conservation. Creevy *et al.* (2016) showed that considering only a small number of microbial groups can dramatically increase the known diversity recorded within a nature reserve.

Photoautotrophy (that is, autotrophy involving light) in which carbon taken up from the atmosphere as $CO_2$ is fixed is the focus of this review. While other forms of autotrophy are carried out by microbes, the fixation of carbon via photosynthesis far dominates these other processes in peatlands (Gilbert *et al.*, 1998b). Therefore, only the key taxa which carry out photoautotrophy will be discussed.

### 1.5.1 Key taxa

Most studies on peatland microbial communities and populations have considered certain key groups, rather than communities as a whole (Andersen *et al.*, 2013). The same is largely true for studies on photosynthetic microbes in peatlands. Taxonomic groups which have been particularly well studied in peatlands or for their functional roles in other systems will be discussed in more detail.

#### *1.5.1.1 Green algae*

Viridiplantae (green algae) are a large and diverse group, with complex evolutionary relationships. Genetics of the Viridiplantae are complicated by many conflicting gene trees, likely due to incomplete lineage sorting (Yang *et al.*, 2023).

Traditionally, the Viridiplantae were split into chlorophytes, containing the majority of green algae, and streptophytes, containing the land plants and charophyte algae. Palmophylophyceae and Prasinodermaphyceae, formerly included as prasinophytes within the Chlorophyta, are now thought to be a distinct clade, which has been called the Prasinodermaphyta (Yang *et al.*, 2023).

Many Viridiplantae algae are marine, but the group likely diversified in freshwater environments, with transitions to terrestrial (and marine) environments occurring many times during their evolutionary history (Yang *et al.*, 2023). Whilst most prasinophytes are

marine, they can also be found in freshwater and terrestrial systems (Yang *et al.*, 2023). Ulvophyceae are generally considered marine (Cardon *et al.*, 2008) but do also contain freshwater and terrestrial species (Yang *et al.*, 2023). Several of the key algae found in peatlands belong to the Viridiplantae.

#### 1.5.1.1.1 Desmids

Desmids are members of the Desmidiales, an order of green algae. They are unicellular algae, some of which can aggregate to form colonies. Desmids are the dominant group of microalgae in peatlands (Gilbert and Mitchell, 2006). As many desmids are slow growing, they are likely to be most abundant in ombrotrophic conditions, where they have less competition from faster growing microbes with higher nutrient demands. They are also sensitive to how wet sites are and show clear microtopographic distribution patterns. Community diversity is slow to recover even after peatlands are rewetted following drainage (Goodyer, 2014).

#### 1.5.1.1.2 Mixotrophic testate amoebae

Testate amoebae are single celled protists which produce shells known as tests. They are abundant in peatlands and have a well-documented taxonomy, at least in part due to their use in palaeoecological studies (e.g. Swindles *et al.*, 2016; van Bellen *et al.*, 2014; Payne *et al.*, 2006). Testate amoebae are no longer thought to be a true taxonomic clade, but instead consist of members of the eukaryotic supergroups SAR and Amoebozoa (Creevy *et al.*, 2016; Adl *et al.*, 2012). They are well studied in peatlands, due to their tests being well preserved and varying in shape between morphospecies, making high level identification possible (Gilbert and Mitchell, 2006). Many species of testate amoebae are heterotrophic, preying on smaller protists, bacteria, fungi and micrometazoa. However, some species are mixotrophic, meaning they are capable of obtaining carbon through the autotrophy of endosymbiotic algae (Herbert *et al.*, 2019).

Mixotrophic testate amoebae are taxonomically diverse, but all have symbiotic algae from a single taxon, Trebouxiophyceae, in the genus *Chlorella* (Gomaa *et al.*, 2014). Rather than being obligate autotrophs, these amoebae are also capable of heterotrophic metabolism when conditions are unsuitable for photosynthesis, either

due to extended dark periods, lack of symbionts, or requirements for nutrients to enable photosynthesis (Worden *et al.*, 2015; Hartmann *et al.*, 2012). However, symbiont photosynthesis does seem to be the dominant energy source of mixotrophic testate amoebae (Jassey *et al.*, 2015), and reliance on autotrophy may vary between species. It has been suggested that *Archerella flavum*, which has never been observed consuming other organisms, may be entirely reliant on symbiotic autotrophy (Herbert *et al.,* 2019; Schönborn, 1965). This flexibility in nutrition may give mixotrophic testate amoebae greater resilient to changes in availability of nutrients and food resources than heterotrophic species (Herbert *et al.*, 2019).

Mixotrophs can make up a large proportion of the total testate amoeba community; up to 70% of testate amoebae identified in a French peatland were mixotrophs (Jassey *et al.*, 2015). Testate amoebae can be the dominant microbial predators in peatlands (Jassey *et al.*, 2015), but how functionally important mixotrophs are as predators does not seem to be widely known.

#### *1.5.1.2 Diatoms*

Diatoms are unicellular algae of the class Bacillariophyceae (members of the SAR eukaryotic supergroup) which are found globally in aquatic environments, both fresh and saltwater, and damp soils. They are characterised by frustules, cell walls made of silica, and so require sufficient environmental silicon to grow (Creevy *et al.*, 2016; Sheath and Wehr, 2003). The importance of diatoms in carbon cycling is well recognised in aquatic systems, especially in the deep oceans where they are estimated to be responsible for 45% of total carbon fixation (Mann, 1999 via Yool and Tyrrell, 2003). They can be found both in salt water and freshwater, as well as semiaquatic or simply damp environments. In peatlands they can exist at high biomasses, with both biomass and abundance increasing with the addition of nitrogen, alongside changes in community composition (Gilbert *et al.*, 1998b).

While diatoms are an autotrophic phylum, parasitic autotrophic diatoms have been recorded in Antarctic marine systems (Bavestrello *et al.*, 2000). Although it has never

been described, it is possible that terrestrial diatoms are also capable of mixotrophy to survive where light is limiting.

Diatoms are also notable for having acquired their chloroplasts through secondary endosymbiosis from a Rhodophyte (red algae) endosymbiont (Oudot-Le Secq *et al.*, 2007; Gibbs, 1981). Therefore, their chloroplast genes are distinct from those of the green algae (Oudot-Le_Secq *et al.*, 2007).

### *1.5.1.3 Cyanobacteria*

Cyanobacteria are the only group of bacteria which carry out oxygenic photoautotrophy, that is, photosynthesis which produces oxygen (Hamilton, Bryant and Macalady, 2016). They are functionally important in the oligotrophic ocean, where it is estimated they can sometimes account for over half of all planktonic primary production, although due to the methods used this may also include some very small photosynthetic eukaryotes (Porter *et al.*, 1985; Platt, Subba Rao and Irwin, 1983). While cyanobacteria are single celled, in peatlands they are often found in filamentous forms, with many cells aggregating together (Gilbert *et al.*, 1998b). Cyanobacteria make up a smaller proportion of the peatland microbial community than microalgae; Gilbert *et al.* (1998a) found that they accounted for 3% of the total microbial biomass. Nevertheless, cyanobacteria are functionally important in peatlands. As well as carbon fixation some species are capable of producing heterocysts, specialised cells which fix atmospheric nitrogen (Basilier, 1980). This is a significant nitrogen input for nutrient limited peatlands, and is estimated to be equivalent to the nitrogen input to peatlands from atmospheric deposition (Hemond, 1983). Cyanobacteria can also form direct associations with *Sphagnum* mosses, either on the outside of the plant as epiphytes, or within the hollow hyaline cells (Basilier, 1979; Granhall and Selander, 1973). The nitrogen fixed by the cyanobacteria is used by the *Sphagnum* (Berg, Danielsson and Svensson, 2013), possibly to prevent nitrogen being limiting for a plant which, due to its lack of roots, cannot have the same adaptations to low nitrogen environments seen in some higher plants.

## 1.6 Photosynthetic microbes and methodology

Traditionally, light microscopy has been used to study algae. It has the advantage of being relatively inexpensive and not requiring specialist laboratory resources outside of the microscope itself. However, the knowledge and skill required for microscopic identification are not inconsiderable, and it can also be extremely time consuming.

Many photosynthetic microbes are also difficult to identify using light microscopy. Within the green algae, vegetative cell features are less predictive of evolutionary relationships than internal cell features, which require electron microscopy to observe in sufficient detail (Cardon *et al*., 2008). In addition, alternating life stages and simple morphology during the non-motile vegetative stage mean that identification can be challenging without isolation and culturing (ibid). This demonstrates the value which molecular methods such as DNA metabarcoding can supply.

Depending on the molecular method used, taxonomic assignment to the species level may not be accurate. Higher taxonomic levels, however, are less dependent on the database used, and therefore more replicable. If potential indicators at these higher taxonomic levels can be identified this could be extremely useful for monitoring and for understanding community dynamics.

## 1.7 Environmental variables and changes in microbial communities

The way in which the communities of photosynthetic microbes can be driven by environmental variables means that they have been suggested as potential bioindicators (reviewed in Gilbert and Mitchell, 2006). Peatlands are nutrient limited, and so it is unsurprising that many photosynthetic microbes have strong relationships with nutrient levels and the peatland microbial community structure can be changed by nutrient input. Gilbert *et al.* (1998b) found that microalgae other than diatoms and cyanobacteria, which included the desmids and *Euglena*, were less abundant in plots fertilised with NPKCa and PKCa than in unfertilised controls. Increased nitrogen levels have also been shown to lead to a relative increase in cyanobacteria, *Euglena*, diatoms and ciliates compared to other microbes (Gilbert *et al.*, 1998a). However, this change in

community structure was only seen when the level of nitrogen deposition was very high. The differences between these two studies demonstrate the potential complexity of microbial responses to nutrients. Nitrogen pollution also leads to loss of biodiversity in plant communities as species face competitive exclusion by more nitrophilous taxa; lichens and bryophytes seem to be particularly affected (Payne *et al.*, 2012; Bobbink, Hornung and Roelofs, 1998). The specificity of the *Sphagnum* microbiome between species (Bragina *et al.*, 2012; Opelt *et al.*, 2007) implies that vegetation changes are likely to correspondingly effect microbial communities which live on the vegetation.

Microbial abundance in peatlands shows seasonal changes, with abundance seeming to be particularly negatively affected by dry periods (Potila and Sarjala, 2004). This is a logical pattern for communities which are adapted to the waterlogged conditions of healthy peatlands. However, all groups of photosynthetic organisms in a study of a French peatland had higher biomasses between June and September than for the rest of the year, presumably due to increased light levels, longer day lengths and higher temperatures (Gilbert *et al.*, 1998b). Some studies have found that peatland photoautotrophs have two periods of increased growth, in the spring and autumn (Gilbert and Mitchell, 2006).

It could be presumed that mixotrophic testates, as facultative rather than obligate phototrophs, are more resistant to changes in conditions as they can fall back on heterotrophy when conditions are not suitable for photosynthesis. Indeed, the ability to produce tests may be important as protection from seasonal environmental changes. It is possible that mixotrophic testates could dominate in drier conditions and desmids dominate in wetter conditions. Perhaps broad scale environmental differences can be detected by changes in dominance at a coarser taxonomic level, whereas patterns in community assemblage at the species level may be more easily obscured by spatial heterogeneity and unidentified confounding factors.

Dry conditions can also be caused by land management, such as the drainage cutting which is frequently seen where peatlands have been the sites of commercial peat extraction, forestry, or other intensive activities. When vegetation is allowed to recover after disturbance, *Sphagnum* mosses and their associated microbial communities

recolonise, which leads to patchy distributions of mixotrophic testate amoebae (Mitchell and Gilbert, 2004).

Alongside the direct effects on microbial taxa briefly detailed above, environmental variables could also change the microbial community indirectly. One of the probable pathways for indirect environmental effects is via changes in the vegetation. Photosynthetic microbial community dynamics are inherently linked to light. While mixotrophic organisms can acquire carbon via heterotrophy when lack of light makes photosynthesis less productive, obligate photoautotrophs which are dependent on photosynthesis for all of their carbon will be able to acquire less carbon under lower light conditions. More vegetation growth, particularly of taller plants, will create shading, which would be expected to make obligate photoautotrophs less competitive.

However, experimental shading of peatlands suggests that this relationship may be more complex. Counterintuitively, the proportion of mixotrophic to heterotrophic testate amoeba increased in highly shaded experimental plots (where light intensity was reduced by 50-60%) compared to unshaded controls (Herbert *et al.*, 2019). While this could be due to changes in relative proportion being driven by changes in population size, which was not measured in this study, there may also be mechanistic explanations. Symbiotic zoochlorella within mixotrophic testate amoebae are protected by their hosts, whereas free living microalgae are subjected to a proportionally higher grazing pressure, which could be compounded by slower reproduction with shading due to lower photosynthetic rates. These reductions in the activity of free living primary producers could therefore affect heterotrophs more than mixotrophs where shading is partial enough that zoochlorella symbionts can survive and maintain at least some photosynthesis (Herbert *et al.*, 2019). This experiment was also carried out over two growth seasons in relatively heterogeneous floating *Sphagnum* lawns (Herbert *et al.*, 2019), and so it is also possible that any changes to shading in relative testate amoebae community composition would happen at longer timescales than measured in this experiment, especially as changes in testate amoeba community composition from land use changes can be long lasting (Creevy *et al*., 2023).

The moisture levels in the local environment are often considered to be an important driver of microbial communities in peatlands (Elliott *et al.*, 2015; Andersen *et al.*, 2013). As microbes operate on much smaller scales to plants, they are likely to be less affected by water table depth, and more by soil moisture levels, which is a more direct measurement of the environmental conditions at the surface of a peatland where photosynthetic microbes are found. However, a number of potentially confounding variables could also play a role in this relationship, making it more complicated. Vegetation is also likely to be different between wetter and drier areas, meaning that shading, nutrient uptake and deposition of organic matter are also likely to differ, all of which could affect the microbial community. Furthermore, where there are specific associations between plants and the microbial community, environmental changes can affect the microbial community directly through vegetation changes. These complex interactions make understanding the causal drivers of community changes in these systems a challenge.

Soil microbes also show high levels of spatial heterogeneity. For example, testate amoebae on a seemingly uniform 0.25m$^2$ carpet of *Sphagnum magellanicum* showed significant spatial structuring at broad and fine scales (Mitchell *et al.*, 2000). Community structure was significantly correlated to the microtopography of the *Sphagnum* carpet, despite a maximum height difference of only 6.6cm. This spatial heterogeneity means that there is the potential for these microbes to be used as indicators of ecological gradients at very fine resolutions which may otherwise be difficult to detect; however, it also makes drawing general conclusions about peatland communities, or even the communities of one specific peatland, potentially challenging. Soil protists may have the ability to colonise sites which are not optimum conditions when present in nearby refugia, though this leads to patchiness of species distribution (Mitchell and Gilbert, 2004).

## 1.8 Functional roles

Key to understanding the functional role of photosynthetic microbes in peatlands is quantifying the amount of carbon which is fixed by photosynthesising microbes rather than plants. Jassey *et al.* (2015) found that a 46% reduction in abundance and 40%

reduction in biomass of mixotrophic testate amoebae alone resulted in the photosynthetic capacity of the surface of a peatland being 13% lower. This is a substantial difference, particularly considering that the biomass of other algae did not seem to be affected by the treatment, and so suggests that photosynthetic microbes may make a significant contribution to peatland carbon cycling. Gilbert *et al.* (1998b) found levels of photosynthetic assimilation varying between 5 and 261 mg C $m^{-3}$ $h^{-1}$ from water samples of peatland microbes. These were highest during the summer, and correlated to the biomass of both cyanobacteria and diatoms, as well as the concentration of chlorophyll a. These studies quantifying the amount of carbon fixation which can be attributed to peatland microbial communities build on knowledge from aquatic systems, where photoautotrophic microbes are known to contribute functionally important carbon fixation (Ma *et al.*, 2023).

Carbon fixed by photosynthetic microbes in peatlands may have a variety of fates. It could be transferred through the peatland soil food web, either via algae being consumed by algivores, or indirectly as extracellular nutrients produced as algal exudates or from cell lysis upon death of alae calls as taken up by other organisms. Alternatively, it could be lost from the peatland, either as $CO_2$ or $CH_4$ or via water as dissolved organic carbon (DOC) (Clymo *et al.*, 1998). Any carbon fixed by photosynthetic microbes which is not consumed or lost may be sequestered in the peat and stored. The fates of carbon fixed by microbial primary producers rather than plants does not seem to have been studied in peatlands.

Alongside measurements of microbial photosynthesis, it is also important to be able to quantify microbial biomass, diversity and population sizes in order to understand microbial processes and their links to ecology at a larger, systems scale (Creevy *et al.*, 2016). Microbial biomass in particular allows us to draw conclusions about the size of the microbial carbon pool. The microbial carbon pool in peatlands is smaller than the vegetation carbon pool; where microbial biomass is increased by the addition of nitrogen, it still comprises only 1% of the total carbon within the *Sphagnum* (Payne *et al.*, 2013). However, the microbial carbon pool may play a disproportionately important role in carbon cycling. Carbon fixed by plants is generally considered to be less labile

that microbial carbon. Plant litter contains complex organic molecules such as cellulose, hemicellulose and lignin, which are recalcitrant and can only be broken down into smaller, more biologically available hydrocarbons via the action of specialised enzymes (Killham, 1994). Much of the initial decomposition in non-peatland soils is therefore carried out by fungi, which are capable of producing these enzymes (A'Bear *et al.*, 2014; Burke, Smemo and Hewins, 2014). In peatlands, this relationship is more complex. While saprotrophic fungi are generally considered to be the dominant decomposers in surface layers where oxygen is rarely limited, some studies suggest that bacterial biomass may be higher and bacteria more active than fungi in some peatlands (reviewed in Andersen *et al.*, 2013). Regardless of the main decomposers, the polymers which make up microbial cell walls, in contrast, are generally considered to be more labile, as they appear to be relatively rapidly broken down into smaller hydrocarbons and so can be more quickly taken up by other organisms when microbes degrade (Kögel-Knabner, 2002).

Microalgae also release fixed carbon to the surrounding environment as exudates; Wyatt and Turetsky (2015) found that 38% of carbon fixed by algae was lost as exudates. Old plant material, especially *Sphagnum*, is incorporated into accumulating peat, and it seems reasonable to expect that the more recalcitrant complex polysaccharides would be the most likely to be preserved. Microbial carbon, tending to be more labile, may be more likely to be used as a carbon source, particularly by other microbes. If microbial carbon does have a different fate to plant carbon on a generalisable scale in this way, then the carbon fixed by microalgae could be an important unquantified factor in our knowledge of peatland carbon cycling.

Photosynthetic microbes play an important role in soil nutrient cycling through the carbon they fix as primary producers being passed on to higher trophic levels. They are likely to be used as a food source by heterotrophic microbes, especially microbial grazers, which do not graze extensively on plants due to their size. Many protists are phagotrophs, acting as predators of other protists and bacteria (Geisen *et al.*, 2018), and prey taken up are likely to be in a size range which allows them to be engulfed, which plants and larger organisms are not (Boenigk *et al.*, 2002). Specialist algivorous

protists are common in soils, reinforcing the idea of algae as important soil primary producers for microbial soil food webs (Busch and Hess, 2017; Seppey *et al.*, 2017). Nematodes predate small soil organisms including algivores, as well as directly filter feeding on bacteria including cyanobacteria and possibly other microalgae (Rønn *et al.*, 2012). This potentially important nutrient flow from microbial primary producers to microbial omnivores and parasites has rarely been considered in soil food web models (Geisen and Bonkowski, 2018). Photosynthetic algae also act as a measurable and potentially ecologically significant carbon source for soil macrofauna including earthworms and Collembola in experimental systems (Schmidt, Dyckmans and Schrader, 2016). It is likely that this relationship also exists in peatlands. However, the complex food webs of peatland microbes make investigating predator-prey relationships challenging, and so currently we have to extrapolate from interactions demonstrated in other soil systems.

The microbial loop is also likely to be important when considering the functional roles of photosynthetic microbes in peatland nutrient cycling. This is a description of nutrient and energy flows within microbial communities. It originates in studies of aquatic systems, but is also found in semiaquatic environments and in soils (Azam *et al.*, 1983). It is likely to be important for peatland nutrient cycling, as organic matter is abundant but inorganic nutrients are often limited (Gilbert *et al.*, 1998b). In summary, heterotrophic bacteria which are able to use dissolved organic matter which is not bioavailable to higher trophic levels as a carbon source rapidly increase in abundance when nutrient levels are high. This then allows increases in populations of bacterial grazers such as ciliates, flagellates and heterotrophic amoebae, which in turn keep the microbial population in check and act as food sources for larger organisms (Gilbert *et al.*, 1998b). The microbial loop allows essential nutrients to be recycled within the microbial community, as well as potentially promoting growth of the primary producers (Porter *et al.*, 1985). In very nutrient limited ecosystems such as peatlands, it is possible that the initial seasonal increase in nutrients in summer is due to increased abundance of photosynthetic microbes. These release labile carbon as exudates, as well as acting as a source of a wider range of nutrients when they die (Wyatt and Turetsky, 2015). Nutrients from algae can also become bioavailable due to cellular contents lost as a

result of ‘messy eating’ by algivores, as is documented in aquatic systems (summarised in Williams, 1981). If these nutrient inputs include DOCs then they could provide the carbon source for the heterotrophic bacteria which are the foundation for the traditional view of the soil microbial loop. The soil microbial loop has historically considered heterotrophic microbes to be key as the main food source for protists, but algivory is now being shown to be a potentially important feeding strategy for protists in soils (Seppey *et al.*, 2017). This suggests that algae may feed into soil microbial food webs in a similar way.

The microbial loop is also expected to be different in peatlands to aquatic environments, where it has been studied in more detail, because of the presence and importance of testate amoeba in peatlands, which have complex trophic interactions (Gilbert *et al.*, 1998b). As peatlands are nutrient limited systems where algae are known to be diverse and abundant, the role of algae as a direct carbon source through grazing or as a source of nutrients to feed into the soil microbial loop seems likely to be functionally important. Indeed, increases in algal abundance due to nutrient enrichment have been shown to allow increases in bacterial abundance in a fen, where bacteria are limited by labile carbon (Wyatt and Turetsky, 2015).

Grazing is also likely to directly affect microalgae populations. Grazing reduces bacterial biomass and abundance, but it can also stimulate bacterial activity. This may be because of grazing acting as a selection pressure for greater microbial activity, or because grazed populations are maintained in a growth stage rather than reaching a stationary phase where a limiting factor can reduce their rates of activity (Ekelund and Rønn, 1994). Furthermore, when microbes (presumably including photosynthetic microbes) are predated by phagotrophic protists, bioavailable nutrients are released to their surroundings. Where grazers (usually heterotrophic protists) have a similar C:N ratio to the microbes being grazed, the carbon lost by grazers through respiration will result in them gaining excess N from their prey. This N is then excreted by the grazers as ammonia (Rønn *et al.*, 2012; Sherr, Sherr and Berman, 1983). This mineralised N is then available for uptake by other organisms including plants (Geisen *et al.*, 2018). This could be a form of priming, where increased N levels in normally ombrotrophic peat bogs

allow for the loss of old stored carbon through microbial activity. Changes in bacterial communities and their activity through grazing can increase vegetation growth in experiments on non-peatland soils (reviewed in Rønn *et al*., 2012); it does not seem unreasonable that similar processes could be observed with grazing of microalgae in peatlands. Any stimulation of vegetation growth could in turn affect the microalgae by increasing shading, further increasing the complexities of interactions between photosynthetic microbes, heterotrophic microbes and plants.

Algae, by releasing DOC and other nutrients as exudates and when undergoing lysis, can be expected to contribute to the soil microbial loop. Algae will also act as primary producers and so provide an input of carbon which moves through the soil food web via algivory by soil protists and soil animals (Seppey *et al.*, 2017; Schmidt *et al*., 2016). The role of soil algae may be important even if they are at relatively low biomass levels due to their relatively rapid turnover time when compared to microbial predators (Seppey *et al.*, 2017).

The function of peatland microbial communities may be susceptible to changes in environmental conditions. The relatively high turnover rate of microbial communities may also make them more sensitive to environmental perturbations, and result in them responding more rapidly than other organisms (Gilbert *et al.*, 1998a). When nitrogen levels are increased, total microbial biomass increases, with autotrophs particularly showing increases (Payne *et al.*, 2013; Gilbert *et al.*, 1998a). Addition of nitrogen also results in increases in photosynthetic assimilation by microbes (Gilbert *et al.*, 1998a), which could be due to increases of biomass but may also indicate functional changes. Seasonal changes in the community structure of photosynthetic microbes may also result in functional changes. Potila and Sarjala (2004) found that the decrease in microbial abundance due to dry periods was associated with decreases in gross N mineralisation, demonstrating a link between microbial abundance and function.

## 1.9 Links to thesis

The effects of environmental stressors on peatlands tend to be complex, with impacts on microbes varying by taxonomic group and site. Therefore, understanding their impacts on peatlands in good condition is vital for accurate planning of peatland restoration and ensuring that conservation efforts are well allocated. Even when restoration is working with peatlands that are far from pristine, understanding the processes in more near-natural peatlands allows us to make informed decisions about what we aim to achieve with restoration.

The studies in this thesis have been carried out on generally nutrient poor and ombrotrophic peatlands in Scotland, with intact mire vegetation communities and limited active drainage. This allows us to investigate the impact of both long- and short-term environmental stressors on bogs in otherwise good condition. For fire and nutrient pollution in particular, there is reason to believe that these will have strong effects on the peatland algal community; however, this has not been widely studied in the field across all algal taxa. It is generally considered that UK peatlands have limited natural tree cover, and changes in hydrology and shading from trees are likely to impact photosynthetic microbes. Ensuring that our understanding of these processes is evidence-based is vital in assessing how to protect peatlands, and the biodiversity and carbon they contain.

Bioindicators can provide a helpful tool for monitoring environmental conditions and the progress of restoration. If algae are sensitive to the kinds of environmental stressors which concern peatland management, they may provide a valuable indicator of peatland surface conditions. Mixotrophic testate amoeba (MTA) have been identified as useful indicators (Gomes da Silva *et al*., 2022), including for forest to bog ecotone transitions (Creevy *et al.,* 2018), but algae considered as a whole may be less suitable. MTA have been identified as an indicator in microscopy studies of testate amoeba, rather than as part of the wider autotrophic community. Non-mixotrophic testate amoeba are heterotrophic, so mixotrophic species are distinct as being highly dependent on light availability. However, amongst algae, which are inherently autotrophic (or mixotrophic if extending the definition of algae as I have here), this is no

longer a noticeable trait, and so individual algal taxa may be less likely to be useful indicators. The greater numbers of algal species when considering all taxonomic groups, and therefore greater variation, may make total algal community more difficult to use as indicators than testate amoeba which are well known and well described. Therefore, understanding the relationships that the algal community has with different environmental stressors is key to determining if they can serve as useful bioindicators.

# 2 Photosynthetic microbial community changes following fire

## 2.1 Introduction

Fires are a threat to temperate and boreal peatlands. Within the UK, recent years have seen multiple large fires on peatlands, especially in periods of hot and dry weather. Wildfires across the UK are predicted to increase with climate change (Belcher *et al.*, 2021; Shotter and Wentworth, 2019). Whilst there has been little specific research on peatlands, they are likely to experience the same pattern. Whilst damage to the vegetation communities and changes in carbon cycling in peatlands following wildfires have been widely studied (Lukenbach *et al.,* 2015; Magnan, Lavoie and Payette*,* 2012; Worral *et al.,* 2011, etc), the impacts of fire on the surface microbial communities are less well understood.

Photosynthetic microbes, which are presumed to be largely constrained to the peatland surface by light requirements (Geisen *et al.*, 2018), would be expected to be affected even by shallow fire damage, especially if the bryophytes on which they live are damaged or destroyed. Photosynthetic microbes (microalgae) are abundant and diverse on the surfaces of peatlands (Gilbert and Mitchell, 2006). As well as providing a large pool of cryptic diversity, they also contribute to peatland carbon fixation via photosynthesis (Jassey *et al.*, 2015; Gilbert *et al.*, 1998b). In addition, some photosynthetic microbes such as cyanobacteria carry out nitrogen fixation. N fixing cyanobacteria are frequently found as symbiotic partners of peat-forming *Sphagnum* mosses and as such may mediate their growth in what are otherwise environments of extremely low nutrient availability (Berg *et al.*, 2013).

Healthy peat bogs often have complex microtopography, encompassing hummocks, lawns and hollows (Rydin and Jeglum, 2006). These provide a range of niches for microbes, with different vegetation, moisture and shading all likely to be important drivers of the microbial community. Fire impact on peatland vegetation is determined in part by hydrology, with wetter conditions associated with a high water table associated

with less fire damage (Lukenbach *et al*., 2015). Logically, therefore, drier hummocks, which sit further above the water table than lawn or hollow features, would be expected to be subject to greater fire damage. Conversely, however, studies of peatlands in boreal forests in Canada show hummocks are protected by the high moisture content and capillary action of *Sphagnum* mosses, and thus less damaged by fire than the surrounding lawns (Nelson *et al*., 2021; Thompson and Waddington, 2013). However, the *Sphagnum* species mentioned in Nelson *et al*. (2021) are not universal (and were not found in the field site for this chapter). Therefore, to fully understand the effects of fire on peatland microbes, we need to understand their responses and potential recovery at a scale which accounts for these microtopographic landforms, and to understand this relationship in a local setting.

Photosynthetic microbial response to fire may be complex; there are also possible positive interactions from fire, such as a burst of nutrients, or removal of taller vegetation reducing shading and therefore competition for light. Therefore, as well as changes in community composition and structure from fire, there is also potential for increased abundance and biomass of microbial primary producers due to lack of shading competition from plants.

The May 2019 Forsinard wildfire gives the opportunity to investigate the impact of wildfire on a near natural bog with complex microtopography, and the varying levels of fire damage associated with this. In this study we will attempt to identify fire impacts on the whole photosynthetic microbial community, rather than focusing on core algal groups. The goal is for this to allow the study to encompass variation between larger microalgae such as diatoms and desmids, as well as filamentous algal forms and cyanobacteria.

### 2.1.1 Aims and hypotheses

(Wild)fires are predicted to increase in boreal systems due to climate change, but their effects on photosynthetic microbial communities are largely unknown. This chapter aims to better understand to what extent intact near-natural peatlands, and their

photosynthetic microbial communities, are vulnerable to or resilient to fire, and if more varied microtopography can increase resilience.

In a hummock- hollow system, more fire damage would likely be expected on the drier hummocks than the *Sphagnum* filled hollows and lawns (unless the pattern found in boreal forest peatlands also applies to more open peatlands; see above). Therefore, as more of the bryosphere is lost, more changes in the photosynthetic microbial community would be expected from hummocks, even within the remaining bryophytes if fire damage is not total. If this is less extreme for hollows, these could act as a refugia for microbial species which can be found across the range of microhabitats (likely those able to withstand inundation and periods of drying) to recolonise from after fire. This would result in hummocks and hollows being more similar to each other after the fire than in unburnt sites.

Hypotheses to test can be split into three broad categories: environmental heterogeneity, fire severity, and microbial community changes.

Environmental heterogeneity:

- Hollows and hummocks have different photosynthetic microbial communities.
- These are driven by differences in environmental conditions between microtopographies, especially soil moisture content.

Fire severity:

- Statistical analysis matches anecdotal field observation that, in Forsinard, hummocks were more severely impacted by fire than hollows.

Microbial community:

- Fire causes an overall decrease in microbial diversity indices, as well as changes in community composition.
- There are differences in resistance or resilience to fire for different microtopographic landforms. Wetter hollows will be more resilient to community shifts from fire disturbance than hummocks.

## 2.2 Methods

### 2.2.1 Site description

The RSPB Forsinard Flows NNR is a large nature reserve in the Flow Country of Caithness and Sutherland (Fig. 2.1). In May 2019, it experienced a significant wildfire, which burned over a number of days. The fire damage pattern was complex, with changes in wind direction meaning that there was no single fire front, and some areas were burnt more than once. In total, the fire burnt over 60km$^2$ of the Flow Country, including designated blanket bog in the Special Area of Conservation (SAC).

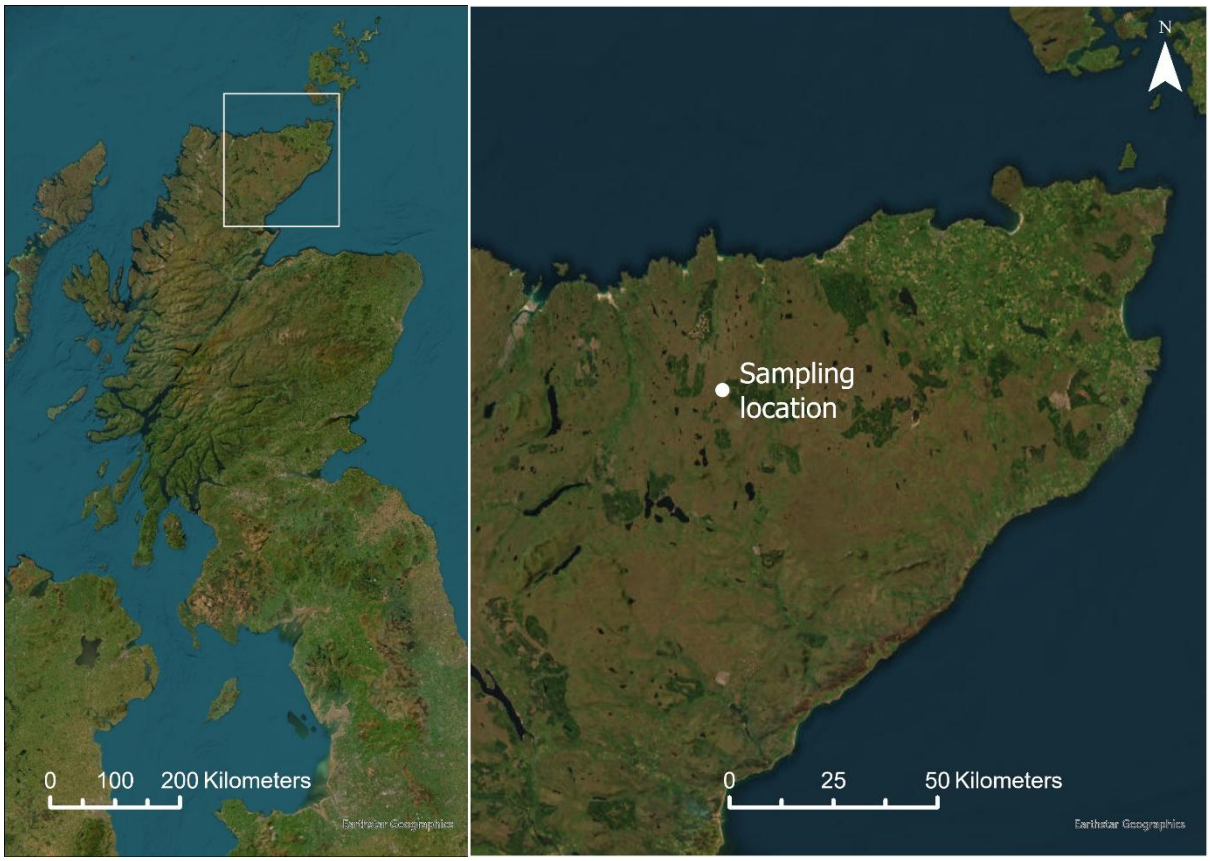



**Figure 2.1.** The location of the RSPB Forsinard Flows NNR study site (sampling location) within the UK and within the highlands of Scotland. Satellite imagery from ArcGIS World Imagery (Sources: Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community).

Field work was carried out in early August 2019, around 2 and a half months after the fire. The area sampled consisted of near natural blanket bog managed by the RSPB. Although there had been some history of drainage, this had largely been restored. Vegetation was *Sphagnum* dominated, with a significant pool system to one side (Fig. 2.2). Microtopography was varied, with complex hummock-hollow-lawn forms, alongside deeper pools. *Sphagnum* communities were diverse, with many species being observed.

### 2.2.2 Field methods

A transect was taken across the fire margin, with sampling sites every 100m to allow for the assumption of independence (Fig. 2.2). 6 sampling sites were included in the burnt area and 6 in the unburnt area (Fig. 2.3). Sampling sites were selected by walking along the transect for 100m from the previous site. A coin was flipped to determine left or right, and the first hummock on this side of the transect was the centre of the sampling site. The nearest identifiable hollow to this hummock was then identified and also sampled. The nearest *Sphagnum* lawn was also identified and sampled from. However, heavy rainfall over the course of field sampling made it increasingly difficult to distinguish lawns from pools, and so the lawn data was discarded.

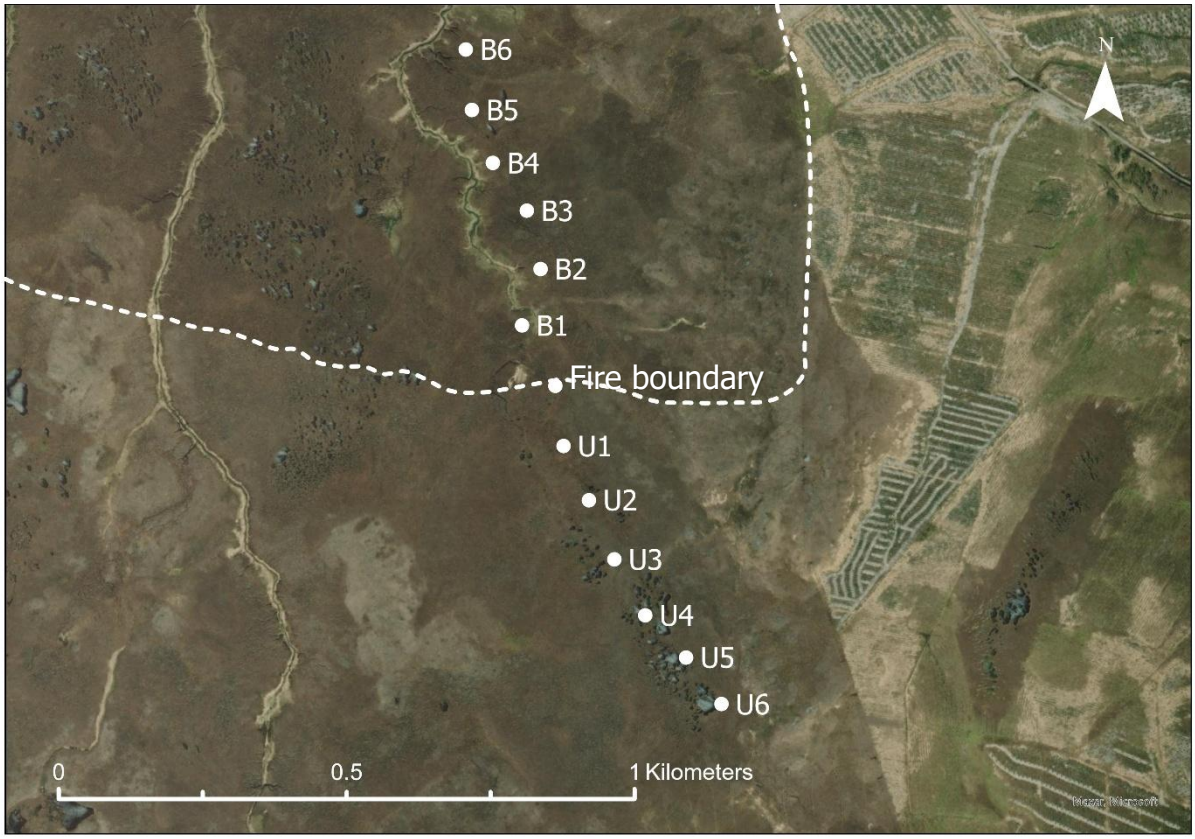



**Figure 2.2.** Location of hummock sampling points at RSPB Forsinard Flows NNR either side of the margin between burnt and unburnt areas. Each hummock is paired with the closest identifiable hollow. Points prefixed with B are burnt, and those prefixed with U are unburnt. The fire margin is visible from the satellite imagery, and highlighted with a dotted line. Also visible is a pool system, especially around points U5 and U6. Satellite imagery from ArcGIS World Imagery (Sources: Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community).

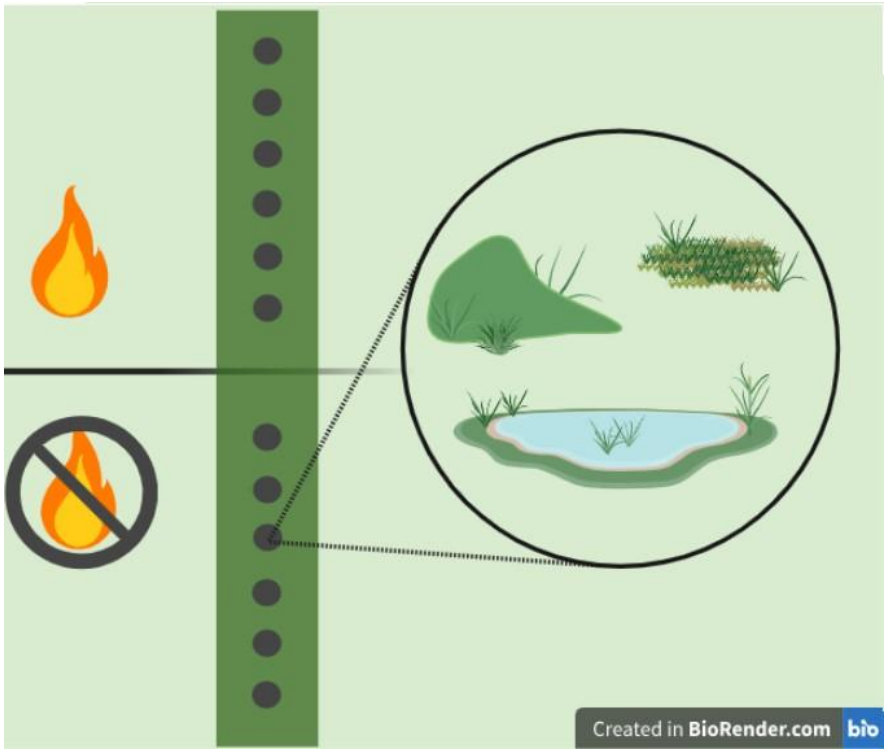



**Figure 2.3.** A schematic of field sampling methodology for Chapter 2. At each sampling point, a hummock, hollow and lawn were visually identified and sampled.

*Sphagnum* samples for laboratory analysis were taken from the dominant *Sphagnum* species found in each quadrat. The top 10cm, including the capitulum, was sampled to incorporate the photic layer to prioritise photosynthetic microbes. Soil moisture of the bryophyte layer was measured with a conductivity probe (HH2-Theta Probe, type ML2x, Delta-T Devices, Cambridge, UK), operated according to the manufacturer's instructions, with a mean taken across three measurements. Vegetation community assemblage and percentage cover was estimated using a 30cm$^2$ quadrat, and maximum vegetation height above the upper surface of the *Sphagnum* was measured within the quadrat.

### 2.2.3 Peatland composite burn index

A modified peatland composite burn index (PCBI) was used to assess severity of fire impact for burnt sites (adapted from Davies *et al.*, 2016). This has been linked to microbial and ecosystem functions in other settings (e.g. Holden *et al.*, 2016). This was

calculated as the mean of scores for *Sphagnum* damage (both discolouration and capitulum loss), moss scorch/consumption and overall *Sphagnum* survival. These figures were based on percentages of moss affected (as detailed in Table 2.1) in the 30cm$^2$ vegetation survey quadrats. The PCBI score ranges from 0 (no fire damage to moss) to 3 (all *Sphagnum* consumed, with no survival).

**Table 2.1.** Details of peatland composite burn index scoring, as used in the field to quantify fire damage to *Sphagnum*.

| Peatland CBI score | | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |
|---|---|---|---|---|---|---|---|---|
| *Sphagnum* damage | Discolouration | unchanged | -- | 30% | -- | 50% | 75% | 100% |
| | Capitulum loss | none | <10% | -- | 30% | -- | 70% | 100% |
| Moss scorch/ consumption | | none | -- | 50% scorch | 100% scorch | -- | 50% consumption | 100% consumption |
| *Sphagnum* survival | | 100% | -- | 90% | -- | 50% | <20% | none |

For example, for sample B1la (burnt site 1, lawn) there was no visible moss scorch or consumption, but visible capitulum discolouration of between 30% and 50% of the quadrat area and <10% capitulum loss (Fig. 2.4). Therefore, *Sphagnum* damage was scored as 1.0 (the mean of the scores for discolouration and capitulum loss), moss scorch/consumption was scored as 0 and *Sphagnum* survival was scored as 0.5 (between 90% and 100% survival). This gave a total PCBI score of 0.5.



**Figure 2.4.** Sample B1la (burnt plot 1 lawn), with a peatland composite burn index score of 0.5. The quadrat area is 30cm$^2$. *Sphagnum* capitula are discoloured but there is no capitula loss or visible scorching, and almost complete survival.

In contrast, sample B2hk (burnt plot 2 hummock) had nearly complete *Sphagnum* discoloration (but not 100% capitulum loss), under 50% consumption (with the entire moss burnt away rather than simply scorched) and only around 50% *Sphagnum* survival

(Fig. 2.5). Therefore *Sphagnum* damage was scored as 2.5, moss scorch/consumption as 2 and *Sphagnum* survival as 2, giving a total PCBI score of 2.17.



**Figure 2.5.** Sample B2hk (burnt plot 2 hummock), with a peatland composite burn index score of 2.17. The quadrat area is 30cm$^2$. Almost all *Sphagnum* is discoloured but not all capitula have been lost. Consumption (complete destruction of moss) is visible.

Sample B5hk had worse *Sphagnum* capitulum damage, with complete discolouration, but less scorch and only limited consumption, with more than 50% *Sphagnum* survival (Fig. 2.6). Therefore *Sphagnum* damage was scored as 3, moss scorch/consumption as 1.5 and *Sphagnum* survival as 1.5. The PCBI score was 2.



**Figure 2.6.** Sample B5hk (burnt plot 5 hummock), with a peatland composite burn index score of 2. *Sphagnum* discolouration and capitulum loss is extensive, but scorch is incomplete and consumption is limited.

## 2.2.4 Microscopy

### *2.2.4.1 Microbial extraction*

*Sphagnum* leachate was collected for microscopic analysis, using a method adapted from Jassey *et al*. (2011) and Goodyer (2014).

In brief, a small handful of *Sphagnum* was squeezed whilst wearing gloves to capture pore water. 10ml $dH_2O$ was then added to the *Sphagnum* and allowed to absorb. The *Sphagnum* was squeezed again. 30ml $dH_2O$ was then added to the *Sphagnum* and shaken on a benchtop shaker at 180RPM for 90 minutes, before squeezing again to extract liquid. All leachate fractions were combined and passed through a 500µm sieve

to remove large vegetation fragments. A combination of squeezing and shaking was chosen to maximise yield of microbes from the *Sphagnum* surface, without damaging the more fragile microalagae by excessive force (Goodyer, 2014).

As particles settled to the bottom of the sample bottoms during storage, 1ml of the settled material was subsampled from each sample. This subsample was pipetted 5 times until visibly well mixed. 40µl was pipetted from this subsample onto the slide and mixed with 40µl of glycerol. Slides were sealed with clear nail varnish.

#### *2.2.4.2 Count number for microscope work*

Target count number was calculated by rarefaction, to calculate the impact of sampling effort on species richness. Seven samples at random across all sample types were counted under the microscope to a minimum of 300 individuals. Cells were identified and classified as different morphospecies. Species accumulation curves were calculated in iNEXT (Hsieh, Ma and Chao, 2020) in R version 3.6.0 (2019-04-26). The asymptote (modelled species richness) was calculated using iNEXT for each sample. The modelled sample size (number of individuals counted and identified) required to exceed the lower 95% confidence interval of the species richness prediction were found. The mean sample size required to exceed the lower 95% confidence interval was 374.

In order to simulate randomly ordered species accumulation curves, to remove the impact of sampling order, individuals were randomly sampled without replacement from each sample of microscope data to give a simulated sample size between 0 and the total number of individuals counted in the sample, increasing in intervals of 10. The number of unique species encountered in each simulated sample was calculated, and this was iterated 100 times for each simulated sample size. A species accumulation curve was constructed from this simulated data for each microscope sample by applying the Michaelis Menten equation using the DRC package (Ritz *et al.*, 2015).

A minimum count of 375 individuals was the target for all microscope work. Samples which had previously been counted to a minimum of 300 were revisited and further cells counted up to the new minimum. However, the data analysed still includes some of the

original, lower count numbers. Fixing the samples with Lugol's Iodine was not sufficient to prevent all fungal growth over long term storage, and when new slides were prepared to allow for new counts up to the total, there were very few non-fungal microbes present. The impact of adding data biased towards only those taxa such as diatoms and testate amoeba with shells resistant to fungal decomposition was judged to be greater than the impact of lower count number. This highlights the need to be cautious when using Lugol's Iodine for fixing of samples which are likely to contain larger amounts of fungi than the aquatic samples for which the method is more typically used.

### 2.2.5 Data analysis

All data analysis was carried out in R version 4.1.2 (2021-11-01). Biodiversity calculations and ordinations were carried out using 'Vegan' version 2.6-4 (Oksanen *et al.*, 2022). Soil moisture data resisted attempts at normalisation; therefore Wilcoxon rank sum was used to test for the effect of microtopography on soil moisture across burnt and unburnt samples. For samples in the burnt area, statistical significance of microtopography on measured fire severity was calculated with a Wilcoxon rank sum test in R.

### 2.2.6 Microbial community data

To determine if fire caused an overall decline in biodiversity, inverse Simpson's Diversity Index was calculated for all samples using the 'diversity' function in Vegan. A Wilcoxon test was used to compare burnt and unburnt sites, and for burnt sites a Spearman's rank correlation was carried out between inverse Simpson's diversity and fire severity score.

Simpsons diversity index was chosen over other diversity indices (Shannons etc.) due to relative robustness at smaller sample sizes (Magurran, 2004), and less emphasis on rare species, which we considered to be likely to be less biologically important, especially for the functioning of microbial communities.

Microbial community data was converted to relative abundance, a Bray-Curtis distance calculated, and an NMDS carried out using the function MetaMDS in Vegan. NMDS was

selected as the ordination method due to resilience to zeros and non-parametric data, making it suitable for the microbial dataset. Stress testing was carried out, with 4 dimensions selected as the lowest number with a stress below the threshold of 0.10 (Fig. 2.7).

PERMANOVA analysis was carried out using the adonis2 function in vegan to test for statistical significance in microbial communities between burnt and unburnt samples and hummock and hollow microtopography. Adonis2 was also used to test for statistically significant differences in microbial community with soil moisture, maximum vegetation height and fire severity (for continuous variables, this is analogous to a linear modelling approach).

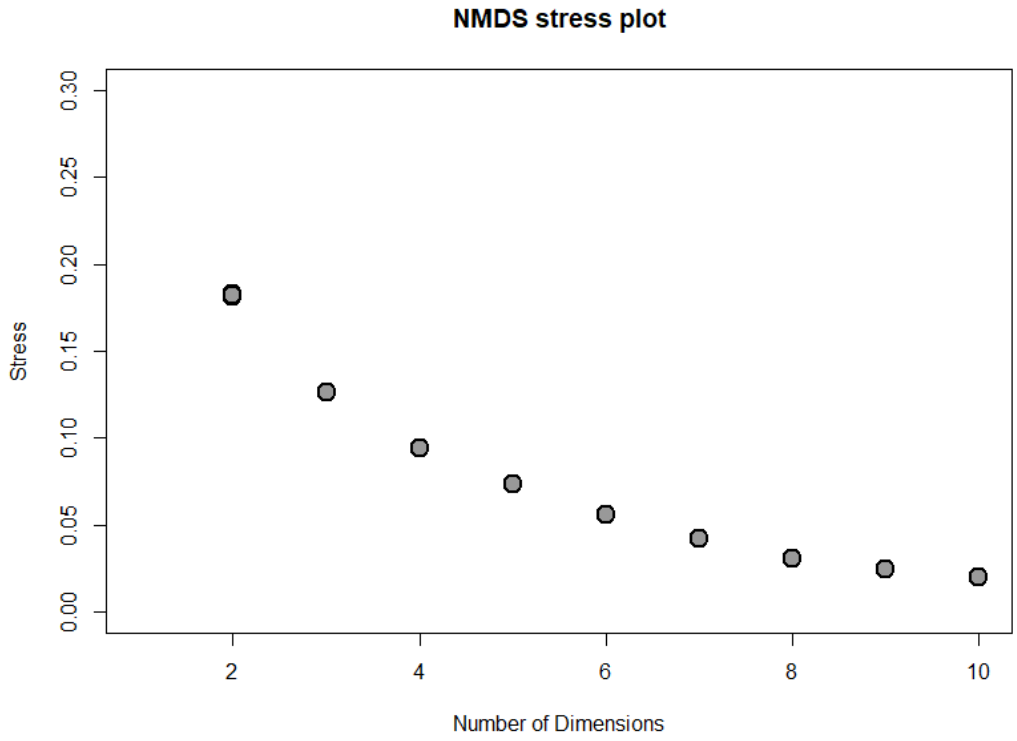



**Figure 2.7.** NMDS stress plot for microbial community Bray-Curtis distance matrix. Stress decreases with increasing number of dimensions for microbial community data. 4 was the lowest number of dimensions below the 0.10 stress threshold.

### 2.2.7 Indicator species analysis

Indicator species analysis was carried out in R using function multipatt from the indicspecies package (De Cáceres and Legendre, 2009). Briefly, it calculates the indicator value of species (indval) found across the groups of sites. The indicator value is the product of values of specificity (A: the probability that, when a species is found in a site, the site will be of the group that the species 'indicates') and fidelity/sensitivity (B: the probability that, for a given site of the correct grouping, the indicator species will be found) (Dufrêne and Legendre, 1997). The statistical significance was tested using a permutations test (999 permutations). As indval could not be calculated for two different site groupings simultaneously, this analysis was repeated for microtopography (hummock and hollow) and fire status (burnt and unburnt).

## 2.3 Results

### 2.3.1 Environmental data

As expected, hollows have statistically higher soil moisture % than hummocks across both burnt and unburnt samples ($W = 0$, $p < 0.001$) (Fig. 2.8).

Microtopography had a significant impact on fire severity ($W=36$, $p < 0.005$), with hummocks more severely affected than hollows (Fig. 2.9). Whilst damage for hollows was generally low, hummocks on average had a fire severity score equivalent to between 50% and 20% *Sphagnum* survival. This may be explained by the relationship between fire severity and soil moisture. All six hollow samples had a mean soil moisture of 100%, with four of these also having a mean fire severity score of 0 (Fig. 2.10). The relationship between fire severity and soil moisture could not, however, be statistically tested for. As all hollows had the same soil moisture, a correlation approach was not suitable. As the grouping was between the different microtopographic landforms, a Wilcoxon rank sum test would have simply repeated the test of fire severity differences between microtopographies described above.

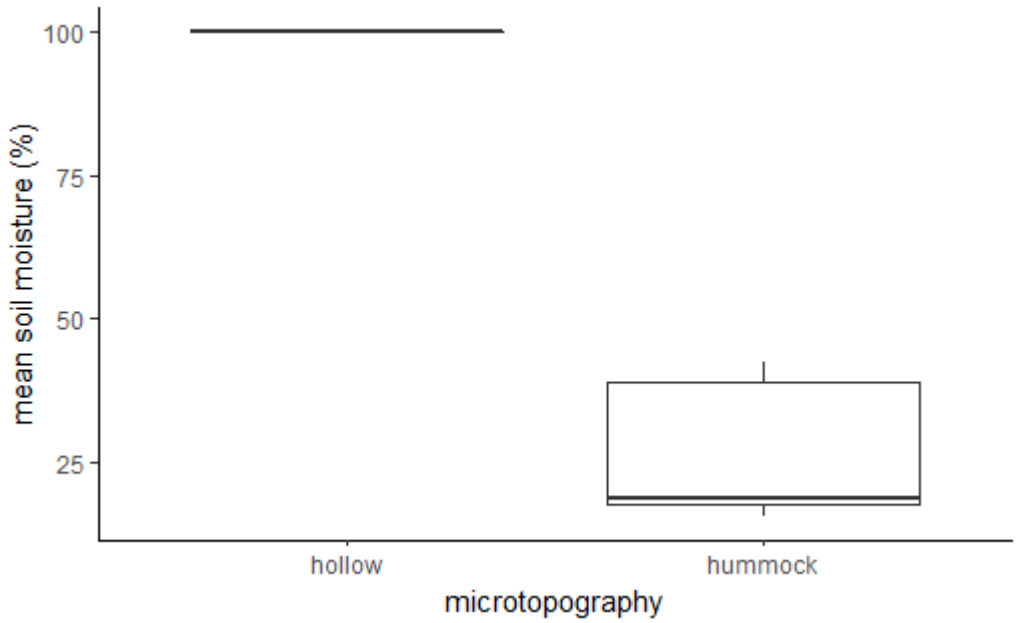



**Figure 2.8.** Box plot showing the mean soil moisture of different microtopographic landforms. Hollows had a higher mean percentage soil moisture than hummocks, with all hollows having a mean soil % moisture of 100%. Boxes show median and interquartile range.

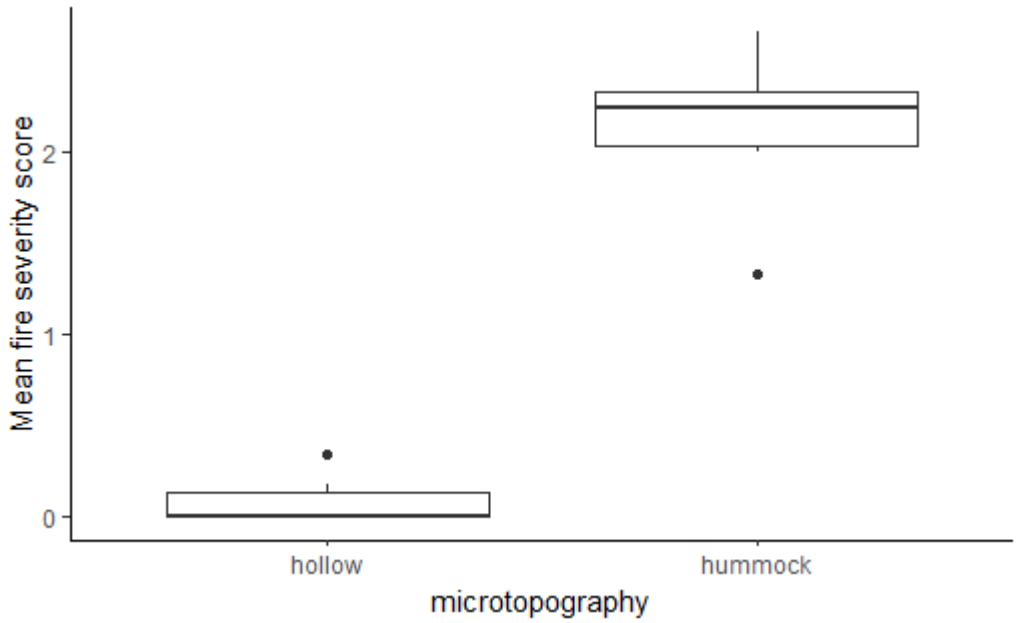



**Figure 2.9.** Box plot showing the effect of hummock/hollow microtopography on fire damage, calculated as a mean of fire severity scores. Boxes show median and interquartile range, and points show outliers.

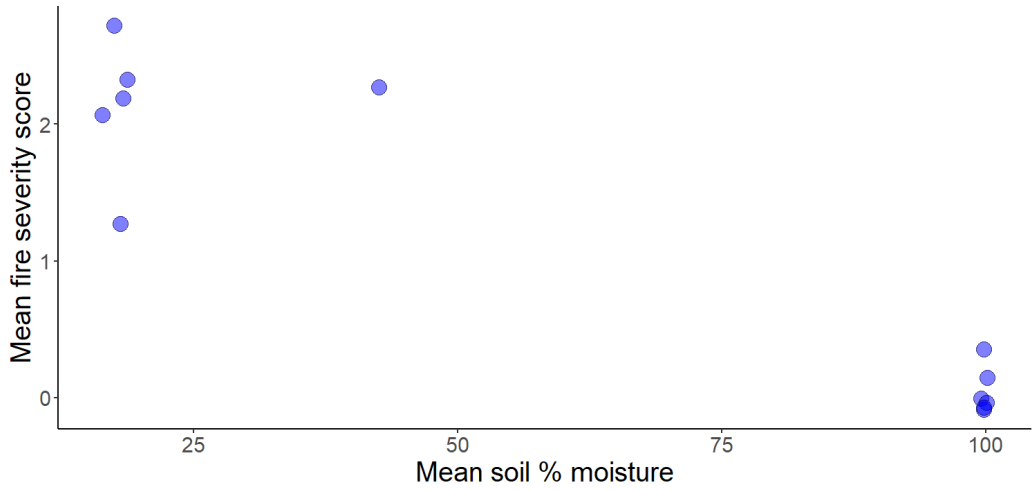



**Figure 2.10.** Fire severity plotted against mean soil moisture content. Fire severity is calculated as a mean of three distinct fire impact measurements (see Section 2.2.3). All six hollow samples had a mean soil percentage moisture of 100%, with four of these also having a mean fire severity score of 0. This caused overplotting- the points have been jittered using geom_jitter in ggplot2 to reduce this and to make all the points visible.

### 2.3.2 Diversity

There were no significant differences in Simpson's Biodiversity between burnt and unburnt samples (W = 70, p = 0.9, Fig. 2.11). There was also no significant correlation between fire severity and biodiversity for burnt samples (R = -0.3779693, p = 0.23).

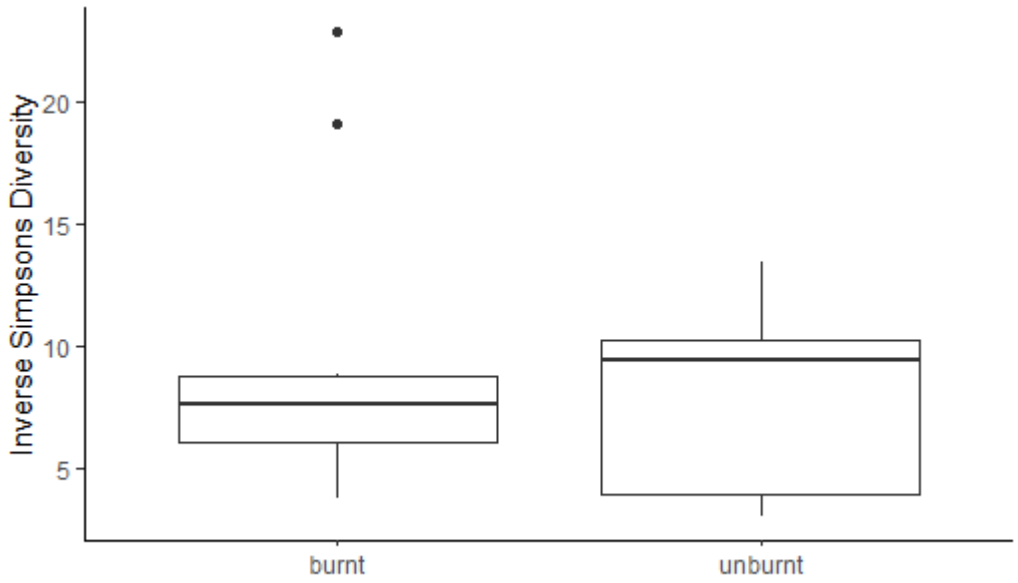



**Figure 2.11.** Boxplot showing difference in inverse Simpson's biodiversity with fire. There was no significant difference in inverse Simpson's biodiversity between burnt and unburnt samples. Boxes show median and interquartile range, and points show outliers.

### 2.3.3 Microbial communities

Though the clustering is imperfect, the NMDS plots show microbial communities are more similar to each other within the same microtopographic feature type (Fig. 2.12). On average, hummocks seem more similar to each other, with hollows showing greater variation. The NMDS plot shows no clear clustering between burnt and unburnt samples, although unburnt hollows appear to be more distinct towards the left of the plot.

PERMANOVA analysis showed a small but significant effect of microtopography on microbial community ($R^2 = 0.122$, $p = 0.002$ for 999 permutations). No other variables had a statistically significant effect on the microbial community. The photosynthetic microbial community showed no significant differences with fire or fire severity. Alongside the lack of changes to Biodiversity indices, this suggests resilience of the surface photosynthetic microbial community in a hummock-hollow peatland system.

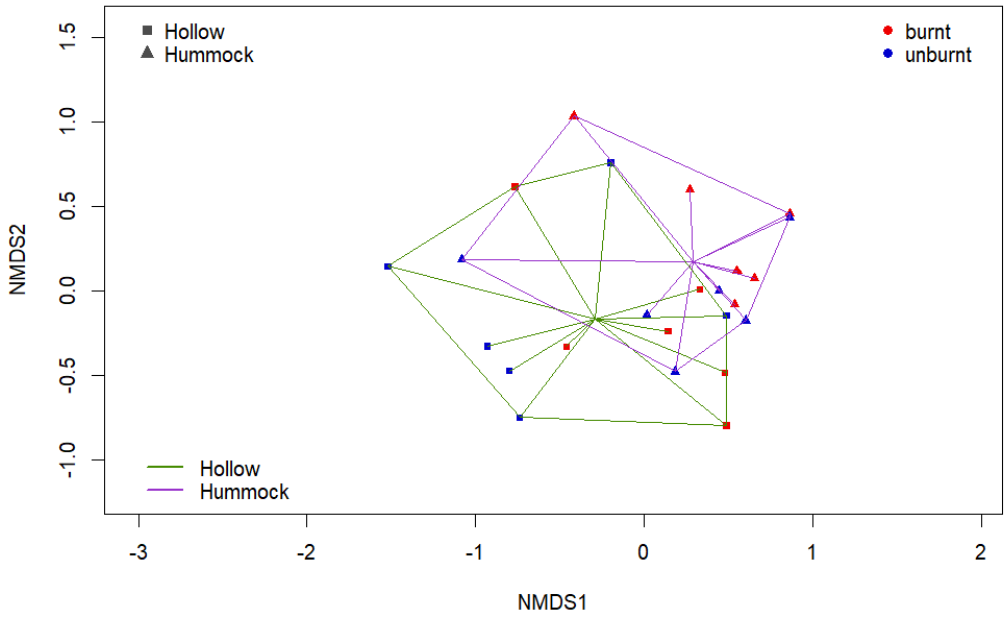



**Figure 2.12.** An NMDS plot showing clustering of the photosynthetic microbial community by microtopography and burnt/ unburnt status. Burnt/ unburnt status is shown by colour and microtopography is shown by point shape. Centroids to microtopography (as the only significant variable: envfit goodness of fit $r^2 = 0.1734$, $p = 0.023$) are included as spiders.

### 2.3.4 Indicator species analysis

Indicator species analysis showed no statistically significant indicator species for burnt vs. unburnt sites. There were 6 statistically significant indicator species for both hummocks (Table 2.2.) and hollows (Table 2.3.). Of these, all of the highlighted indicator species for hollows were either diatoms or desmids. Indicator species for hummocks tended to be smaller and simpler cells (thus harder to identify to species). In general, hummock indicators had higher IndVal scores, suggesting that those species have greater predictive power.

**Table 2.2.** Hummock microtopography indicator species as identified by indicator species analysis. The IndVal (indicator value) is made up of specificity and fidelity. See Section 2.2.7 for more details.

| Species | IndVal | A (specificity) | B (fidelity) | p-value |
|---|---|---|---|---|
| single_1 | 0.916 | 0.8391 | 1.0000 | 0.003 |
| single_23 | 0.891 | 0.9537 | 0.8333 | 0.001 |
| euglena_3 | 0.866 | 0.8991 | 0.8333 | 0.001 |
| euglena_2 | 0.826 | 0.8178 | 0.8333 | 0.008 |
| colony_4 | 0.798 | 0.8501 | 0.7500 | 0.015 |
| single_26 | 0.764 | 1.0000 | 0.5833 | 0.006 |

**Table 2.3.** Hollow microtopography indicator species as identified by indicator species analysis. See Section 2.2.7 for more details.

| Species | IndVal | A (specificity) | B (fidelity) | p-value |
|---|---|---|---|---|
| *Cylindrocystis* spp. | 0.891 | 0.9517 | 0.8333 | 0.001 |
| diatom_type_1 | 0.810 | 0.9844 | 0.6667 | 0.003 |
| diatom_type_2 | 0.738 | 0.9336 | 0.5833 | 0.013 |
| diatom_4 | 0.694 | 0.9645 | 0.5000 | 0.015 |
| diatom_7 | 0.637 | 0.9730 | 0.4167 | 0.037 |
| Desmid sp. (S22) | 0.619 | 0.9209 | 0.4167 | 0.044 |

## 2.4 Discussion

Peatland microbial surface communities show a surprising resilience to the effect of fire, with no significant changes in diversity or in community composition. This suggests that peatland microbial processes mediated by surface microbial communities may be resilient to fire. Ecosystem functioning could either be unaffected, or microbial resilience could prompt a more rapid return to peatland ecosystem functioning. *Sphagnum* was sampled where possible, so if severe fire resulted in complete loss of *Sphagnum*, this might be different. Indeed, as *Sphagnum* provides the moisture and low pH which are characteristic of peatland microbial niches, it would be expected that a complete loss of moss would cause more distinctive damage. However, this work demonstrates that the action of fire (heat, scorch, etc.) does not result in a significant difference in the total photosynthetic microbial community where moss is not totally destroyed.

Field sampling was carried out two months after the fire damage, presenting the possibility that the photosynthetic microbial community had undergone rapid recovery rather than persisting unaffected by fire. Previous studies on other microbial groups generally show long recovery times following fire (Meiczan *et al.*, 2022; Holden *et al.*, 2015; Ferrenberg *et al.*, 2013). However, it is perhaps possible that algae, adapted to the moss surface which can have large variance in moisture levels over a season (Küttim *et al.*, 2020) show this resilience due to evolutionary strategies to survive periods of drying (e.g. encysting mechanisms) allowing a certain amount of fire resilience. This could be particularly true for hummock species, which showed no changes despite significantly higher fire damage, and which are routinely at greater risk of drying than those found more frequently in wetter microtopographic environments. Indeed, it has been suggested that this is one of the main ways in which the community assemblage of algae on hummocks is determined (Goodyer, 2014). This possibility of rapid recovery could be investigating with a time series study beginning immediately after fire damage; unfortunately, for this thesis both immediate sampling following the fire and repeated measures through 2020 were not possible.

Care must be taken when interpreting these results, as it is likely they cannot be generalised across all UK peatlands. Conditions in Forsinard may be different to those found elsewhere (for example, this sampling area around a pool system is extremely wet), and it is likely that a drier peatland would have different patterns of damage. Where fire causes complete surface peat loss, microbial persistence is much less likely.

### 2.4.1 Importance of microtopography for microbes

The importance of microtopography for microbial communities has been demonstrated previously (e.g. Goodyer, 2014). It is expected that fire would impact on the microbial community if fire causes microtopography to be lost, as can occasionally happen in high intensity fires (Nelson *et al.*, 2021). Therefore, post fire restoration for photosynthetic microbes should focus on restoring a natural range of microtopographies, with hummocks and hollows providing different microbial niches. This could also be applied to restoration following other damage as well as fire. The reintroduction of hummock forming *Sphagnum* species, or even small scale surface level manipulation to introduce a range of microtopography, may be helpful.

This reinforces the importance of considering microtopography at a relatively small scale for studies of peatland microbes in the surface layer. Most of the hummocks and hollows in this study area were <1m$^2$. Experimental designs for work on surface dwelling peatland microbial communities should consider microtopography, and caution should be used when making conclusions about peatlands as a whole if microtopographic variance is not considered.

The impact of microtopography on microbial communities, and the interaction with fire, is likely due to hydrology- this is a wet area of bog, near a substantial pool system. Extreme caution should be used before applying these findings to other UK peatlands, especially those which are drier, either due to management or rainfall envelope. However, this does provide an indication that restoration towards a wetter system may prevent fires from causing significant changes to the surface microbial community.

*Sphagnum* moss provides a distinct niche for the microbial community, and sampling was carried out of the most abundant *Sphagnum* species in each quadrat. The majority

of *Sphagnum* hummocks were composed of *Sphagnum capillifolium*, which could have determined the differences between hummock and hollow, which were dominated by different and more varied *Sphagnum* species. Microbial specificity to *Sphagnum* species has been previously demonstrated (Bragina *et al.*, 2012; Opelt *et al.*, 2007). However, even if the difference between hummock and hollow algal communities is largely driven by the dominant *Sphagnum* species in each microtopography and the microbial microhabitat conditions that they create, this is still an impact of microtopography.

### 2.4.2 Microtopographic and hydrological variation in fire

These results emphasise the importance of considering scale in experimental design for post fire studies and monitoring. *Sphagnum* hollows had significantly lower fire damage than hummocks, likely due to raised water table hampering ignition (and perhaps providing thermal insulation). The same pattern was also recorded in a separate study on the vegetation impacts of the same fire (Andersen *et al.*, 2024). This is in contrast to studies which show hummocks as more resistant to fire than lower microtopographies (e.g. Nelson *et al.*, 2021). This may be because of different *Sphagnum* species having different levels of susceptibility to fire; the species in Nelson *et al.* (2021) were not found in this study. Alternatively, site-level differences may be responsible, for example the high water table and proximity of pool systems in this area of the Forsinard Flows. It is therefore important for future studies to account for this, as fire impacts on a microtopographically varied site will likely not be uniform, even at the scale of <1m. This recommends an experimental design which considers ways to systematically account for this small-scale variance in microtopography.

### 2.4.3 Indicator species

Indicator species analysis showed no significant candidate indicators for fire damage, reinforcing the similarity of burnt and unburnt communities. However, it did suggest candidate species for hummocks vs. hollows. All six significant indicators of hollows were desmids or diatoms; this is unsurprising as they are known to be especially abundant in wetter peatlands, including pool systems. All indicator species for

hummocks were smaller and simpler cells. These would not always be included in microscope studies, which tend to focus on one or a small number of taxonomic groups. Identification of these types of microbes is difficult with light microscopy, and identification is generally only possible to morphotype rather than species. In many cases, genetic analysis has replaced microscopy for these more cryptic species; however, cost constraints mean that this is not always possible. These results suggest that morphotypes can be useful as indicators with statistical significance, even if species cannot be identified with confidence. This is especially useful for hummocks and other, drier, areas of peatlands, where diatoms, desmids and other more traditional indicators are generally only found at very low abundance. When a target count number approach is used, sampling across a wider range of taxonomic groups can also reduce the time required to locate individuals of the taxa of interest.

As hollows and hummocks can be easily visually distinguished, the value of this may seem limited. However, when there is no set definition of the different microtopographies, and peatland water table is variable, the microalgal communities may provide an indication of how hydrological conditions on sampling days relate to longer term water table levels. Monitoring of microalgae over time would be necessary to see how stable these communities are to short term changes in water table. They may also allow us to test if smaller scale and less visually apparent microtopographical variation has the same potential to increase beta diversity within the algal community.

## 2.5 Conclusions

In a near natural blanket bog, surface communities of photosynthetic microorganisms can show surprising resilience to fire where combustion is incomplete and bryophyte material survives.

Healthy peat bogs have complex microtopography, encompassing hummocks, lawns and hollows. These provide a range of niches for microbes, with different vegetation, moisture and shading all likely to be important drivers of the microbial community. They are also affected differently by fires, which are predicted to pose an increasing threat to peatlands with climate change. Therefore, to fully understand the effects of fire on

peatland microbes, we need to understand their responses and potential recovery at a scale which accounts for these microtopographic landforms.

The majority of the evidence for peatland restoration reducing fire risk and severity in the UK is largely anecdotal (Andersen *et al*., 2024). Therefore, any data on the response of UK peatlands to fires can inform planning and management. In this case, managing to maintain (or increase where appropriate) peatland moisture levels and *Sphagnum* cover, and to ensure that microtopographic variation is allowed to develop, are likely to both protect cryptic algal biodiversity and also confer some resilience to wildfire.

# 3 Bioinformatics method development

## 3.1 Introduction

Chapters 4 and 5 of this thesis involve microbial community data collected via DNA analysis. This requires the raw sequence data to be processed in order to produce community data which can be analysed. The bioinformatics methods used for this processing can have a significant influence on the end results of the analysis (e.g. Hleap *et al*., 2021; Alberdi *et al*., 2017), and therefore should be carefully considered.

Initially, bioinformatics was carried out using the DADA2 pipeline (Callahan *et al*., 2016) and Phyloseq (McMurdie and Holmes, 2013) in Rstudio (see Section 3.3), which is a standard approach for this type of amplicon sequencing. DADA2 clusters groups of reads into assigned sequence variants (ASVs) rather than the previously more common operational taxonomic units (OTUs). Whereas OTUs are clustered according to a predetermined sequence similarity threshold (often 97%), ASVs are inferred without a similarity threshold and can separate sequence variants differing by a single nucleotide. This allows for improved resolution of clustering, error filtering and comparability between studies (Callahan, McMurdie and Holmes, 2017).

However, for the datasets in this thesis, the standard DADA2 pipeline resulted in very poor rates of taxonomic assignment- that is, the identification of DNA sequences into taxonomic classifications which allow both meaningful ecological analysis and, perhaps more importantly, non-target taxa to be filtered out (see Section 3.3.3). This is likely due to poor reference database performance. Therefore, a new approach to bioinformatics was required.

A phylogenetic approach was initially considered (see Section 3.4), before the decision was made to combine the DADA2 pipeline with Kraken2 (Wood, Lu and Langmead, 2019), a different method of taxonomic assignment used more frequently in metagenomics. Section 3.5 discusses the development of this methodological pipeline and the methodological testing of its performance. Finally, Section 3.6 describes *in silico* testing of primer complementarity.

## 3.2 Laboratory work

### 3.2.1 Primer choice

Because photosynthesis in microorganisms is not a taxonomically conserved trait (Fehling, Stoecker and Baldauf, 2007), the use of more common primers such as 18S/16S was considered to present too many challenges in ascribing photosynthetic capability, especially for taxa which are less well studied. rbcLa was therefore selected as a functionally significant photosynthetic gene for primer targeting (Table 3.1). rbcL primers (Kress and Erickson, 2007) have been used to barcode communities for both plants and algae (Seymour *et al*., 2020; An, Choi and Noh, 2020; An *et al*., 2018). It has been shown to be an effective primer for diatoms, with a good relationship between rbcL metabarcoding and traditional identification via light microscopy (Pérez-Burillo *et al*., 2020). It also shows better discrimination at genus level of some microalgae than standard ribosome-amplicon methods (Ballesteros *et al*., 2021). Shared endosymbiont origin of chloroplasts for many algae early in eukaryotic evolution (Yoon *et al*., 2004) means that a reasonable amount of conservation can be expected across all algae.

**Table 3.1.** Primer sequences (both raw primers and with Illumina tagging) used in PCR reactions.

| **Primer** | **Sequence 5' - 3'** |
|---|---|
| rbcLa- F | ATGTCACCACAAACAGAGACTAAAGC |
| rbcLa - R | GTAAAATCAAGTCCACCRCG |
| rbcLa- F (Illumina tagged) | TCGTCGGCAGCGTCAGATGTGTATAAGAG ACAGANNNHNNNWNNNHATGTCACCAC AAACAGAGACTAAAGC |
| rbcLa - R (Illumina tagged) | GTCTCGTGGGCTCGGAGATGTGTATAAGA GACAGTGTAAAATCAAGTCCACCRCG |

### 3.2.2 PCR reactions

PCRs were carried out following the protocol described in Fazekas *et al*. (2012), as described in Table 3.2.

**Table 3.2.** PCR cycle used for rbcL primers.

| **Stage** | **Temperature (°C)** | **Time** |
|---|---|---|
| Initial denaturing | 94 | 4 minutes |
| 35 cycles of: | | |
| Denaturing | 94 | 30 seconds |
| Annealing | 55 | 30 seconds |
| Extension | 72 | 1 minute |
| Final extension | 72 | 10 minutes |

Amplification with 18S TAREuk primers (Stoeck *et al*., 2010) was attempted but was unsuccessful. Interestingly the same primer sets also failed in Burn (2021), suggesting possible methodological issues when used for peatland samples. An accidental repeat PCR showed a band of correct length, suggesting the possibility that low DNA yield for those primers could be the cause.

*Sphagnum* represented the majority of the rbcL amplicon reads found (see Section 3.5.2), though TAREuk would also be expected to amplify other DNA which was not amplified by rbcL, such as invertebrates, protists and other non-photosynthetic microeukaryotes, as well as the non-*Sphagnum* plant and microbial sequences which were detected with rbcL. This raises the possibility of other issues, such as a particular

sensitivity of PCR with these primers to inhibitors found in peatlands and *Sphagnum*. Complex organic acids are abundant in peatlands (Rydin and Jeglum, 2006) and can act as inhibitors to PCRs (Schrader *et al.*, 2012; Matheson *et al.*, 2010). Despite use of BLOTTO (Bovine Lacto Transfer Technique Optimiser; a solution of fat-free milk powder in water which can help to reduce PCR inhibition (De Boer *et al.*, 1995)) to prevent this, both in this study and Burn (2021), inhibition of PCR may still be the cause of failure.

### 3.2.3 Sequencing

Sequencing was carried out by the Biology Technology Facility, University of York. In preparation for sequencing, amplicons were amplified and barcoded using an 8 cycle PCR with NEB Q5 high fidelity mastermix and the Illumina Nextera XT indexing primers, then cleaned with AMPure XP beads (Beckman Coulter, USA), quantified using the invitrogen Quant-it kits and pooled at approximately equimolar ratios. Dual end 250 bp sequencing was performed on an Illumina NovaSeq instrument.

## 3.3 Original bioinformatics method

Bioinformatics was initially carried out using DADA2 (Callahan *et al.*, 2016) and Phyloseq (McMurdie and Holmes, 2013) in R (version 4.1.2) in Rstudio, following a pipeline adapted from Burn (2021). Nextera adapters were removed with cutadapt version 4.6 (Martin, 2011). The expected insert size was 599bp. As forward and reverse sequences did not have sufficient overlap to merge, sequences were treated as single-end throughout.

Reads were trimmed using DADA2 to remove primers, and truncated to a length of 210bp following visual inspection of quality plots. Reads were also truncated at the first instance of quality score less than 2, and after truncation reads with higher than 2 expected error were discarded. Error rate was calculated, dereplication carried out and chimeras removed. Taxonomy and species names were assigned using a custom reference database (see Section 3.3.1).

Assigned sequence variants (ASVs) were inferred using DADA2 (analogous to clustering sequences). Samples with fewer than 40 reads were removed, and data rarefied to even sequencing depth.

### 3.3.1 Reference database preparation

Whilst rbcL is frequently used as a primer for metabarcoding approaches, it is usually used to focus on a few taxa rather than in studies encompassing all algal groups. Therefore, no reference databases were available and a custom one had to be created. This combined data from two existing curated databases with publicly available data downloaded from NCBI (Fig. 3.1).

A curated database exists for diatoms (Bacillariophyceae). The Rsyst::diatom database diatbarcode v. 11.1 (Rimet *et al.*, 2018) was used, with formatting adapted for use in DADA2. This provided the framework for the taxonomic levels to be used for the custom database as a whole. Taxonomy of the Rsyst::diatom database had been checked against Algaebase during curation (ibid) and so was considered to be sound.

Plant data was taken from the taxonomic, genetic and ecological data resource for the vascular plants of Britain and Ireland (Henniges *et al.*, 2021). This curated database was trimmed to remove extinct species and species which had no associated rbcL sequence. This was converted to the format required for DADA2, with higher taxonomic levels added from the wider literature.

These databases were supplemented with data downloaded from NCBI. The NCBI gene database (accessed December 2023) was searched using Geneious Prime (version 2023.2.1) for records where the gene name included rbcL. This generated 14454 records. These were then filtered in R to remove diatoms, to prevent redundancy between different datasets. Viridiplantae sequences were then separated out. Viridiplantae were filtered to remove the Embryophyta (land plants), with all other Viridiplantae taxa (various groups of green algae) added back to the database. As the plant database only contained vascular plants, the Spermatophyta were separated from the NCBI vascular plant reads, and also added to the database (Fig. 3.1).

Taxonomy for each algal record from NCBI was manually checked against Algaebase (Guiry, M.D. & Guiry, G.M., accessed December 2023). Taxonomy for non-algal NCBI records was visually checked for consistency based on botanical and microbial taxonomic naming conventions. Six NCBI records were filtered out due to limited taxonomic information. The three databases were combined, formatted for DADA2 and converted to FASTA format.

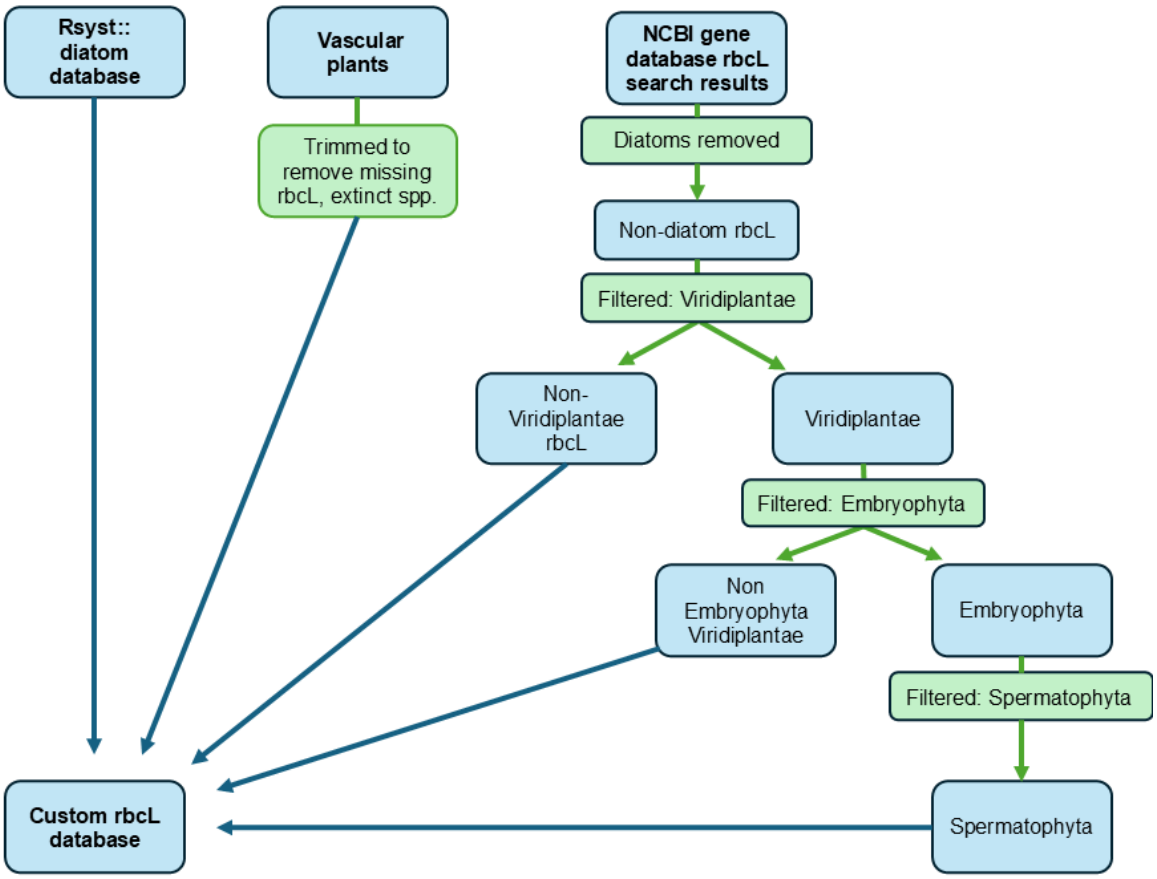



**Figure 3.1.** A flowchart showing custom reference database construction for analysis of rbcL amplicons with DADA2. Blue boxes show databases or subsets of databases. Green boxes show filtering or trimming steps.

### 3.3.2 Data processing

Data were filtered to remove all ASVs identified as plants (vascular and non-vascular) before further analysis. Whilst ASV identification is not perfect, the reference database included almost all UK vascular plants, so assignment for these was likely to be accurate. There was more potential for incorrect ASV assignment for non-vascular plants, which were less well represented in the database.

### 3.3.3 Level of taxonomic identification

The original bioinformatics approach, with taxonomic assignment using the custom rbcL database and taxonomic assignment with DADA2 in R, yielded very low proportions of ASVs with any level of taxonomic identification.

The majority of assigned sequence variants were unclassified at phylum level (Fig. 3.2).

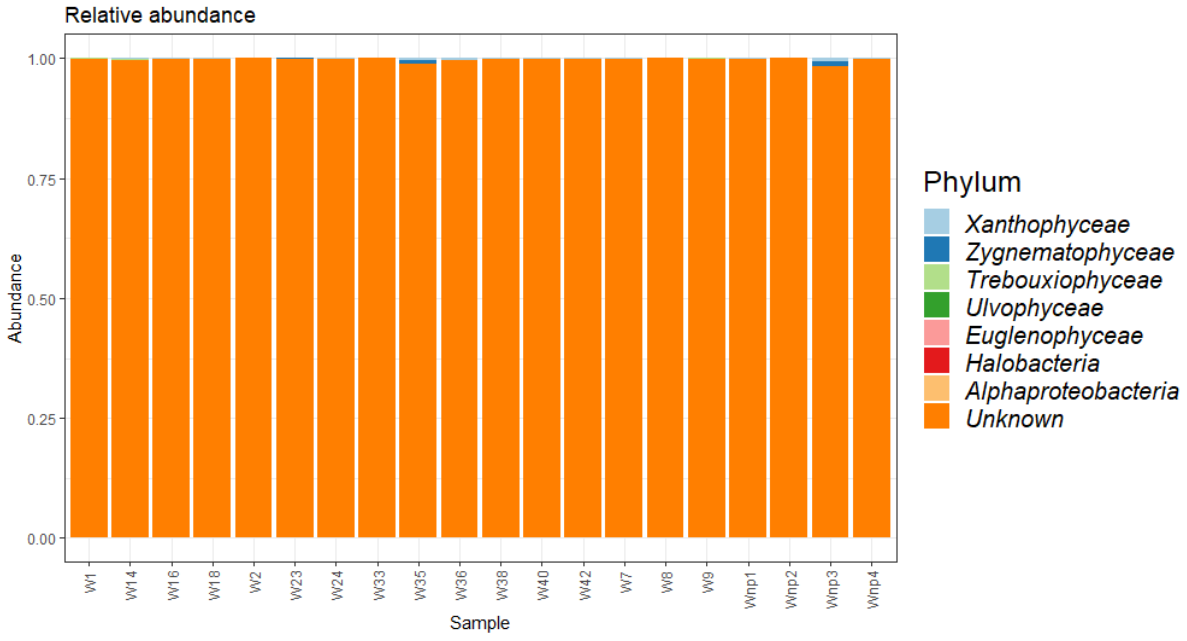



**Figure 3.2.** Relative abundance of dominant microbial phyla, as identified with taxonomic assignment using the custom rbcL database following the DADA2 pipeline, with no filtering for ASVs unclassified at the phylum level. Data taken from Chapter 4.

In order to better see any patterns, samples unclassified at phylum level were filtered out. However, even with ASVs unclassified to phylum removed, a large proportion remained unclassified at 'class' level (Fig. 3.3).

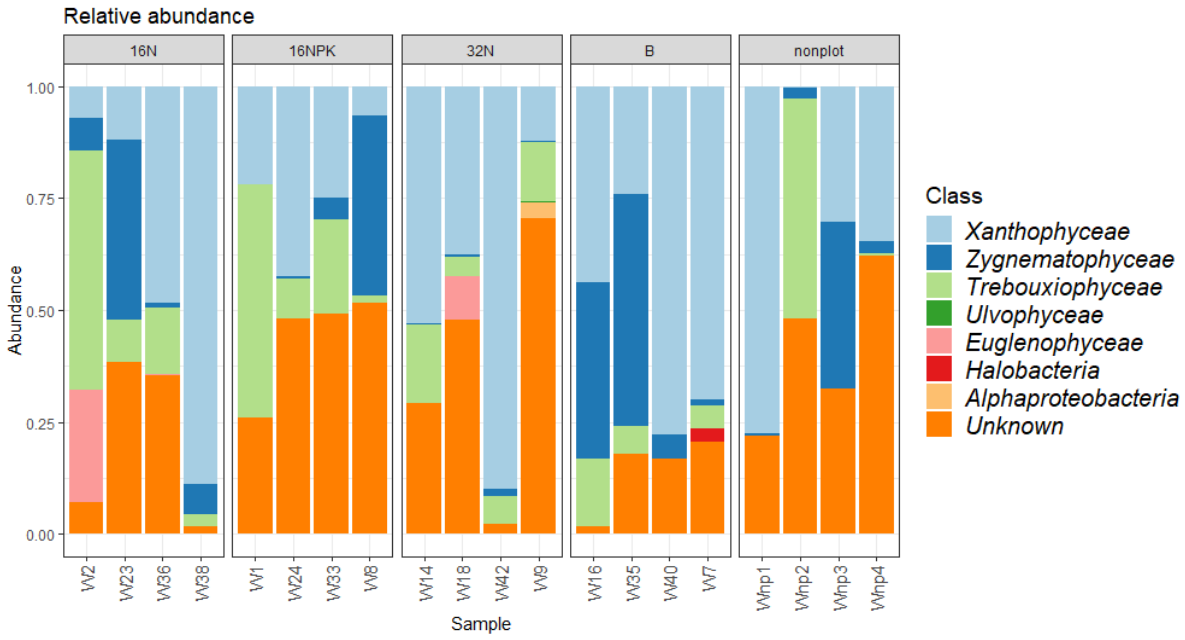



**Figure 3.3.** Relative abundance of dominant microbial classes, as identified with taxonomic assignment using the custom rbcL database following the DADA2 pipeline, after filtering to remove ASVs unclassified to phylum level. Many ASVs remain unclassified to class level. Data taken from Chapter 4, faceted by nutrient treatment.

Poor levels of taxonomic identification made it challenging to infer any changes in taxonomic groups caused by treatments. Therefore, the decision was made to attempt other methods to increase the proportion of ASVs with assigned taxonomic information.

## 3.4 Benefits and issues with a phylogenetics approach

Initially, a phylogenetic approach was selected as being likely to improve taxonomic identification. However, reading the literature suggested that there can be problems with the robustness of using a *de novo* approach to phylogenetic tree building (Czech *et al.*, 2022; Dunthorn *et al.*, 2014; Moret, Roshan and Warnow, 2002). This is worse for the short reads produced by metabarcoding, as shorter sequence lengths (especially when combined with many ASVs) have reduced phylogenetic signal, resulting in a less robust

phylogenetic tree than would be possible with longer sequences such as full 16S reads (Czech *et al*., 2022). A phylogenetic placement approach can be useful in these cases. Phylogenetic placement is the mapping of the unknown sample sequences onto a fixed reference tree, by calculating which branch of the reference tree a sample sequence is most closely related to (ibid). Following trimming for quality, read lengths used in the analysis for this thesis were very short at 138bp, and therefore any conventional phylogenetic tree is likely not to be robust.

However, whilst some reference trees for microbial eukaryotes do exist, none that can be found in the literature are suitable for the full range eukaryotic microalgae, or plastid primers. For example, Berney *et al*. (2017) and Del Campo *et al*. (2018) are both protist only reference trees, and Rajter and Dunthorn (2021) and Rajter *et al*. (2021) have reference trees for only ciliates. All of these reference trees also use ribosomal genes rather than plastid genes. Therefore, phylogenetic placement approaches would only be possible following building my own reference tree. This in turn would require significant time spent curating my custom rbcL database, and would be computationally intensive. Therefore, a different approach to improving the level of taxonomic identification was required.

## 3.5 New bioinformatics method- Kraken2

Kraken2 (Wood *et al*., 2019) is a taxonomic assignment software designed for metagenomics, which rapidly and accurately carries out classifications using exact alignment of kmers. Kraken2 can handle unassembled sequences (Valenzuela-González *et al*., 2016), and should therefore not be limited to only metagenomics uses. Kraken2 can be used for metabarcoding amplicon sequencing approaches (e.g. Odom *et al*., 2023; Mlaga *et al*., 2021).

Kraken takes a LCA (last taxonomic common ancestor) approach, which is useful here as a species level classification is unlikely to be achievable for all taxa considering gaps in the reference databases. The LCA approach may allow for taxa unidentifiable at the species level to still be identified at a higher taxonomic level, which is still useful for analysis (Hleap *et al*., 2021). It will also leave sequences unclassified if they do not

meet its criteria for classification (Wood and Salzberg, 2014). Some authors have found the accuracy of Kraken2 to be lower than some alternatives (e.g. Pipes and Nielson, 2024), however I judged that the ability to use a very large database to classify taxa with poor representation in databases outweighed this negative, especially where identification could be agglomerated to higher taxonomic levels without causing issues.

Kraken2 outputs data in a format showing taxonomic hierarchies. However, the data can be extracted into more traditional taxon abundance tables at the taxonomic level of choice using Bracken (Bayesian Reestimation of Abundance with KrakEN), which estimates the number of reads from each species or taxon at a higher taxonomic level (Lu *et al.*, 2017).

### 3.5.1 Database

Kraken2 was run using the NCBI nt database nt_20240530 from 30th May 2024. This was downloaded from https://genome-idx.s3.amazonaws.com/kraken/k2_nt_20240530.tar.gz on 2024-08-20 using Wget. This is a very large collection, including the GenBank, RefSeq, TPA and PDB databases, with an archive size of 684GB and an index size of 889GB.

### 3.5.2 Methodological testing

An initial test of Kraken2 in quick mode with raw sequence data post demultiplexing but prior to adaptor cleaning and primer removal was carried out on a single sample from Chapter 4, selected at random. This was done to check if the software would be able to use this style of data and would run without failure or error, as it was originally designed for metagenomics datasets.

Whilst Kraken has been used in the literature on metabarcoding data rather than the metagenomic data for which it was originally designed (e.g. Hleap *et al.*, 2021) there were no clear examples demonstrating how to include it in an existing pipeline. Therefore, we carried out a series of tests to attempt to find the best way to combine taxonomic identification using Kraken2 with data cleaning via the DADA2 pipeline. Testing was carried out using Whim sample 14. A sample from Whim (see Chapter 4 for

details of sampling) was chosen for ease of testing as all of these samples were taken from *Sphagnum capillifolium*, so host species would be known, and the number of non-*Sphagnum* plants should be more limited than in the mixed-host samples from Chapter 5. Sample number 14 was selected at random.

#### *3.5.2.1 Kraken test 1- DADA2 cleaning, no dereplication or primer removal*

Kraken2 was run in quick mode on data which had undergone primer removal, adapter trimming and quality trimming using the DADA2 pipeline (see section 3.1.4 for details). The data was removed from the DADA2 pipeline before dereplication, and therefore before any chimera removal could take place. This is because dereplication changes the file format in DADA2, so that it cannot be input into Kraken2 without manual editing. The resulting Kraken output showed very low proportions of reads with unassigned taxonomy. This improvement is unsurprising, as primers, adapters and low-quality reads would be expected to result in sequences which could not be identified in the database. However, the vast majority of sequences were *Sphagnum* spp.. This is not unexpected, as even with filtering steps the method of DNA extraction (especially with prior freezing) would likely yield very high levels of host DNA.

Of the identified sequences from this output, some seemed likely to be spurious, and may be of chimeric origin. Chimeras occur when non-pair fragments of DNA anneal during PCR reactions, and are a known source of methodological bias in metabarcoding (Haas *et al*., 2011; Kopczynski *et al*., 1994), which can potentially lead to falsely inflated diversity estimates (Kunin *et al*., 2010). Chimera removal has been shown to increase the precision of Kraken for metabarcoding sequences, though does also affect the sensitivity (Mlaga *et al*., 2021).

#### *3.5.2.2 Kraken test 2- DADA2 cleaning, clustering and chimera removal in Vsearch*

Chimera removal was attempted on the same data as Kraken test 1 (which had undergone primer removal, adapter trimming and quality trimming in the DADA2 pipeline) using Vsearch (Rognes *et al*. 2016), but was extremely computationally intensive without dereplication or clustering. For the test sample, chimera removal in

Vsearch without dereplication resulted in no chimeric sequences being detected or removed. Clustering at 97% similarity using Vsearch allowed chimera removal to be carried out. However, when this data was run through Kraken, abundance counts were lost from outputs, which only gave the number of clusters with an ID, not the number of reads within each cluster (Fig. 3.4). Therefore, relative abundance could not be calculated from this data. As percentages of taxa were calculated from the number of clusters, these were very different to the percentages returned by other methods (Table 3.3).

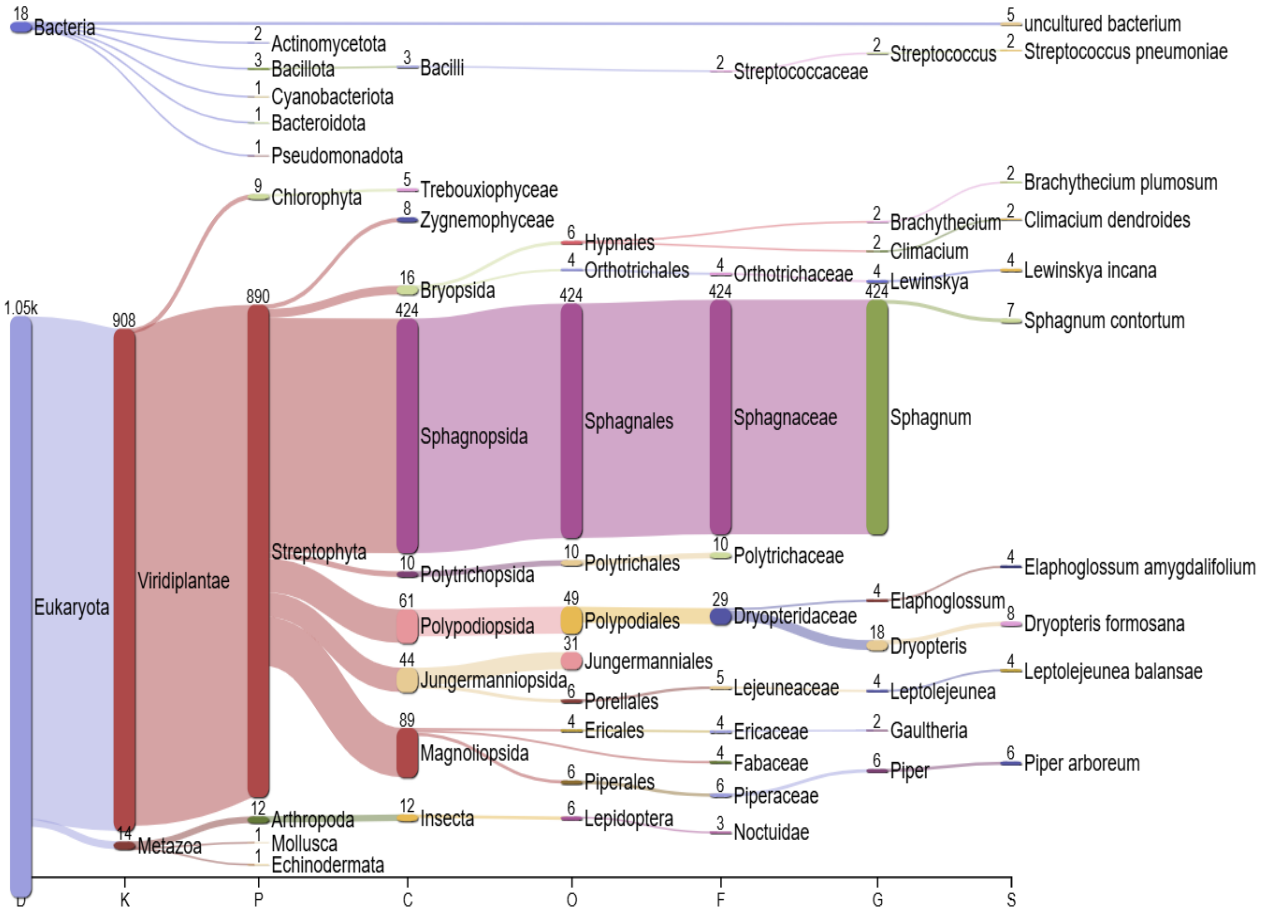



**Figure 3.4.** A Sankey visualisation of the taxonomy assigned by Kraken for samples with data cleaning and trimming carried out following the DADA2 pipeline, then clustering at 97% and chimera removal carried out in Vsearch (Kraken test 2). Numbers above nodes should show the read numbers assigned to this taxonomic grouping, however instead show the number of ASVs (clusters) assigned to the taxonomy, as read count data has

been lost during clustering. Metazoan samples are likely to be errors in classification or in the database (e.g. symbionts recorded as host DNA). Some species level assignments are also clearly spurious, e.g. *Piper arboreum*, tree pepper. The Sankey visualisation was created using Pavian (Breitwieser and Salzberg, 2020).

#### *3.5.2.3 Kraken test 3- DADA2 cleaning, chimera removal and sample inference, quick mode*

Another attempt was made using the output from the DADA2 pipeline, after dereplication, chimera removal and sample inference (DADA clustering into ASVs). Following manual conversion of sequence abundance tables into fasta files, this was able to run successfully on Kraken2. However, the resulting data showed lower diversity than other Kraken tests (Figs. 3.5 - 3.7). There were also some spurious taxonomic assignments of vascular plants (for example in the Ericales, for which the only species assigned was *Burckella oxycarpa,* a tree species from Papua New Guinea). The reason for this difference is unclear, but likely to be because of Kraken2 quick mode, which stops searching after the first database hit rather than searching all k-mers in a sequence, possibly resulting in incorrect assignment to a related taxa.

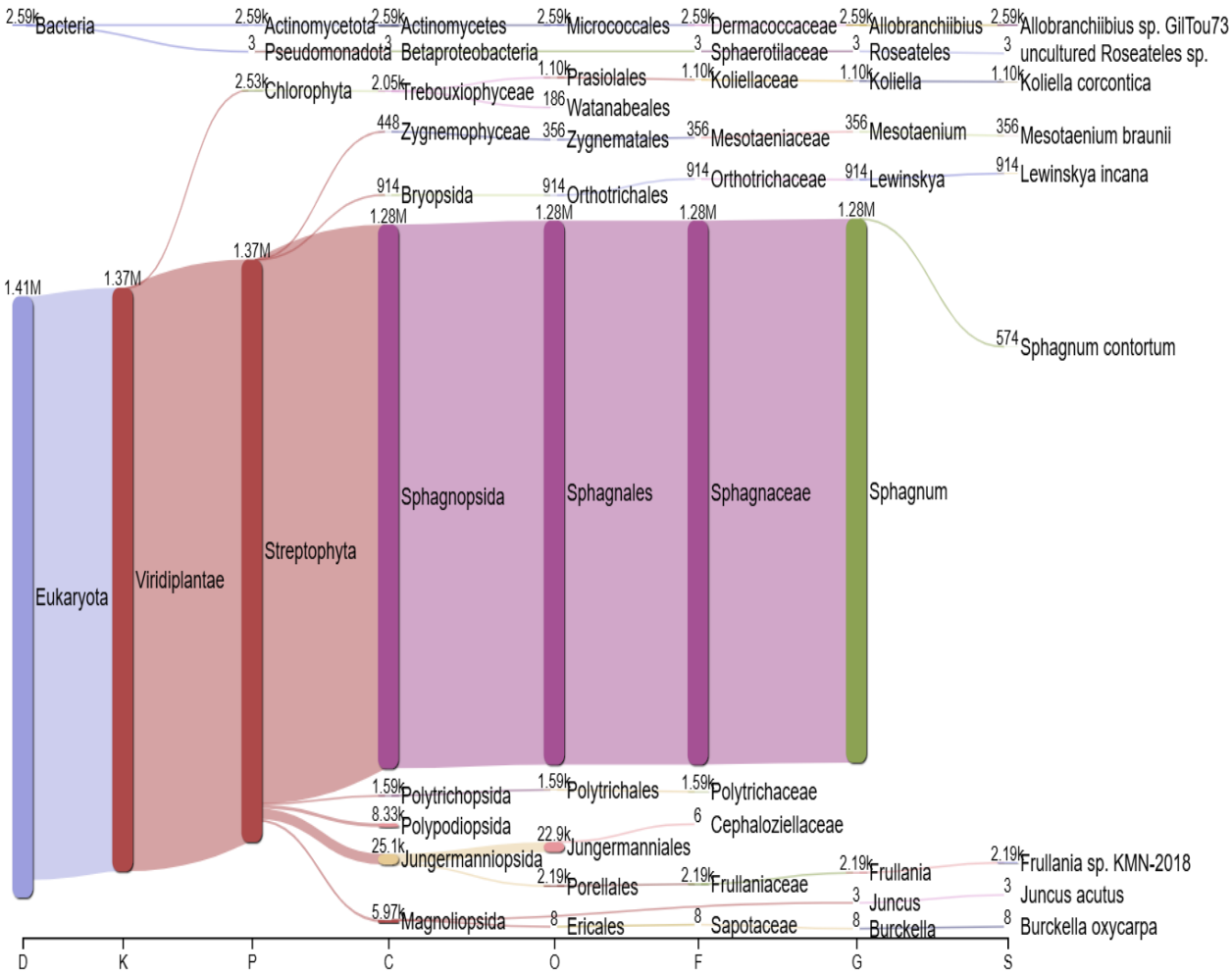



**Figure 3.5.** A Sankey visualisation of the taxonomy assigned by Kraken in quick mode for samples following data cleaning, filtering and clustering into ASVs using the DADA2 pipeline (Kraken test 3). Numbers above nodes show the number of reads assigned to this taxonomic grouping. The Sankey visualisation was created using Pavian.

### *3.5.2.4 Kraken test 4- DADA2 cleaning, primer and adaptor removal. No clustering or chimera removal.*

As Kraken2 allows per-read taxonomic assignments, ASV clustering prior to taxonomic assignment should not be biologically necessary. High levels of taxonomic assignment for data when cleaned using DADA2 before chimera removal/ sample inference mean there would be few unassigned ASVs (Table 3.3), and any ASVs with the same taxonomic assignment would need to be combined anyway. This is especially the case when not

looking at species level taxonomic assignments; for this data differences between higher taxonomic levels are more interesting, due to greater reliability even if species level assignments may not be correct. The risk is losing some cryptic diversity of ASVs not assigned to a taxonomic level, but identification rates are high enough (>99%, Table 3.3) that this could be judged a worthwhile trade-off. Chimera removal was impractically time consuming and computationally intensive without either dereplicating or clustering, but simply removing singletons/doubletons, or anything with fewer than 10 reads across all samples should result in similarly good denoising (Alberdi *et al.*, 2017).

This approach was tested with Kraken2 with quick mode turned off. This still resulted in spurious matches (e.g. *Burckella oxycarpa* and *Piper arboreum* in the Ericales), and still yielded low levels of microbial biodiversity (Fig. 3.6). Therefore, as the available data from sequencing is likely to be limited, the choice was made to aim for the maximum accuracy possible and use DADA2 for the whole process, run in standard rather than quick mode.

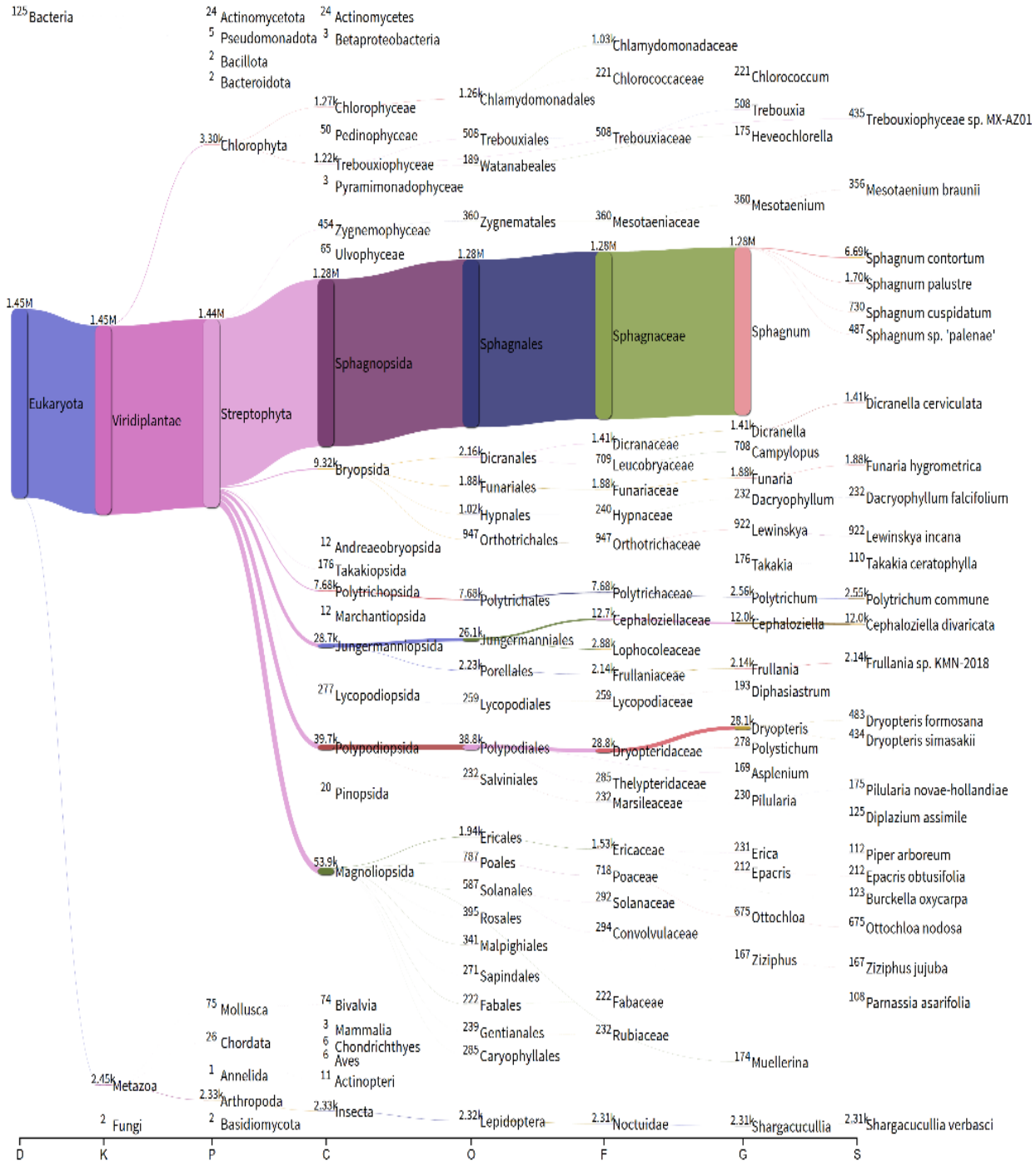



**Figure 3.6.** A Sankey visualisation of the taxonomy assigned by Kraken with quick mode turned off for samples following data cleaning and filtering using the DADA2 pipeline, with no clustering or chimera removal (Kraken test 4). Numbers above nodes show the number of reads assigned to this taxonomic grouping. The Sankey visualisation was created using Pavian.

#### *3.5.2.5 Kraken - Final run*

For the final data set, Kraken2 was run on samples which had been through the full DADA2 pipeline. This included primer removal, adapter trimming and quality trimming, before dereplication, chimera removal and sample inference (clustering to ASVs). Data was then manually converted to FASTA format. FASTA files were run in Kraken2 without quick mode (Fig. 3.7). This is equivalent to Kraken test 3 without quick mode. Compared to when run with quick mode on, this resulted in a slightly higher percentage of reads assigned to Chlorophyta (one of the green algae groups of interest) and a lower percentage of reads assigned to prokaryotes (Table 3.3). This lack of Prokarya is not concerning, as *in silico* PCRs suggested that amplification of cyanobacteria would likely be poor with the primer set used (see Section 3.6).

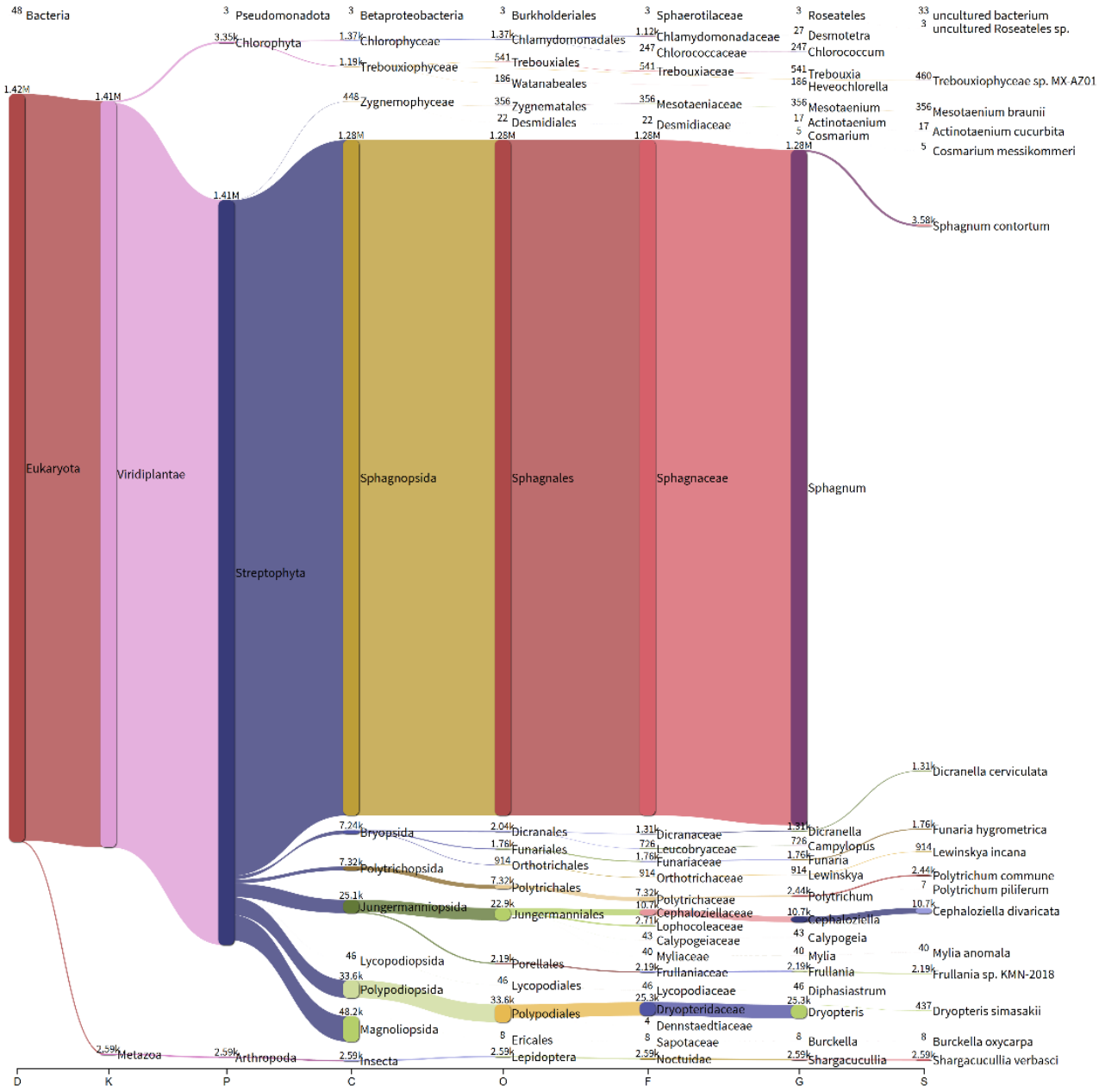



**Figure 3.7.** A Sankey visualisation of the taxonomy assigned by Kraken with quick mode turned off on samples cleaned, trimmed and clustered using the DADA2 pipeline (final Kraken run). Numbers above nodes show the number of reads assigned to this taxonomic grouping. The Sankey visualisation was created using Pavian.

**Table 3.3.** Percentage of reads unclassified and assigned to different taxonomic groups from Kraken2 methodological testing runs (sample W14 from Chapter 4, selected at random).

| Kraken run | % reads unclassified | % reads Eukaryota | % reads Prokaryota | % reads *Sphagnum* spp. | % reads Chlorophyta |
|---|---|---|---|---|---|
| **Kraken test 2** | 4.92 | 89.32 | 1.53 | 35.93 | 0.76 |
| **Kraken test 3** | 0.10 | 99.71 | 0.18 | 90.21 | 0.18 |
| **Kraken test 4** | 0.15 | 99.83 | 0.01 | 88.45 | 0.23 |
| **Final run** | 0.10 | 99.90 | 0.00 (48 reads) | 90.21 | 0.24 |

As the *Sphagnum* species which these samples were taken from was known to be *S. capillifolium*, this provides another method of assessing these results (Table 3.4). All test runs had the same percentage of reads assigned to Sphagnaceae and *Sphagnum*, meaning that all reads assigned to Sphagnaceae were also assigned to *Sphagnum*. However, a very low percentage of these reads (<1% of total read count) was assigned to the species level. Kraken test 3 and the final run (both of which involved cleaning and ASV clustering with the DADA2 pipeline) only had reads assigned to *S. contortum* at the species level. Kraken test 2 also included the unknown species *Sphagnum* sp. *'palenae'*. Kraken test 4 (unclustered, with no chimera removal) showed the most diversity, with reads assigned to five *Sphagnum* species (*S. contortum, S. palustre var. imbricatum, S. cuspidatum*, *Sphagnum* sp. *'palenae'* and *S. ovatum*). The increased diversity is likely to be due to the lack of clustering or ASV inference.

It is notable that, for none of the methods, Kraken assigned any reads to the correct species, *S. capillifolium*. This highlights the need for caution with interpretation of these results, and reinforces my decision to discount species level data and take only higher taxonomic levels through for further analysis.

**Table 3.4.** Taxonomic outputs for the Sphagnaceae from Kraken2 methodological testing runs.

| Kraken run | Kraken taxonomic output for Sphagnaceae |
|---|---|
| **Kraken test 2** | 35.93% Sphagnaceae<br>35.93% *Sphagnum*<br>0.59% *Sphagnum contortum*<br>0.08% unclassified *Sphagnum*<br>0.08% *Sphagnum* sp. *'palenae'* |
| **Kraken test 3** | 90.21% Sphagnaceae<br>90.21% *Sphagnum*<br>0.04% *Sphagnum contortum* |
| **Kraken test 4** | 88.45% Sphagnaceae<br>88.45% *Sphagnum*<br>0.46% *Sphagnum contortum*<br>0.12% *Sphagnum palustre*<br>0.12% *Sphagnum* palustre var. imbricatum<br>0.05% *Sphagnum cuspidatum*<br>0.03% unclassified *Sphagnum*<br>0.03% *Sphagnum* sp. *'palenae'*<br>0.00% (23 reads) *Sphagnum ovatum* |
| **Final run** | 90.21% Sphagnales<br>90.21% Sphagnaceae<br>90.21% *Sphagnum*<br>0.25% *Sphagnum contortum* |

The aim of the new bioinformatics approach was to improve taxonomic assignment rates to enable more biological conclusions to be drawn from the data. Across both Whim (Chapter 4) and Abernethy (Chapter 5) samples, all but two samples had fewer than 2.5% of reads unclassified, and the majority of samples had fewer than 1% of reads unclassified (Fig. 3.8). This is in contrast to the previous method, where the vast majority of non-plant reads were unclassified to the phylum level (Section 3.3.3, Fig. 3.2). Only one sample (A2) had higher than 10% of reads unclassified with the second bioinformatics method.

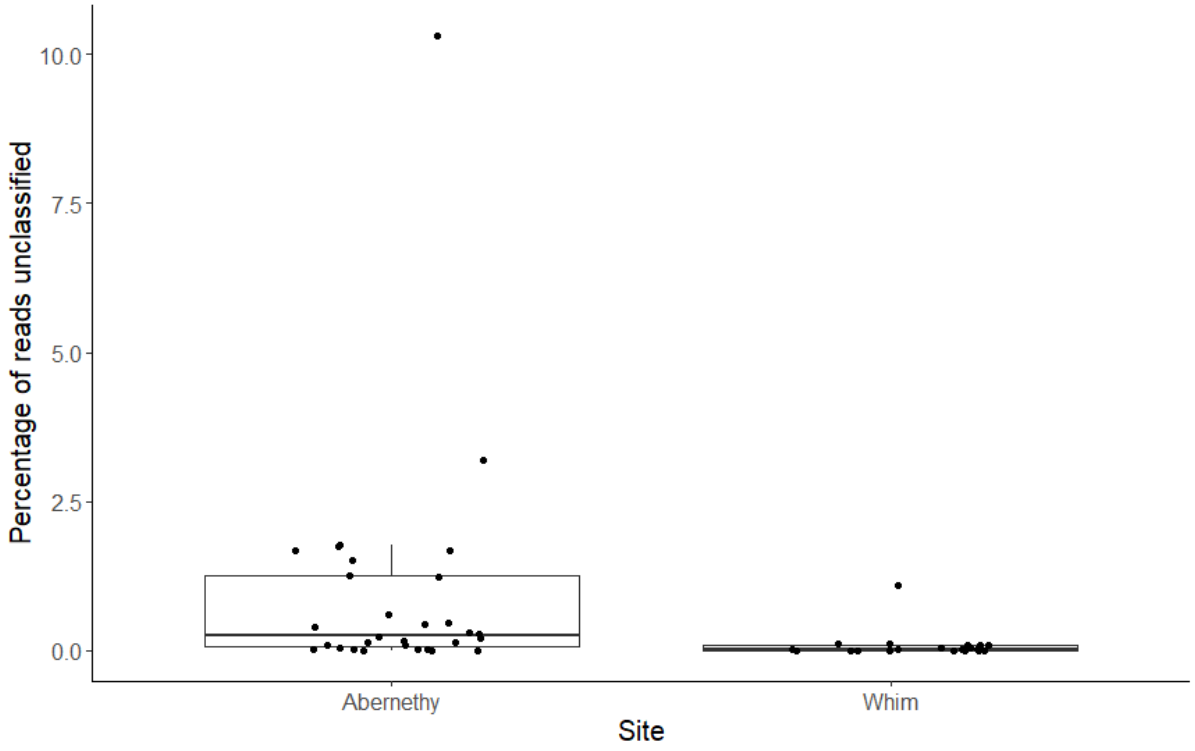



**Figure 3.8.** Percentage of unclassified reads from the final Kraken2 run for all samples. Abernethy is Chapter 5 data, and Whim is Chapter 4 data. Boxes show median and interquartile range. Points show the raw data.

## 3.6 In silico PCR

### 3.6.1 Introduction

An *in silico* PCR approach was used to test for primer complementarity and specificity with organisms which were missing from the identified ASV dataset from my initial bioinformatics methodology. For example, cyanobacteria and diatom ASVs were not identified, despite their complete absence from the samples being highly improbable. An *in silico* PCR approach allowed me to test whether the absence of these ASVs was due to use of unsuitable primers or a different issue, such as sample preparation or a failure in taxonomic assignment during the bioinformatics stage.

### 3.6.2 Methods and results

*In silico* PCR was carried out using the web app detailed in San Millán *et al*. (2013). rbcL primers (as detailed in section 3.2.1) were tested in both 5' to 3' and 3' to 5' orientation against genomes for 7 different cyanobacteria taxa. These were *Nostoc azollae* 0708, *Anabaena cylindrica* PCC 7122, *Anabaena* sp. 90 chromosome chANA01, *Anabaena variabilis* ATCC 29413, *Cyanobacterium* UCYN-A, *Cyanobacterium aponinum* PCC 10605 and *Cyanobacterium stanieri* PCC 7202. All returned no bands, indicating insufficient complementarity for PCR to occur.

In order to avoid any issues with one particular alignment algorithm, *in silico* PCR analysis was repeated using Primer-BLAST (Ye *et al*., 2012). As the rbcL primers which I used were originally designed for plant taxonomy, Primer-BLAST was initially run with specificity for *Arabidopsis thaliana* and *Malus domestica* to ensure settings were correct and the algorithm was working as expected. This resulted in successful simulated PCRs with the default Primer-BLAST settings. This was repeated searching for Bacillariophyta (diatoms) and cyanobacteria in the RefSeq representative genomes database and subsequently the nr database (of which the RefSeq representative genomes database is a subset). No target templates were found, indicating that the primers would not work for these organisms.

In contrast, Primer-BLAST found target templates for Desmidiales (desmids), Chlorophyta and Phaeophyceae (brown algae) when searching in the nr database. Search for target templates for Xanthophyceae (Yellow-green algae), Chrysophyceae (golden algae), Rhodophyta (red algae) and Archaea failed using the nr database. Further exploration of the Viridiplantae phylogenetic tree resulted in target templates found for the Palmophyllophyceae, Prasinodermophyceae, Mamiellophyceae and Charophyta. No target templates were found for the Glaucophyta, which are a sister group to Viridiplantae (Yang *et al*., 2023).

### 3.6.3 Implications

The results of the *in silico* PCR indicate that the rbcL primers which I used in my PCRs do not have sufficient complementarity to amplify cyanobacterial sequences, or the chloroplast sequences of most non-Viridiplantae (green) algal groups. This makes these primers unsuitable for a single primer approach to sequencing all photosynthetic microorganisms. This is unexpected, especially for diatoms, for which rbcL primers are one of the key molecular taxonomic tools. However, it appears that more specialized primers are required for diatom identification (e.g. Pérez-Burillo *et al*., 2020; An *et al*., 2020), and some papers even use a multiplex PCR approach with multiple primer sets in order to amplify DNA from as wide a range of diatom species as possible (Vasselon *et al*., 2017). Any future work attempting to use a functional primer usable across the full range of microalgae may be best carried out with a similar multiplexing approach, with primer sets designed to capture photosynthetic genes from different algal groups.

Further laboratory work to carry out PCRs using a wider range of primers, including diatom specific rbcL primers, has not been possible due to lack of funding/ expiry of funds. Therefore, no downstream analysis can be carried out on photosynthetic microbes of the taxa not amplified.

### 3.6.4 Prokaryotes

Whilst primer-BLAST did not show any target templates for cyanobacteria, it did return target templates for other bacteria using these rbcL primers. These are detailed in Supplementary Table 3.1. This suggests that the primers may have amplified DNA from bacteria which are not cyanobacteria. This may include some heterotrophic bacteria, such as Streptomycetes. However, there is also the chance of amplification of non-cyanobacteria autotrophic bacteria, for example anoxic phototrophs such as the purple bacteria. Whilst anoxic conditions would not normally be expected at the surface of the peatland or in the bryosphere, it is possible that microsite conditions may allow activity by anoxic taxa, especially in wetter microtopographies (Akhtar *et al.*, 2022). Of the other prokaryotic taxa as listed in Algaebank, Clostridiaceae and Spirochaetota had no target templates returned.

# 4 Photosynthetic microbial communities and nutrient loading – data using a long term experiment

## 4.1 Introduction

Peatlands are categorised by nutrient availability, with ombrotrophic raised bogs typically having very low nitrogen and phosphorus availability. Peatlands are N (or occasionally P) limited, but many in the UK and elsewhere have a history of anthropogenic N deposition. This has been identified as a threat to peatlands, not least because of the potential for increased N deposition to shift them to carbon sources (Payne, 2014; Bragazza *et al.*, 2006). P deposition can also weaken peatland carbon sinks by increasing ecosystem respiration (Lu *et al.*, 2023). Agriculture and industry can both be sources of N, and even peatlands remote from immediate sources can still be affected due to the long distances that N pollution can travel. Whilst on the whole N pollution has decreased in the UK over recent decades due to reductions in industrial pollution, some areas, for example the Forsinard Flows NNR (see Chapter 2) are predicted to have higher N pollution levels in 2030 than were recorded in 2010 (Payne, 2014).

Increased N deposition has been shown to change peatland plant communities in terms of structure and functions (Berendse *et al.*, 2001). Peat-forming *Sphagnum* mosses show reduced mass growth under increased N deposition (Sheppard *et al.*, 2014), in part due to competition from vascular plants and taller bryophytes such as *Polytrichium strictum* (Berendse *et al.*, 2001). High N deposition can also be toxic to *Sphagnum* mosses, reducing photosynthetic capacity and moss water content, and eventually leading to tissue damage and necrosis (van der Heijden, Verbeek and Kuiper, 2000). As peatland photosynthetic microbes live on the moss surface, this will then expose them to nutrient effects both directly, and indirectly via changes in the vegetation community composition. Increased anthropogenic N deposition may result in peatlands, and the *Sphagnum* mosses that are abundant in them, becoming P+K co-limited (Lund *et al.*, 2009; Rydin and Jeglum, 2006; Bragazza *et al.,* 2004), further increasing the complexity of responses.

Algal abundance often increases with nutrient deposition (Payne *et al*., 2013); indeed, visible algal growth on Calluna stems and trees is a common field observation on peatlands affected by high levels of nitrogen deposition. However, their responses may be less straightforward. Microalgae (including cyanobacteria) are involved in nutrient cycling. Many cyanobacteria can fix atmospheric nitrogen (Bothe *et al*., 2010) and so may in fact suffer a competitive disadvantage from increased N deposition. Because of the nutrient poor conditions of ombrotrophic peat bogs, many algae commonly found in peatlands have strategies for obtaining phosphate. For example, some algae produce phosphatase enzymes (Whitton *et al.*, 2004). Mixotrophs can feed via phagocytosis alongside autotrophy, which can help to alleviate P limitation (Koppelle *et al*., 2022). *Dicranochaete,* an epiphytic green algae found on *Sphagnum* leaves, has an increased surface area to aid in nutrient uptake, especially P (Caisová, 2016). These algae, therefore, might be expected to be resistant to P limitation. However, community and competition dynamics, including in vegetation (as many are functionally epiphytes) could still cause nutrient shifts to affect the community. There is good evidence in freshwater systems that increased nutrient pollution levels can increase dominance of some algae whilst reducing abundance of others (Simons *et al*., 1994). It therefore seems probable that the combination of the direct effects of nutrients on algal competition dynamics, and the indirect effects via vegetation changes, would result in shifts in the algal community with increased N and P deposition levels.

Molecular methods allow us to investigate if these changes are observed, by providing a potentially easier method for identifying algal groups which are difficult (or even impossible) to identify with light microscopy. Therefore, they expand the potential resolution at which we can investigate algal communities as a whole, and may provide alternative candidates for bioindicators. Even for more easily visually identified taxa, such as diatoms and desmids, this could prove useful. For diatoms, a good relationship has been shown between results of metabarcoding with rbcL and traditional species identification with light microscopy (Pérez-Burillo *et al*., 2020). However, light microscopy may under-represent some species which are indicators for assessment (ibid), suggesting that rbcL DNA barcoding may have potential for more thorough assessment of indicator species for at least some algal groups. In addition to algae

which are difficult to visually distinguish, molecular methods also allow algae which are difficult to physically remove intact from the growing surface to be studied. These include cyanobacteria within the hyaline cells of *Sphagnum* (Bragina *et al*., 2012) and tightly associated epiphytic algae such as *Dicranochaete* (Caisová, 2016). Both of these examples of 'challenging' algae for light microscopy are important for understanding community responses to changes in nutrient statuses. Cyanobacteria are nitrogen fixers, and this N is incorporated in the biomass of their symbiotic partner *Sphagnum*, increasing *Sphagnum* growth (Berg *et al*., 2013). *Dicranochaete* has a seta with increased surface area, likely to aid in the uptake of phosphorus in nutrient poor environments (Caisová, 2016).

Increased levels of nutrient pollution, especially N, are a threat to UK peatlands and may result in changes in peatland functioning. They are also predicted to result in changes in biodiversity in both the vegetation and microbial communities. Algal responses to changes in nutrient loading are likely to be complex, and an attempt to reveal patterns in these responses should allow us to better understand how to manage peatlands and their ecological communities in the face of change. This study aims to address two different but linked hypotheses:

- Photosynthetic microbial communities change with increasing levels of N deposition, with higher levels of N deposition associated with greater changes in the community composition.
- Co-limitation of N and P affects photosynthetic microbial communities, and therefore addition of N and P results in greater changes in community composition than addition of N alone.

## 4.2 Methods

### 4.2.1 Site description and experimental treatments

Whim Bog is a *Calluna vulgaris-Eriophorum vaginatum* dominated blanket bog in the central belt of Scotland (55.765890, -3.270651) (Fig. 4.1). The cryptogam community includes *Sphagnum* mosses, especially hummock-forming *S. capillifolium*. *S. fallax* and

*S. papillosum* can also be found in hollows. Peat is between 3 and 6m deep, with a pH of ~3.5. There is some drainage, including a channel at the side of the experimental site, due to neighbouring commercial peat workings. While it is classed as an ombrotrophic acid bog, it has experienced historic anthropogenic N deposition due to both industry and agriculture taking place nearby. It is now the site of a long-term nutrient addition and monitoring experiment run by the Centre for Ecology and Hydrology (UK CEH) (Sheppard *et al.*, 2004).

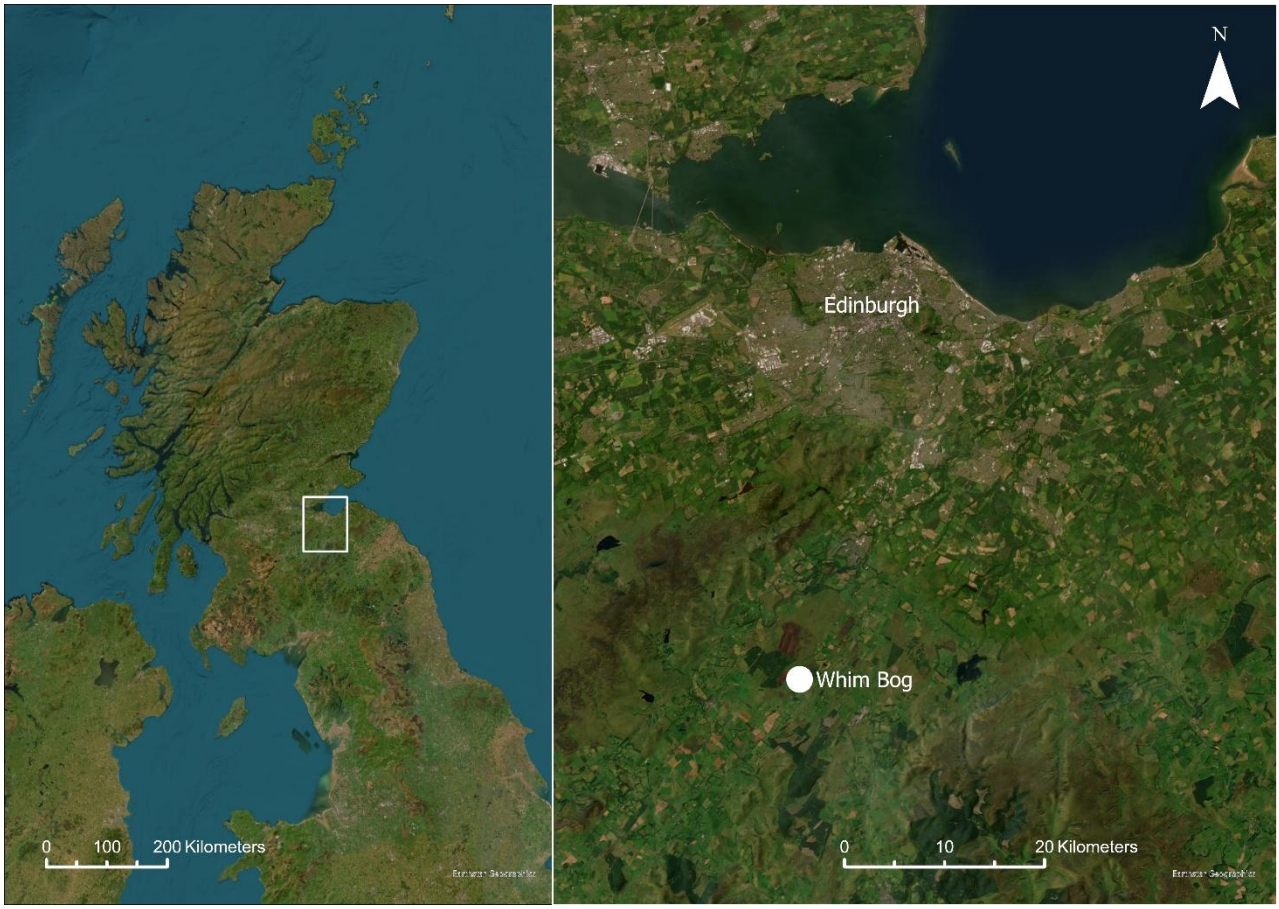



**Figure 4.1.** Map showing the location of Whim Bog within the UK. Satellite imagery from ArcGIS World Imagery (Sources: Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community).

There are 44 wet deposition experimental plots, each measuring 13m$^2$ (Fig. 4.2). Nitrogen is added at a rate of 8, 24, and 56 kg N ha$^{-1}$ y$^{-1}$ in reduced ($NH_4Cl$) and oxidised ($NaNO_3$) forms (Sheppard *et al.*, 2004). There are also N+PK treatments (added as

$K_2HPO_4$ at a 1:14 ratio of P to N) at the 24 and 56 kg N $ha^{-1}$ $y^{-1}$ levels. Each treatment has 4 replicates, arranged in a block design. Plots can be accessed via boardwalk and ladder to minimise disturbance. Wet treatments are carried out by automatic spraying. Rainfall is collected on site; when the reservoirs are filled, N, P and K are added at the correct concentrations for the treatments and piped to plots. Spraying only takes place when wind speed is below 5m $s^{-1}$ and temperatures are above freezing. Control plots are sprayed with rainwater with no nutrient addition.

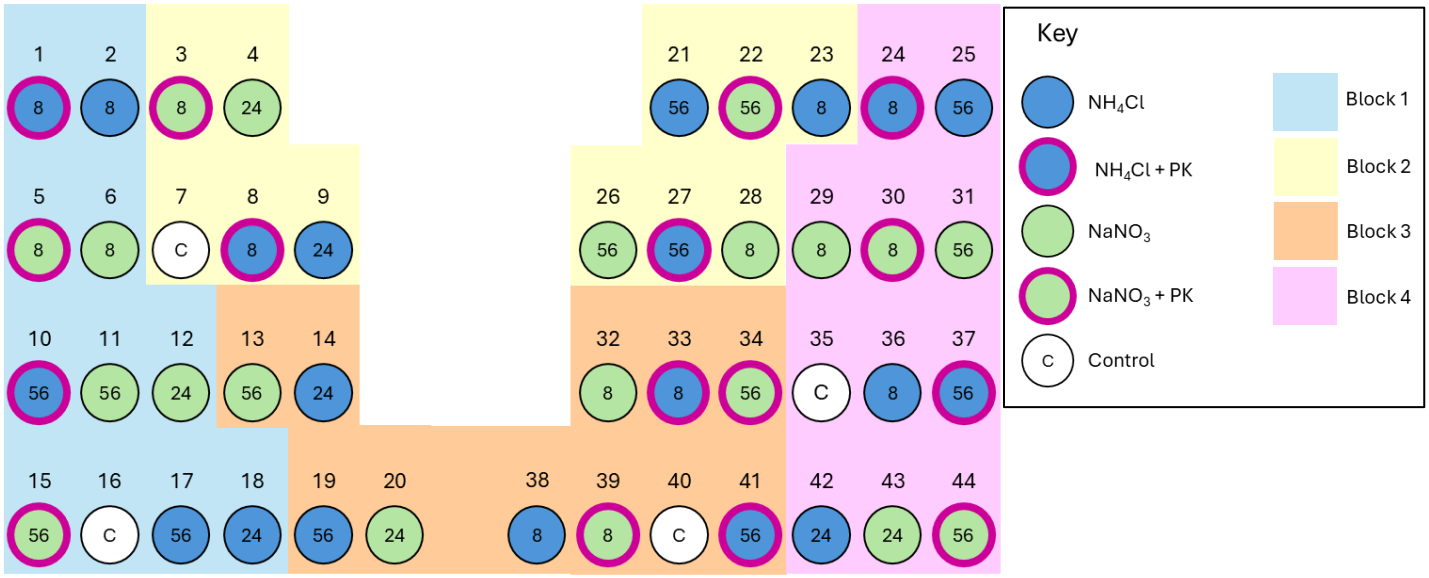



**Figure 4.2.** Diagram of the CEH Whim Bog wet deposition experiment. Circles show experimental plots, with colours showing nutrient treatment (see key), numbers in circles showing N addition in $ha^{-1}y^{-1}$ above background deposition levels, and numbers above circles showing plot ID numbers. Background colour shows experimental block design (see key). Figure adapted from one provided by Dr Matthew Jones, UK CEH (pers. comm., 6 October 2020).

The current background nitrogen deposition rate is 8-11 kg N/ha/yr (Dr Matthew Jones, UK CEH, pers. comm., 6 October 2020). The nearby Auchencorth Moss UK-Air monitoring site had a mean potassium in precipitation of 0.113 mg/l/day and a mean phosphate as P in precipitation of 0.011 mg/l/day between 19/12/2018 and 01/07/2020 (DEFRA, 2020).

Whilst the experiment included both P and K added in addition to N, the decision was made to focus on P colimitation because of the links to algal functional ecology (see

Section 4.1). In addition, while peatlands can be K limited, N or P limitation is more common, and K limitation is associated with an excess of Ca (Rydin and Jeglum, 2006), so the assumption was made that P was more likely than K to be a colimiting nutrient at this site. However, as P and K were added together, truly separating P and K colimitation is not possible from this study.

### 4.2.2 Field sampling

Sampling was carried out over the course of a single day towards the end of the growing season, on 6/10/2020. Samples were taken from wet deposition experimental plots with reduced N addition as $NH_4Cl$. Plots sampled had background N deposition (8-11kg N/ha/yr), background + 8kg N/ha/yr (total ~16kg N/ha/yr) with and without PK (added as $K_2HPO_4$ at a 1:14 ratio of P to N), and background + 24 kg N/ha/yr (total ~32kg N/ha/yr). One sample at each treatment level was taken from each of the four experimental blocks (Fig. 4.3).

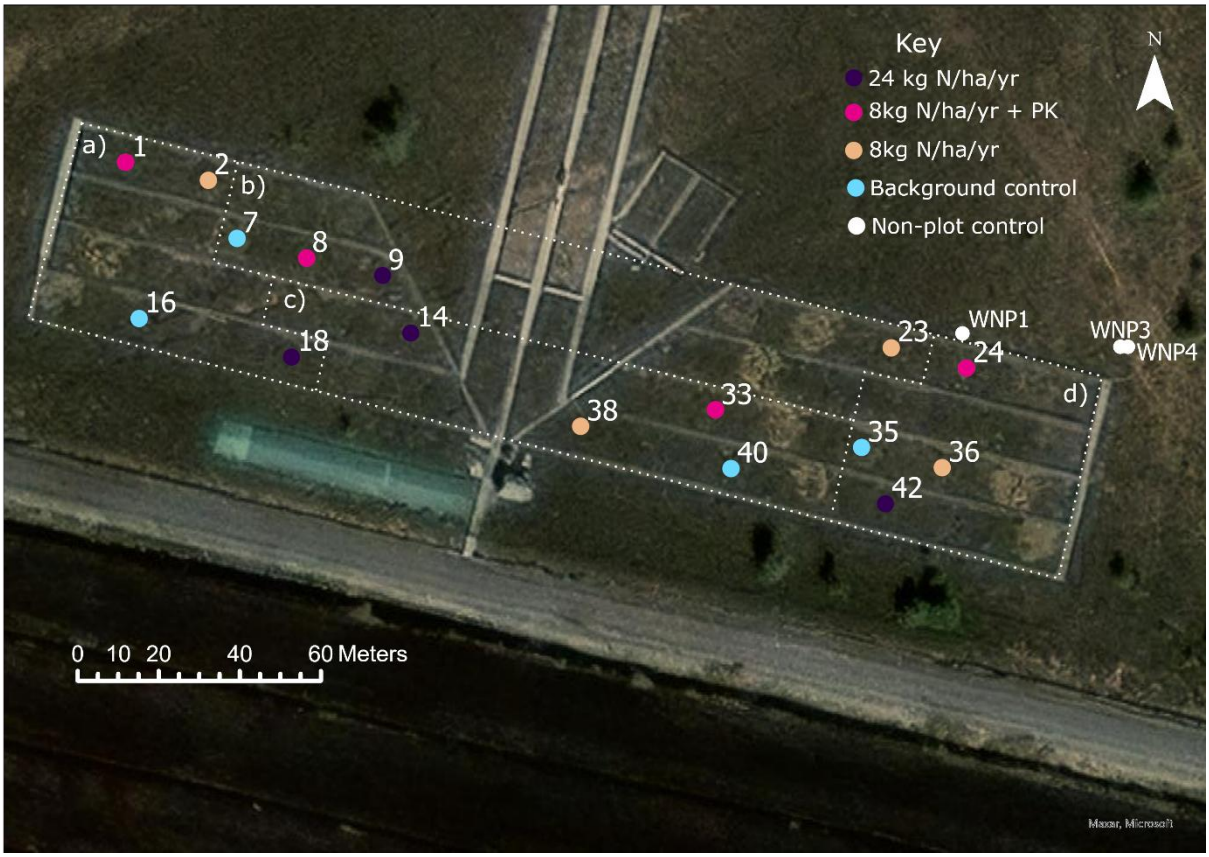



**Figure 4.3.** Satellite imagery showing the location of experimental treatment plots from the CEH Whim Bog wet deposition experiment and non-plot controls used in Chapter 4. Numbers show the plot IDs, and colours show the nutrient addition treatment levels above background deposition. WNP2 (non-plot control 2, located between WNP1 and WNP3) is not shown due to missing GPS data. Blocks are delimited by dotted lines as follows: a) block 1, b) block 2, c) block 3, d) block 4. Satellite imagery from ArcGIS World Imagery (Sources: Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community).

The decision was taken to focus on these levels and exclude the higher 64kg N/ha/yr treatment, as this is far higher than the current exposure level predicted for other sites used within this thesis. Furthermore, *Sphagnum* presence was reduced in 64kg plots, which would have made sampling more difficult; as *Sphagnum* increases the surface moisture and provides a substrate for the algae to grow on, it would be difficult to unpick differences directly due to changes in nutrient loading versus differences due to changes in host vegetation.

Reduced N as $NH_4^+$ was chosen instead of oxidised N due to its predicted greater relevance as a pollutant into the future. Oxidised N deposition levels have fallen more rapidly since the 1990s in the UK and Europe than reduced N (Dragosits *et al.*, 2020; European Environment Agency, 2016). Reductions in burning coal, new technology to reduce $NO_x$ emissions in the energy industry and catalytic converters reducing emissions from road transport have resulted in lower emissions of oxidised nitrogen. However, as reduced N is primarily caused by agriculture, it is more difficult to mitigate and therefore emissions are projected to fall more slowly (ibid).

Four replicate controls for experimental disturbance were also taken from within the Whim site, along the boundary with the neighbouring Site of Special Scientific Interest (SSSI) but away from treatment plots, to control for any impacts of frequent human activity within the experimental plots themselves (Fig. 4.3). This resulted in a total of 20 samples.

#### *4.2.2.1 Sphagnum sampling*

The top 5cm of 8-10 individual *Sphagnum capillifolium* capitula was sampled for each plot to capture the bryophyte photic zone. Samples were taken from at least four *Sphagnum* hummocks within each plot for the background, 16kg N/ha/yr and 16kg N/ha/yr + PK treatments, avoiding long term vegetation monitoring quadrats, to capture spatial heterogeneity. In the 32kg N/ha/yr plots, there were not always four *Sphagnum* hummocks available; in these cases samples were taken from at least two hummocks. *Sphagnum* samples for the plot were then pooled for processing, storage and analysis. Out of plot controls were taken from individual *Sphagnum capillifolium* hummocks further than 1m away from each other and not contiguous, in order to be able to consider them independent.

#### *4.2.2.2 Transport and storage*

*Sphagnum* samples were processed to remove any of the capitula more than 5cm from the surface, and transported back to York on ice in an insulated bag. Samples for molecular work were frozen at -80˚C on the day of sampling.

## 4.2.3 Laboratory work

### *4.2.3.1 DNA extraction*

In order to minimise the amount of moss-derived sequences in samples, and reduce the likeliness of them swamping microbial samples during PCR and sequencing, a pre-filtering step was included to remove larger moss fragments from the samples. This was modified from Heger *et al*. (2018) and Jassey *et al*. (2011).

*Sphagnum* samples were cut into smaller pieces with sterilised scissors and placed in a clean, sterile centrifuge tube. 20ml sterilised d2$H_2O$ water was added and the tube was shaken by hand 30 times. Water and moss were pre-filtered by passing through a 195µm nylon microsieve. The moss was squeezed over the microsieve using a spatula to collect the maximum amount of filtrate. This was repeated a further two times, resulting in a total of 60ml of filtrate for each sample, which was combined. The filtrate was then filtered through a pre-sterilised 0.2µm MCE (mixed cellulose ester) filter under vacuum. All equipment for DNA extraction was sterilised in the autoclave where possible. For equipment which could not be autoclaved, sterilisation was carried out with manual washing with soap and water followed by cleaning with 70% ethanol.

A number of studies show differences in DNA yield and community composition depending on filter selection. Mixed cellulose ester filters were selected for this study because of high DNA yield, possibly due to greater DNA binding capacity than other filter types (Majaneva *et al*., 2018; Hinlo *et al*., 2017; Liang and Keeley, 2013). As my samples had been previously frozen, I considered that the ability to bind extracellular DNA would reduce potential bias between cells which had and had not undergone lysis following freeze-thaw.

Filter paper was cut into small pieces of approximately 2$mm^2$ using sterile scissors. DNA was extracted from the filters using Quiagen Powersoil DNA extraction kits, following standard manufacturers protocols.

Extracted DNA was quantified and quality checked using a nanodrop. This showed mixed DNA yields but relatively high-quality extracted DNA. However, A260/A230 ratios

suggested possible presence of contaminants (e.g. phenols) in the extracted DNA. Therefore the decision was made to include Blotto in PCR reactions (see Section 3.2.2 for more details). Blotto was prepared following Burn (2021) as a 1mg/ml solution of skimmed milk powder and deionised water, vortexed to mix and sterilised by autoclaving.

#### *4.2.3.2 PCRs and sequencing*

PCRs were carried out following the protocol described in Fazekas *et al.* (2012), as described in Table 3.2 (Section 3.2.2). For more details of PCRs and sequencing, see Chapter 3.

#### *4.2.3.3 Pore water nutrients*

Pore water was extracted from two *Sphagnum* capitula by centrifuging at 4000 RPM (Hettich Rotanta 460 centrifuge) for 5 minutes through a Pall Life Sciences 1µm type A/E glass fibre filter (see Fig. 4.4). Pore water was frozen at -20˚C for further analysis.

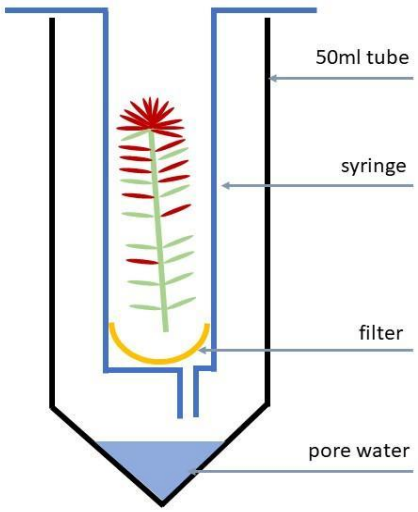



**Figure 4.4.** Diagram showing the setup for pore water extraction from Sphagnum capitula using a centrifuge.

The volume of pore water extracted for each sample varied between 64µl and 326µl. 2000µl of d2$H_2O$ was added to each pore water sample in order to make it up to the required volume for analysis. Pore water was analysed for ammonium N, nitrate/ total organic N and phosphate P using a Seal Analytical AA3 auto-analyser 3 (Seal Analytical, Wrexham, UK). This was operated according to the manufacturers instructions, and calibrated as standard during the run with reference samples of known concentration. However, for all samples, the phosphate measurements were below the theoretical detection threshold as specified by the manufacturer (Fig. 4.5) (Seal AA3 AutoAnalyser Operating Manual). This is likely due to the high dilutions needed to make the samples up to the volume required for analysis. Phosphate levels were calculated to below zero, likely due to the calibration curve. Examining the raw data (Fig. 4.6) showed fluctuations around zero. Neither manually adjusting the point on the peak at which zero was measured or shifting the baseline on the AA3 software was able to correct this.

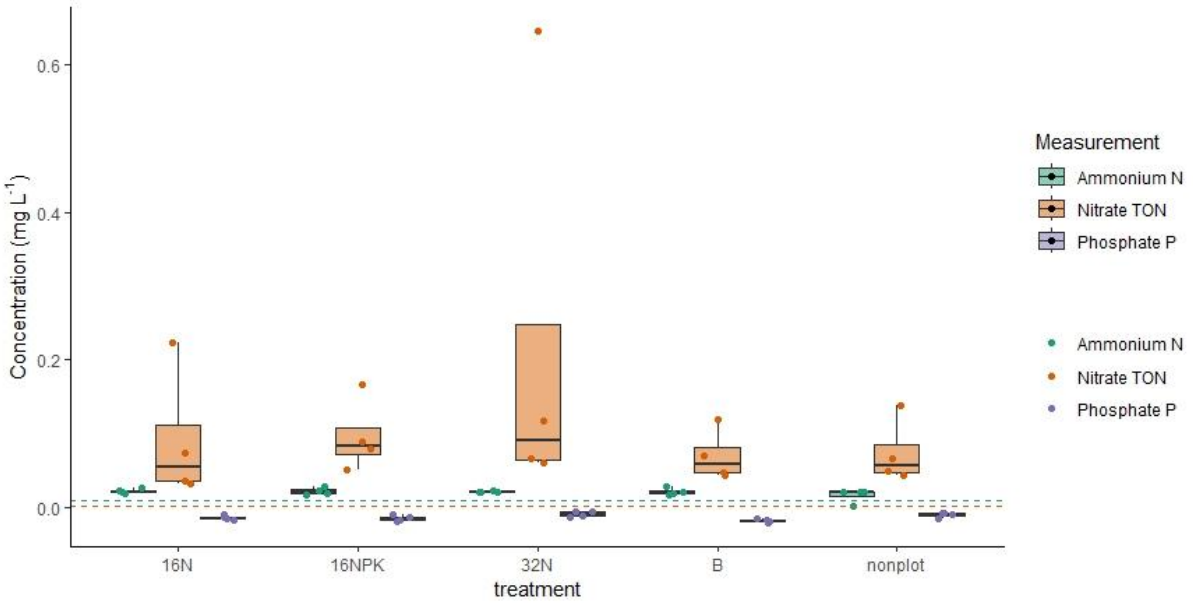



**Figure 4.5.** *Sphagnum* pore water nutrient concentrations with different nutrient addition treatments. Points are the individual sample measurements. The dotted lines show the theoretical detection limit according to the manufacturer's instructions for each nutrient. Nitrate and phosphate both have a theoretical detection limit of 0.003mg/L, so all phosphate measurements are below this threshold, whereas most nitrate measurements are above it. Boxes show median and interquartile range. Points show actual measurements. The treatment labelled B is background deposition control plots and nonplot is the experimental disturbance controls.

While ammonia and nitrate are both above the theoretical detection threshold, the real-world detection threshold has been measured as higher (Hodson *et al.*, 2023; Khan 2022). Ammonia has almost no variability both between and within my samples, and at just above the detection threshold should be treated with caution. Nitrate appears to be more variable but differences between treatments appear to be largely due to individual samples (Fig. 4.5). There is no statistically significant difference in ammonia (Kruskal-Wallis: chi-squared = 0.66, df = 4, p = 0.96) or nitrate (Kruskal-Wallis: chi-squared = 2.53, df = 4, p = 0.64) levels between the different nutrient addition treatments or controls. This may be indicative of a true lack of differences in the pore water nutrient levels, or may simply be a consequence of the high dilution level. When these results were compared to detection levels measured from previous studies which used the same machine (Fig. 4.7), all phosphate measurements were below the real-world detection limit reported in Khan (2022) and Hodson *et al*. (2023), and all ammonium measurements were below the detection limit reported in Khan (2022). All nitrate measurements except one were below the detection limit from Khan (2022), but above the detection limit from Hodson *et al*. (2023), and therefore should be treated with caution. Therefore for the rest of the study, nutrient addition was considered as a treatment factor rather than via the measured pore water nutrients.

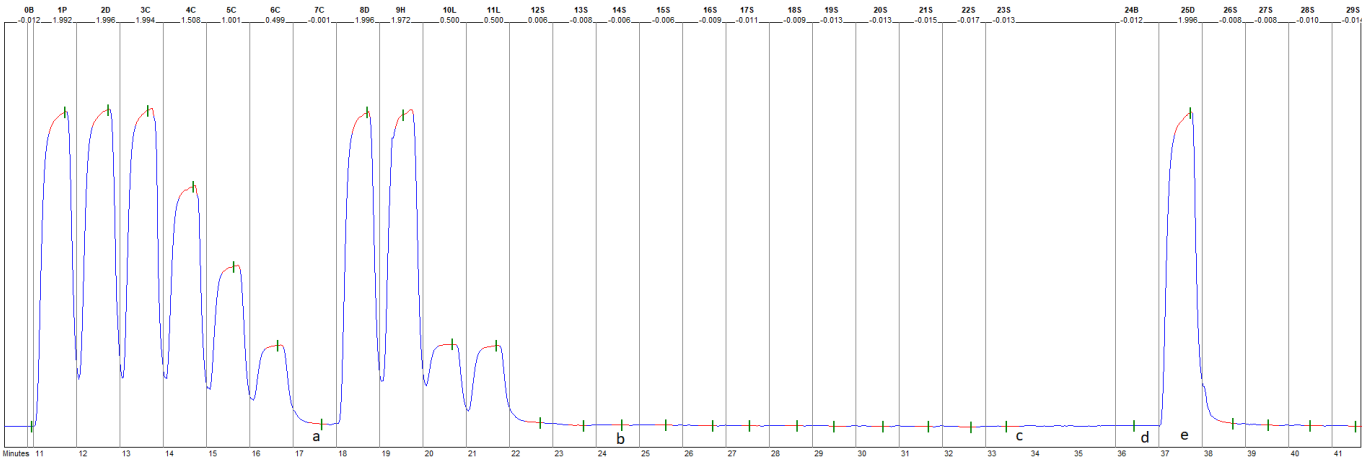



**Figure 4.6.** Chart showing trace from AutoAnalyser3 channel for phosphate. Point a) is the calibration peak for 0.0mg/L, points b-c) are the first 10 samples, d) is the baseline and e) is the drift correction peak. Sample measurements fluctuate around zero.

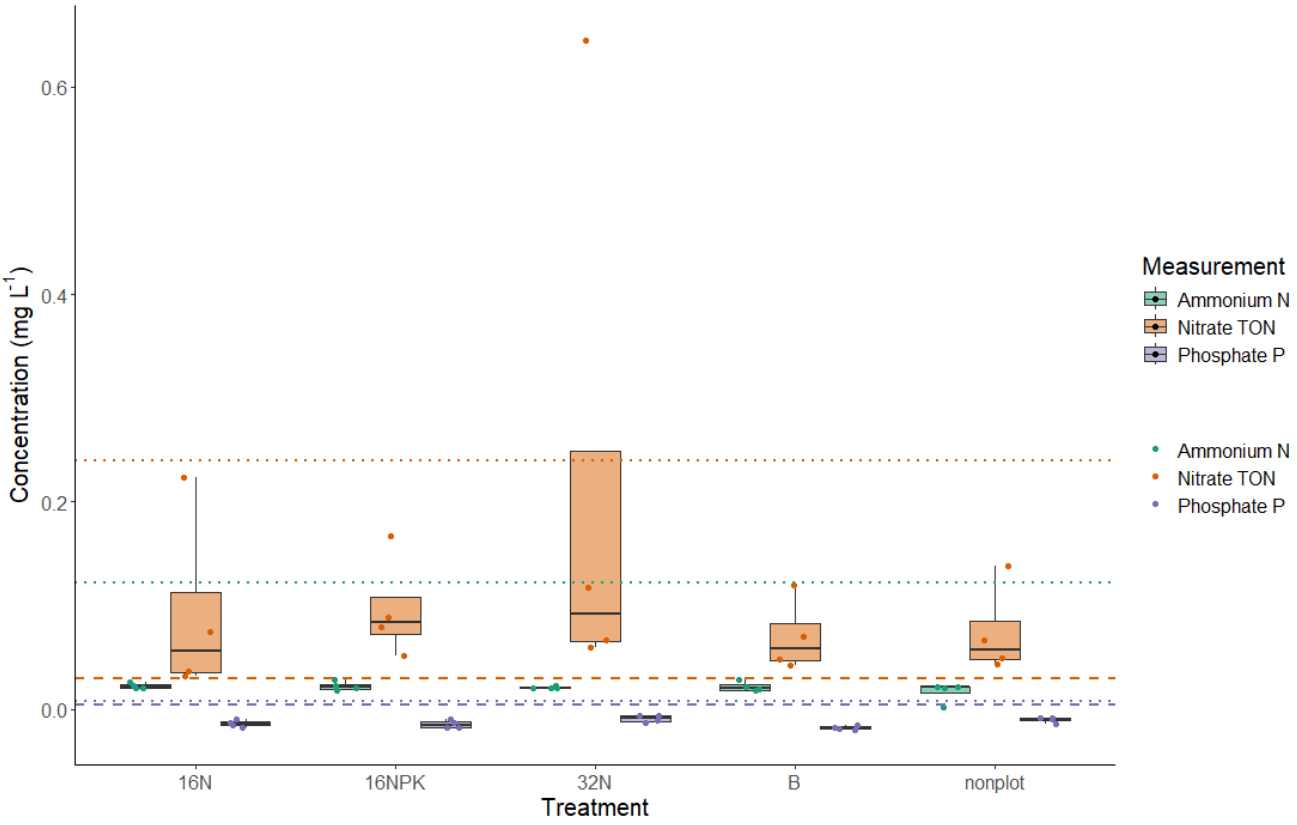



**Figure 4.7.** Sphagnum pore water nutrients compared to measured real-world detection limits from the AutoAnalyser3 on which they were measured. Dashed lines show the detection limits reported in Hodson *et al*. (2023) and dotted lines show detection limits reported in Khan (2022). All ammonium measurements are below the detection limit from Khan (2022), and all phosphate measurements are below both measured detection limits. All nitrate measurements except one are below the detection limit from Khan (2022), but above the detection limit from Hodson *et al*. (2023). Boxes show median and interquartile range. Points show actual measurements. The treatment labelled B is background deposition control plots and nonplot is the experimental disturbance controls.

## 4.2.4 Data analysis

### *4.2.4.1 Original bioinformatics*

As there was no reference database for rbcL available, a custom database was created. This combined the curated Rsyst:: diatom database diatbarcode v. 11.1 (Rimet *et al*., 2022) and the curated taxonomic, genetic and ecological data resource for the vascular

plants of Britain and Ireland (Henniges *et al.*, 2021), supplemented with rbcL records downloaded from NCBI via Geneius Prime (2023.2.1). For full details see section 3.3.1.

Bioinformatics was carried out using DADA2 (Callahan *et al.*, 2016) and Phyloseq (McMurdie and Holmes, 2013) in R (version 4.1.2) in Rstudio, following a pipeline adapted from Burn (2021). For full details of initial bioinformatics pipeline, see Section 3.3.

#### *4.2.4.2 Revised bioinformatics*

The original bioinformatics method resulted in very low levels of identification of ASVs, likely because of insufficient taxonomic coverage from the custom reference database (see Chapter 3 for more details). Therefore, a different method was used for taxonomic assignment.

Bioinformatics was initially carried out using DADA2 (Callahan *et al.,* 2016) and Phyloseq (McMurdie and Holmes, 2013) in R (version 4.1.2) in Rstudio, following a pipeline adapted from Burn (2021), as for the first method.

The cleaned sequence abundance tables from DADA2 were then manually converted into FASTA files in R. These were then inputted into Kraken2 (Wood *et al.*, 2019) for taxonomic assignment, using a custom reference database comprised of a full download of NCBI data (Thurston, in preparation). Kraken2 was called in standard (not quick) mode. Bracken (Lu *et al.*, 2017) was then used to calculate abundances for the identified taxons. Due to uncertainty around identification at species level, the decision was made to collate Bracken outputs to order level. For full details of bioinformatics revision method development, see Section 3.5.

#### *4.2.4.3 Initial data processing and statistics*

Data were filtered to remove all ASVs identified as plants (vascular and non-vascular) before further analysis. Whilst ASV identification is not perfect, the reference database included almost all UK vascular plants, so assignment for these is likely to be accurate. There is more potential for incorrect ASV assignment for non-vascular plants, which were less well represented in the database.

Algal community differences between nutrient addition levels were tested with PERMANOVA performed using adonis from the vegan function (Oksanen *et al.*, 2022) in R. Pairwise PERMANOVA was carried out with the mctoolsr package (Leff, 2022).

Differences in variance (beta-dispersal) of the algal community between treatments were tested using a permutation test for homogeneity of multivariate dispersions, using the function betadisper with 999 permutations, in vegan.

Differences in the algal communities caused by block effect and disturbance effect (experimental plot versus non-plot samples) were also tested for using betadisper and PERMANOVA.

#### *4.2.4.4 Data processing and statistics following revised bioinformatics*

Data were filtered to remove taxa likely to be erroneous, either due to sequencing errors e.g. chimeras which were not removed, identification errors or incorrect entries in the databases. The latter would be possible if for example epiphytic algae or gut content DNA is extracted, sequenced with the host DNA and included in the databases under the host taxon identification. The taxa filtered out were Annelida, Arthropoda and Chordata, which were the animal taxa identified.

Total read counts per sample were calculated before and after removal of these taxa. The removal did not affect the read counts of the six samples with the lowest number of reads (see Table 4.1). No samples needed to be filtered out because of low read count (Fig. 4.8).

**Table 4.1.** Total read counts for the six samples with the lowest number of reads from Chapter 4.

| Sample | Total read count before removal of animal reads | Total read count after removal of animal reads |
|---|---|---|
| W40 | 523993 | 523993 |
| WNP3 | 709310 | 709310 |
| W2 | 1050236 | 1050236 |
| W18 | 1063667 | 1063667 |
| WNP1 | 1072516 | 1072516 |
| WNP2 | 1116488 | 1116488 |

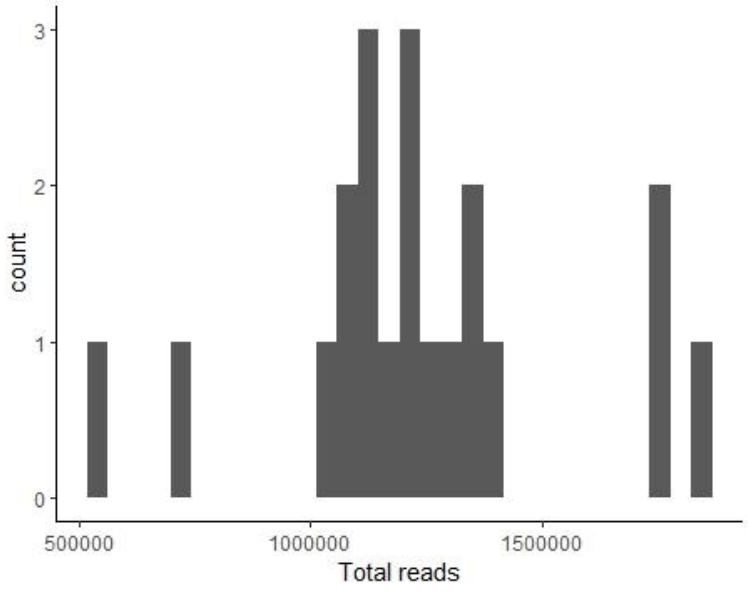



**Figure 4.8.** Histogram of total read counts for all samples from Chapter 4. All samples had a sufficiently high read count to progress through the analysis pipeline.

Rarefaction curves were calculated using the rarecurve function in vegan. A cumulative sum scaling (CSS) transformation was carried out on the abundance table using the metagenomeSeq package (Paulson *et al.*, 2013). This CSS transformed data was then plotted to see if there was any correlation between observed features and sequencing depth. As the rarefaction curves appeared to all be levelling out or to have met their

asymptote (Fig. 4.9) and there was no correlation between observed richness and read count (a proxy for sequencing depth) (Fig. 4.10), the decision was taken that data did not need to be normalised by sequencing depth. This is likely due to pooling the data to order level, reducing the number of ASVs involved.

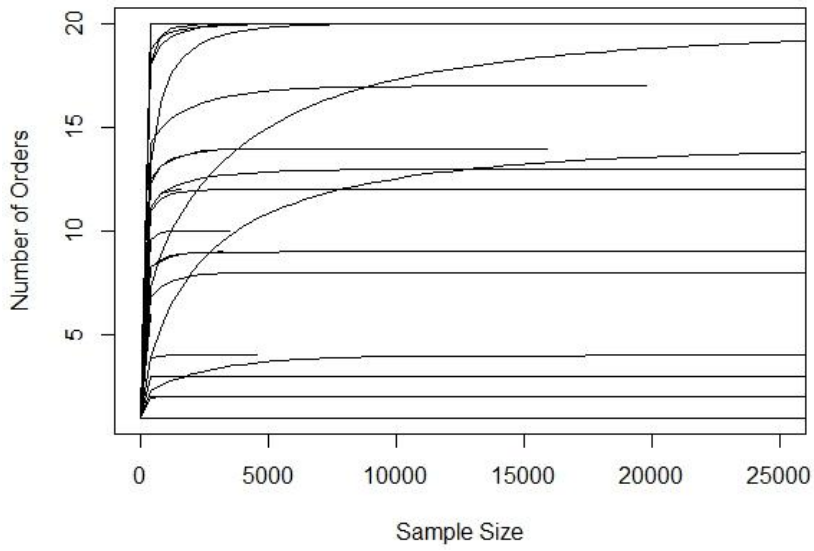



**Figure 4.9.** Rarefaction curves showing biodiversity by sampling effort for each sample in Chapter 4. Note that as the data has been pooled at the order level, the 'ASVs' are orders rather than species/ASVs proper. All samples appeared to be reaching or have reached their asymptote.

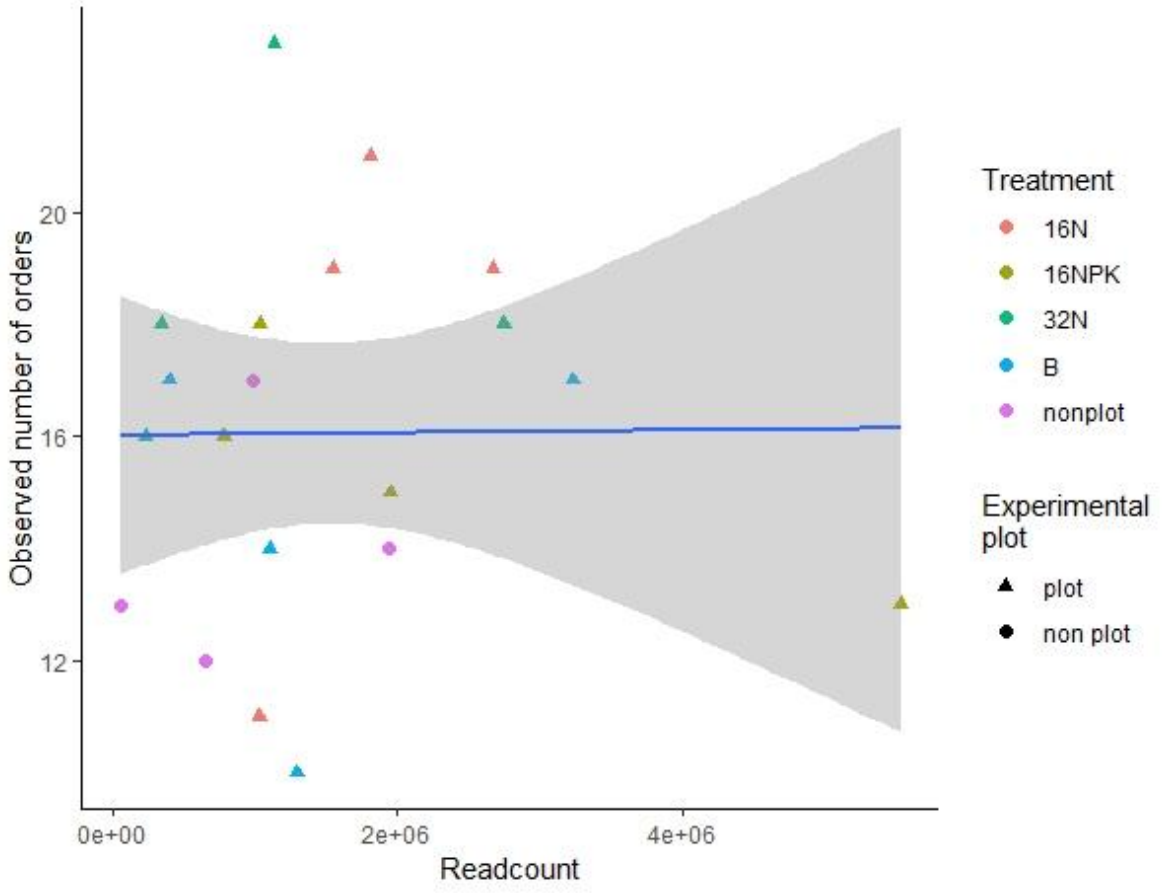



**Figure 4.10.** Correlation plot between observed taxonomic richness and read count (a proxy for sequencing depth) for Chapter 4 data (CSS transformed). There is no indication of a correlation between the two, and therefore no need to normalise data by sequencing depth. The treatment labelled B is background deposition control plots and nonplot is the experimental disturbance controls.

As the bioinformatic data had been pooled at the order level, standard biodiversity indices such as species richness and evenness could not be calculated for this data. However, as the initial bioinformatics was processed in the same way prior to taxonomic assignment for both methods, the first estimates of diversity indices could still be used as long as the taxonomic uncertainty was taken into account.

To investigate the ratio of prokaryotes to eukaryotes between different nutrient treatments, the percentage of all microbial reads (plant taxa excluded) which were prokaryotes of any taxa was calculated. As the data was not normally distributed (Shapiro-Wilk normality test $W = 0.621$, $p \leq 0.001$), a Kruskall-Wallis test was used.

As with the original data sets, algal community differences between nutrient addition levels were tested with PERMANOVA performed using adonis from the vegan function in R. Pairwise PERMANOVA was carried out with the mctoolsr package (Leff, 2022).

Beta-disperal (variance in dispersal) between groups was tested using a permutation test for homogeneity of multivariate dispersions, using betadisper with 999 permutations, in vegan.

Differences in the algal communities caused by block effect and experimental disturbance effect (experimental plot versus non-plot samples) were also tested for using betadisper and PERMANOVA.

To test for spatial patterns, a Mantel test was carried out using Vegan in R between the Bray-Curtis distance matrix for algal communities vs. the geospatial distance. Haversine geospatial distances were calculated from latitude-longitude data using the distm function in the geosphere package (Hijmans, 2024). Coordinates for wet deposition plots were kindly provided by Dr Matthew Jones from UK CEH (pers. comm., 18 April 2025). Coordinates for non-plot controls were collected with iPhone SE internal GPS and converted from degrees, decimal minutes into lat/long using the Ordnance Survey Coordinate Transformation Tool (Ordnance Survey, 2025).

Coordinates were sense-checked in ArcGIS, which revealed a coordinate error for one of the non-plot controls (sample WNP2) (see Fig. 4.11). This may have been due to human error or a drop in GPS signal when recording the coordinates. Therefore, this sample was removed from the data, distance matrices were recalculated, and the Mantel test was repeated.




**Figure 4.11.** Location of incorrect coordinate for sample WNP2, which when plotted appears to be a distance away from the study site. Satellite imagery from ArcGIS World Imagery (Sources: Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community).

## 4.3 Results

### 4.3.1 Algal community changes

#### *4.3.1.1 Original results*

Nutrient addition treatment did not significantly affect algal community (PERMANOVA: $R^2 = 0.33$, $p = 0.19$). The pairwise PERMANOVA showed significant community differences between 16NPK and non-plot controls ($R^2 = 0.75$, $p = 0.035$), and background nutrient addition plots vs non-plot controls ($R^2 = 0.41$, $p = 0.024$). However, when Bonferroni and FDR corrections were calculated, no pairs had significant differences.

There were significant differences in beta - dispersal between the different treatment groups (F = 3.26, p = 0.039). This should not affect the validity of the PERMANOVA results; as the experiment has a balanced design if non-plot and background treatments are considered separately, PERMANOVA should be resilient to this (Anderson and Walsh, 2013). However, it does suggest differences in the levels of variance between plots.

A Tukey HSD test was carried out on the betadispersal distance matrix to test which treatments had significant differences in dispersal. However, this showed no significant adjusted p - values. Nonplot-32N (p = 0.073) and Nonplot-Background (p = 0.064) were the closest to significant (Fig. 4.12).

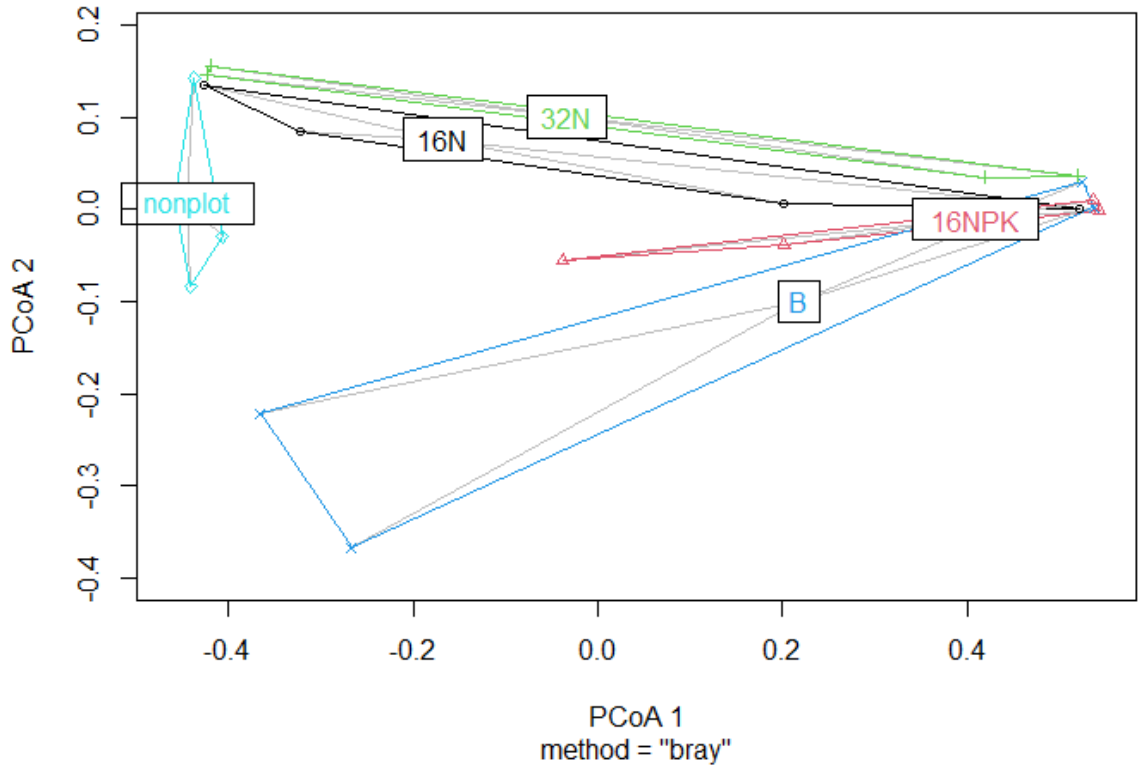



**Figure 4.12.** PCoA plot showing the difference in algal communities between nutrient addition treatments (as identified by the first bioinformatics method). Samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

### *4.3.1.2 Kraken results*

Detrended correspondence analysis reveals no clear patterns in algal community data (Fig. 4.13). As with the first bioinformatics pipeline, nutrient addition treatment does not significantly affect algal community (PERMANOVA: $R^2 = 0.099$, $p = 0.98$). The pairwise PERMANOVA showed significant community differences only between 16NPK and 32N ($R^2 = 0.33$, $p = 0.05$). However, when Bonferroni and FDR corrections were calculated, no pairs had significant differences.

There were no significant differences in beta - dispersal between the different treatment groups ($F = 0.595$, $p = 0.655$) (Fig. 4.14).

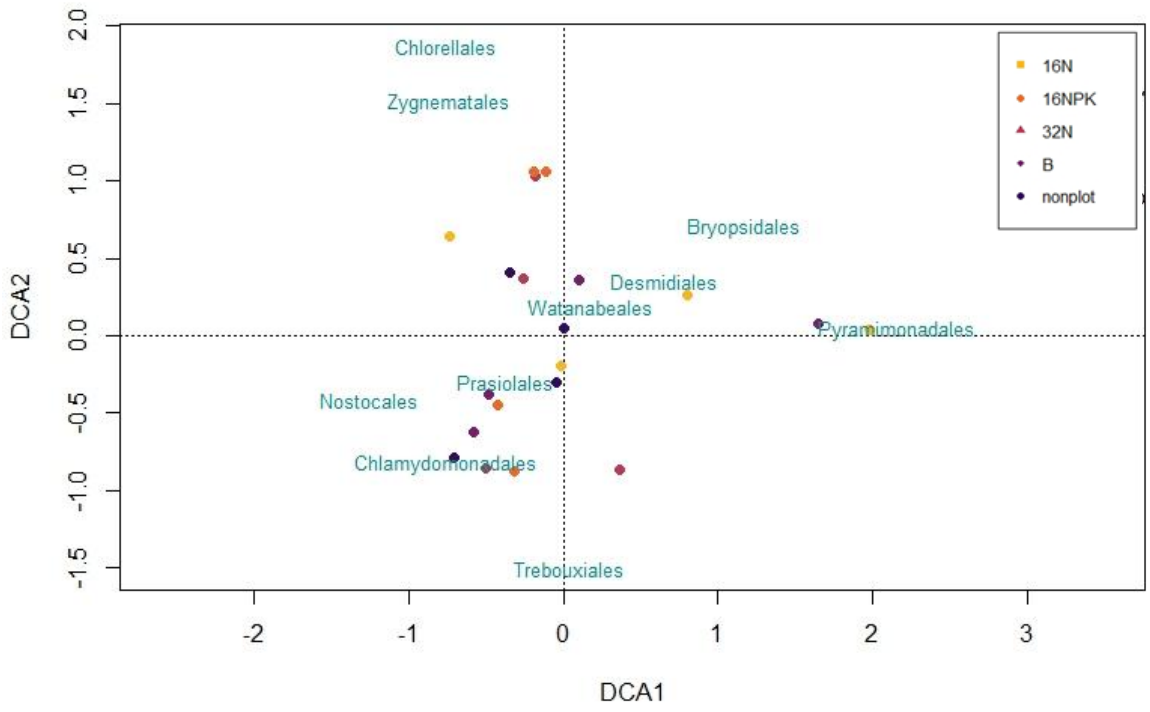



**Figure 4.13.** Detrended Correspondence Analysis (DCA) plot of algal community by treatment, and the algae classes present (as identified with the second bioinformatics method). Samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

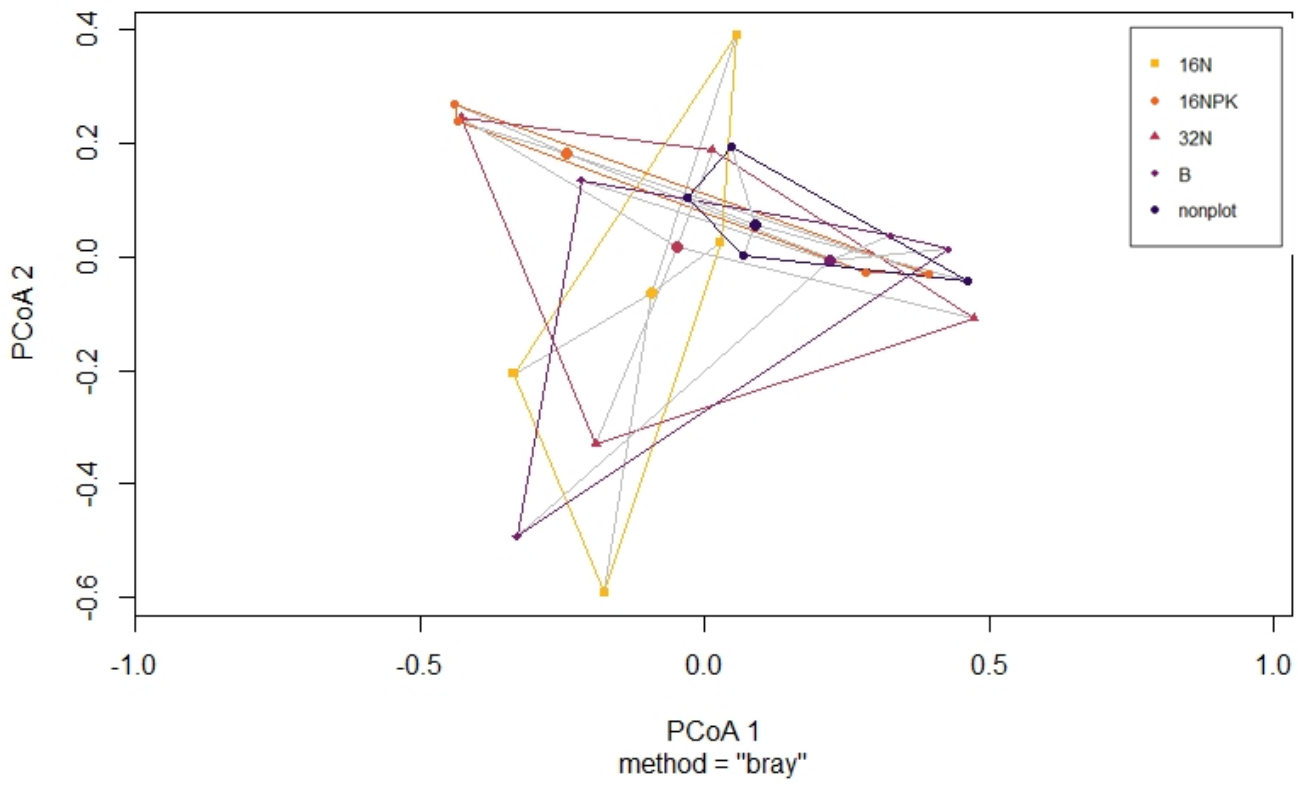



**Figure 4.14.** PCoA plot of beta dispersal of algae communities by nutrient addition treatment (as identified via the second bioinformatics method). Samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

## 4.3.2 Block effect

### *4.3.2.1 Original results*

There was no significant difference in homogeneity of variance between blocks ($F = 2.27$, $p = 0.088$). There was also no significant block effect in Bray-Curtis distance between communities (PERMANOVA: $R^2 = 0.36$, $p = 0.12$).

### *4.3.2.2 Kraken results*

There was no significant difference in homogeneity of variance between blocks ($F = 0.62$, $p = 0.682$). There was also no significant block effect in Bray-Curtis distance between communities (PERMANOVA: $R^2 = 0.25$, $p = 0.28$).

### 4.3.3 Spatial relationships and plot - non plot differences

#### *4.3.3.1 Original results*

There were significant differences between algal communities from the Whim experimental plots and from those sampled as non-plot controls (PERMANOVA $R^2$ = 0.24, p = 0.0158).

There was significant difference in homogeneity of variance between the experimental plots and the non-plot controls (F = 7.47, p = 0.016). Non-plot controls had significantly less variance- perhaps unsurprisingly as there were far fewer samples. Whilst PERMANOVA has an assumption of heterogeneity of variance, when the smaller sample size has lower variance, this is associated with PERMANOVA, like other multivariate tests, being overly conservative (Anderson and Walsh, 2013). Therefore, the significant difference between plot and non-plot communities is likely to be robust.

#### *4.3.3.2 Kraken results*

There were no statistically significant differences between algal communities from the Whim experimental plots and from those sampled as non-plot controls (PERMANOVA $R^2$ = 0.041, p = 0.535). However, once again there was significant difference in homogeneity of variance between the experimental plots and the non-plot controls (F = 7.05, p = 0.018). Non-plot controls had significantly less variance, unsurprisingly as there are far fewer samples (Fig. 4.15).

There was no significant correlation between algal dissimilarity and geospatial distance (Mantel r = -0.165, p = 0.917). Following removal of sample WNP2 (which had incorrect coordinates), this correlation remained non-significant (Mantel r = -0.059, p = 0.768).

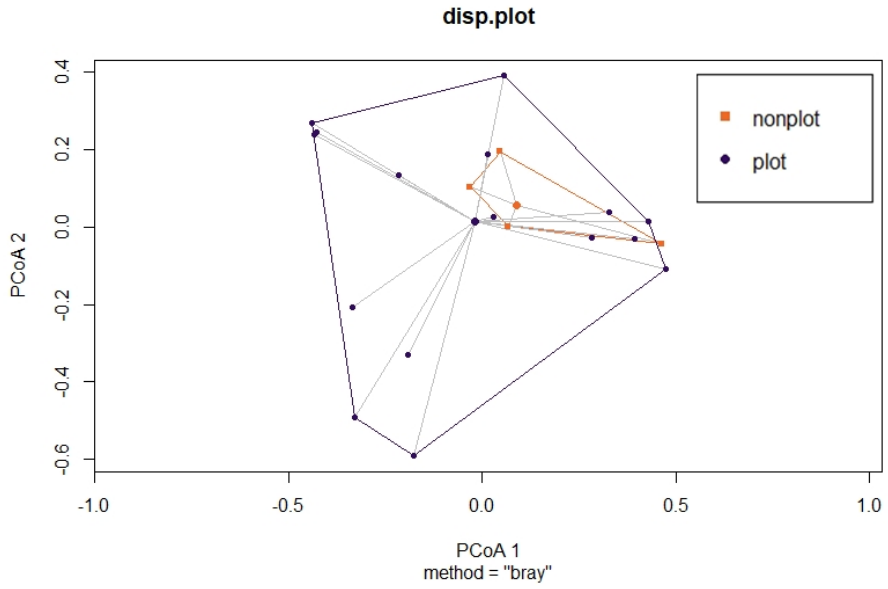



**Figure 4.15.** PCoA plot showing beta dispersal of algal community by plot/ non plot status (as identified via the second bioinformatics method). Plot samples are nutrient treatment plots, and nonplot samples are the experimental disturbance controls outside of the experimental area.

## 4.3.4 Taxonomic groups

### *4.3.4.1 Original results*

The majority of assigned sequence variants were unclassified at phylum level (Fig. 4.16). In order to better see any patterns, samples unclassified at phylum level were filtered out (Fig. 4.17).

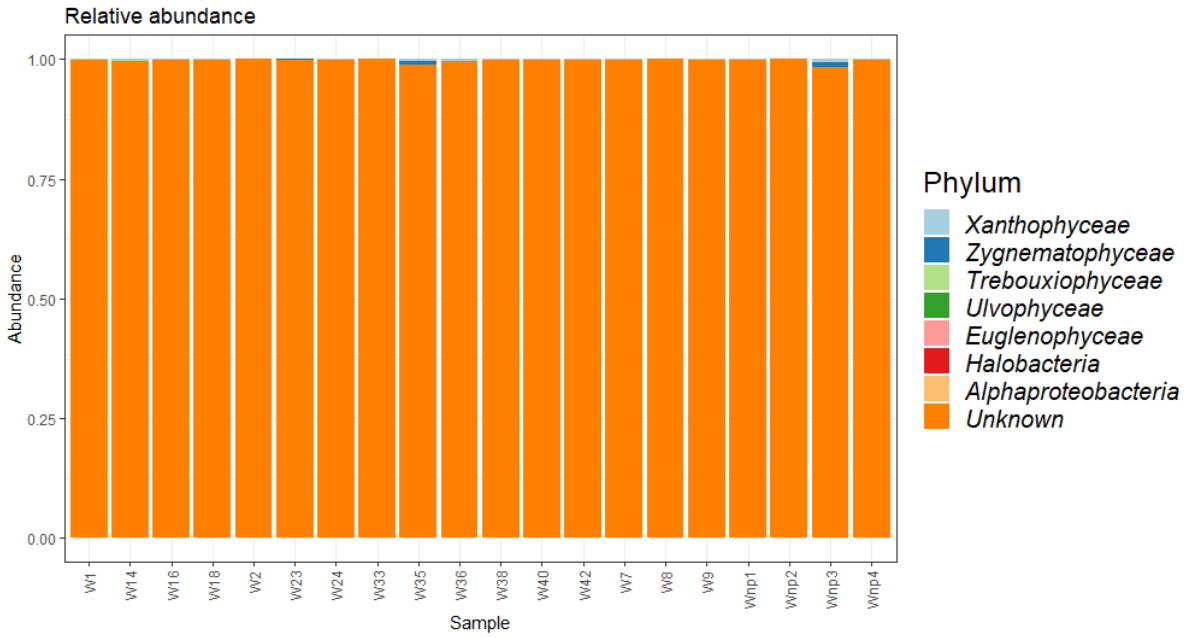



**Figure 4.16.** Relative abundance of dominant microbial phyla without filtering to remove ASVs unclassified at the phylum level (first bioinformatics method).

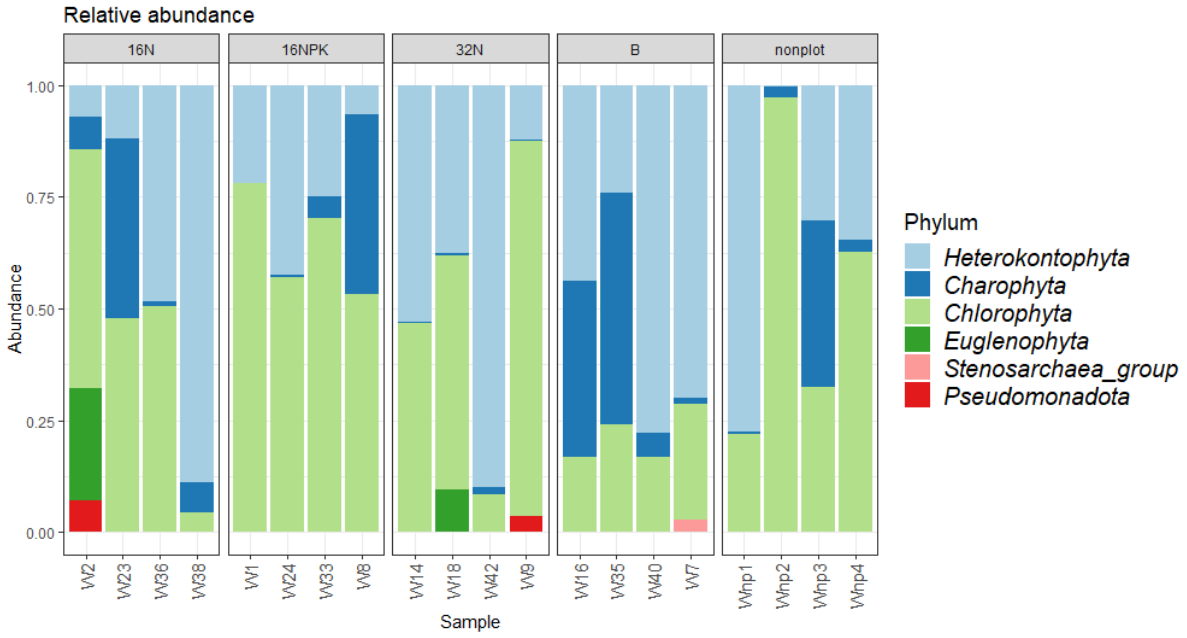



**Figure 4.17.** Relative abundance of dominant microbial phyla when ASVs unclassified to phylum were filtered out (first bioinformatics method). There was no clear effect of treatment on community composition. Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

Even with ASVs unclassified to phylum removed, a large proportion remained unclassified at 'class' level. There were no obvious patterns between treatments for ASVs identified to class (Fig. 4.18).

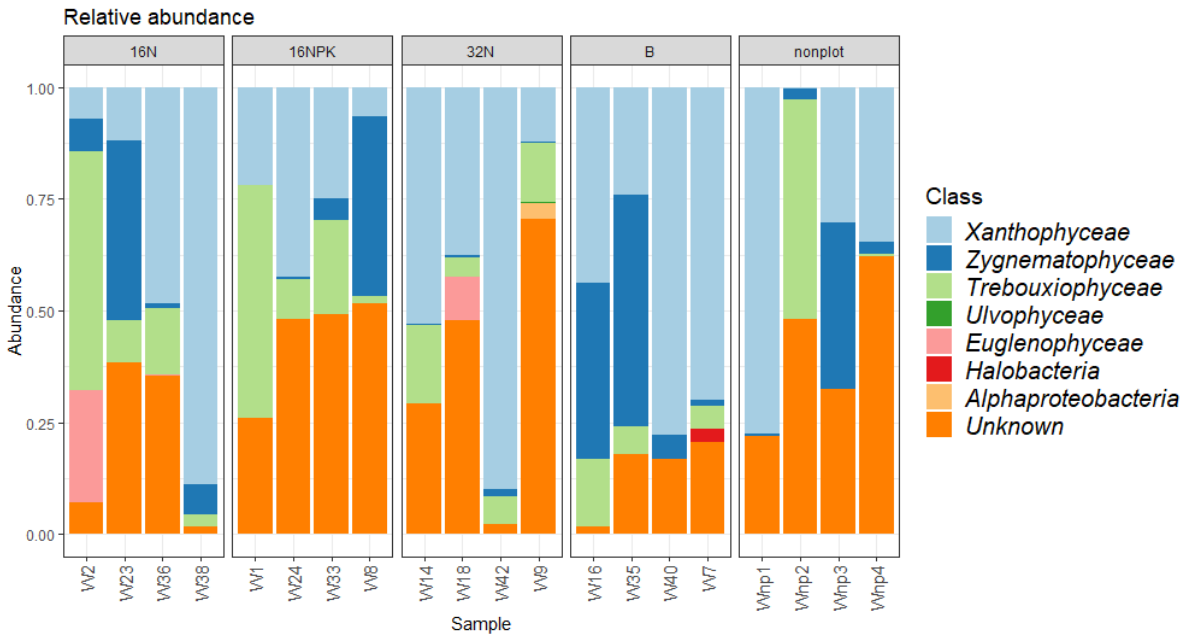



**Figure 4.18.** Relative abundance of dominant microbial classes when ASVs unclassified to phylum were filtered out (first bioinformatics method). Taxa which were unclassified at lower taxonomic levels than phylum were not removed, and these represent a large proportion of ASVs. Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and 'nonplot' are experimental disturbance controls.

### *4.3.4.2 Kraken results*

As a far greater proportion of reads were identified using the Kraken identification method, and Bracken only includes identified reads, it was possible to look at changes in taxonomic composition without having to filter out unidentified taxa. At the phylum level, the majority of algae detected belonged either to the Chlorophyta (including the Chlorophyceae and Trebouxiophyceae) or Streptophyta (including the Zygnematophyceae and Desmidials) (Fig. 4.19). Very few cyanobacteria were detected

(See Section 3.6 for details of non-cyanobacteria prokaryotic groups identified. These were excluded from this analysis due to uncertainty around functional roles). There were large amounts of within-treatment variation, but 32N (the highest N addition treatment) notably had a low relative proportion of Streptophyta in all plots.

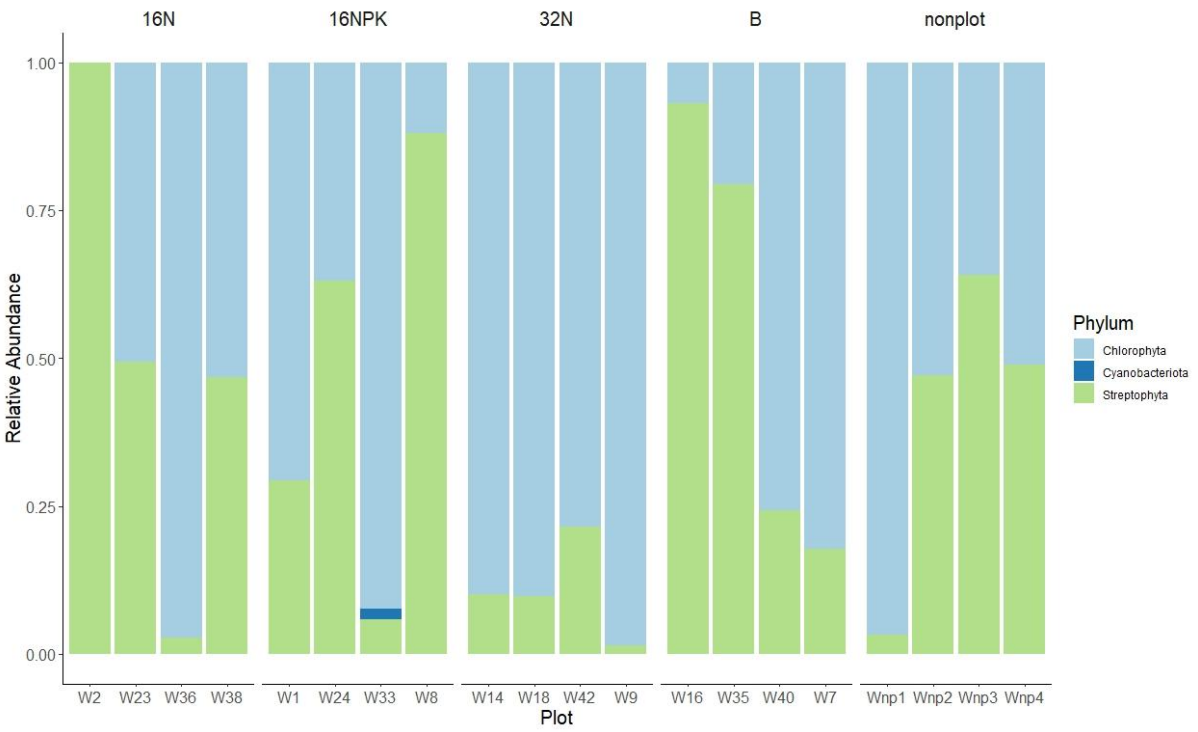



**Figure 4.19.** Relevant abundance of microbial phyla by nutrient addition treatment following the second bioinformatics method (taxonomic assignment with Kraken2). Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

When agglomerated at the class level, the Zygnemophyceae (Streptophyceae) again had low relative abundance in the 32N treatment (Fig. 4.20). Whilst all treatments had at least one plot with a lower relative abundance of Zygnemophyceae than the highest % relative abundance for 32N, it was the only treatment to have no plots with a relative abundance of Zygnemophycea above 50%. 32N and 16NPK treatments also seemed to have a higher relative abundance of Trebouxiophyceae.

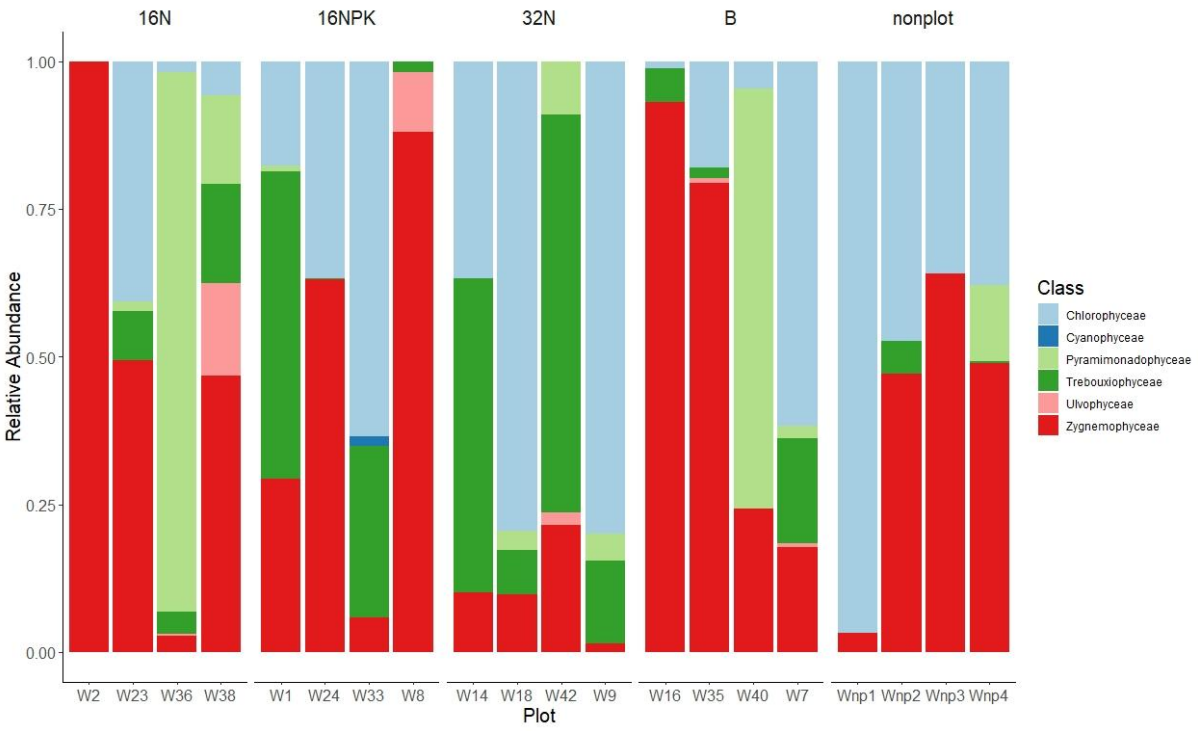



**Figure 4.20.** Relative abundances of the algal community by nutrient addition treatment agglomerated at the class level (second bioinformatics method). Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

When agglomerated to the order level, 32N appeared to have a greater relative abundance of Trebouxiales than all other treatments (Fig. 4.21). 32N and 16NPK also appeared to have lower relative abundance of Desmidiales, although the within-treatment variation was high, as it was for all other taxonomic groups. In general, the eukaryotic algal community appeared to be extremely heterogeneous, even at this relatively low taxonomic resolution.

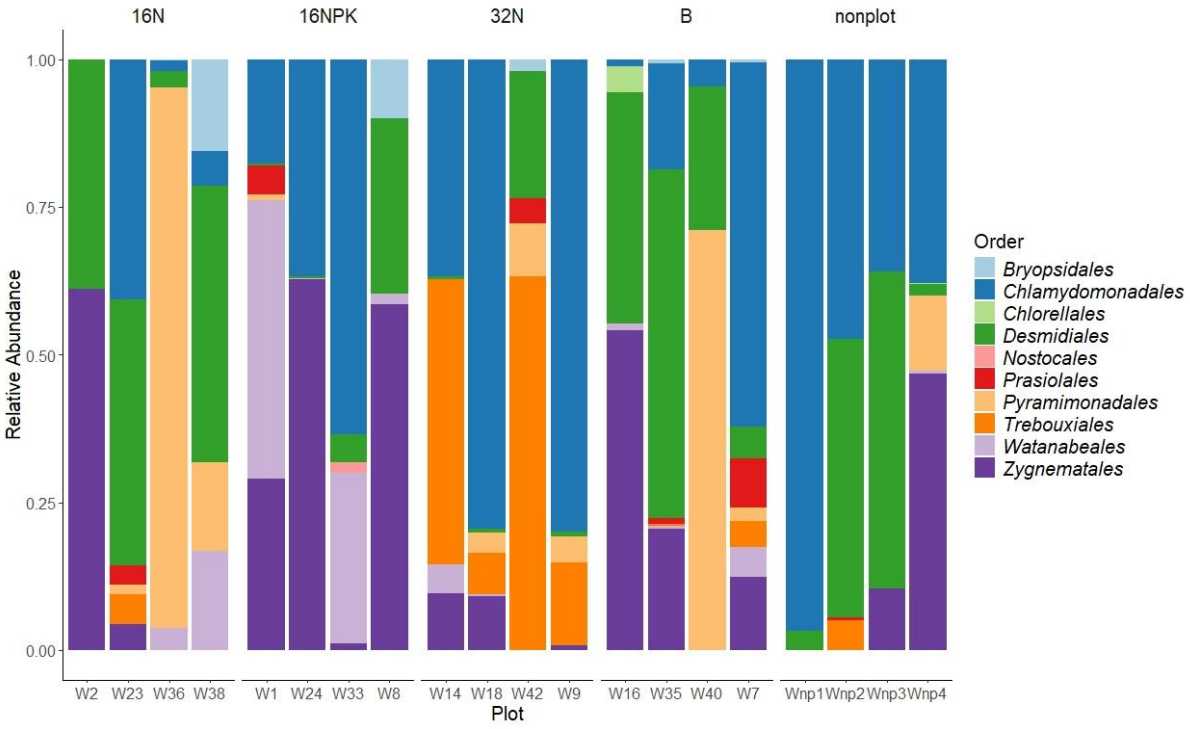



**Figure 4.21.** Relative abundances of the most common orders in the microbial community by nutrient addition treatment (second bioinformatics method). Rare taxa were aggregated at prevalence = 10/100, detection = 0.1/100. Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

## 4.3.5 Prokaryote to eukaryote ratio

### *4.3.5.1 Original results*

The majority of sequences were for eukaryotes, with Prokarya making up a very small percentage of total reads (Fig. 4.22). When filtered to only plot reads identified to phylum, it becomes apparent that proportion of identified prokaryotes was very variable between samples, with no observable treatment patterns (Fig. 4.23).

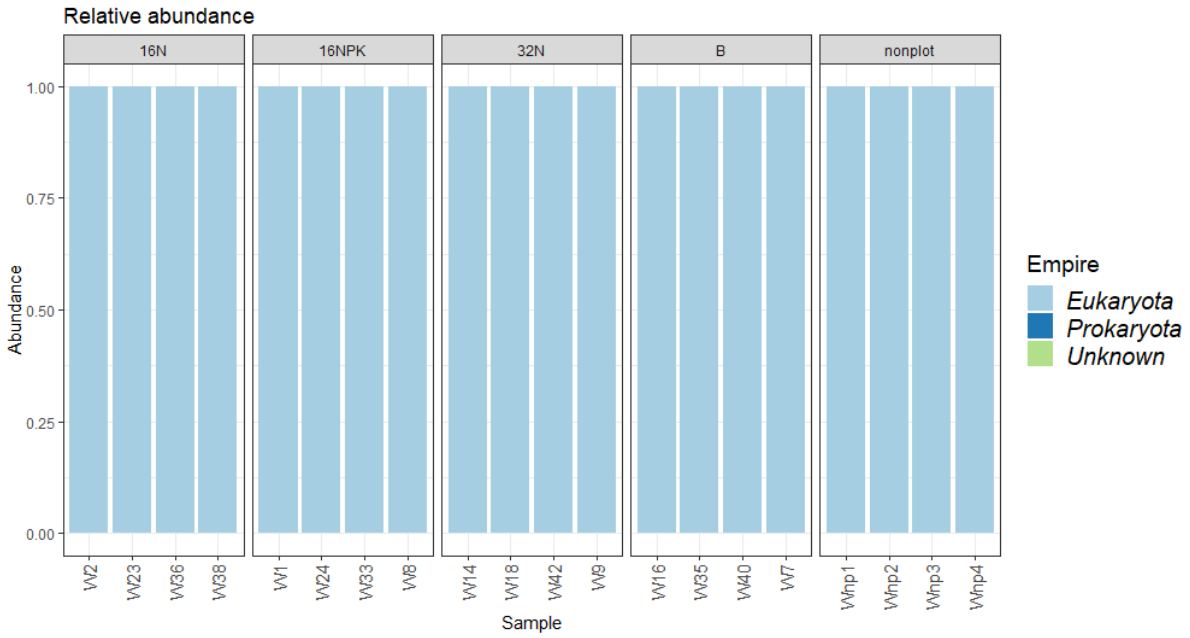



**Figure 4.22.** Relative abundance of Eukaryote vs Prokaryote ASVs, with no filtering to remove ASVs unassigned at any taxonomic level (first bioinformatics method). Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

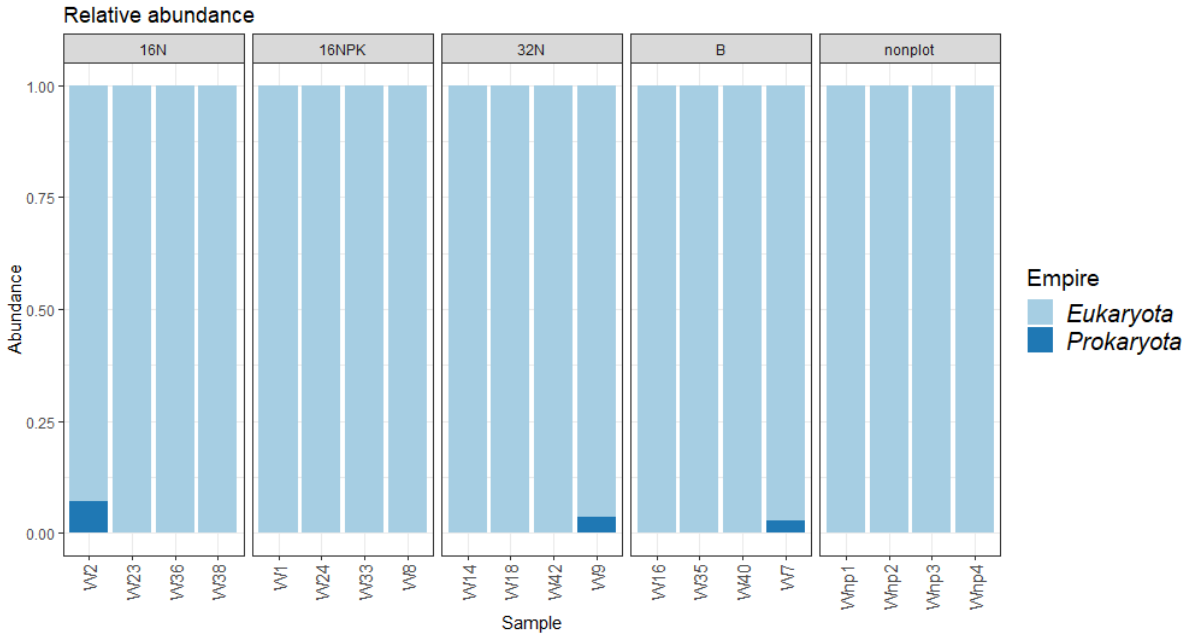



**Figure 4.23.** Relative abundance of Eukaryote vs Prokaryote ASVs, with ASVs not classified to phylum level removed (first bioinformatics method). Labels above the plot

show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

No prokaryote reads were identified as belonging to cyanobacteria. Instead, those which were classified as prokaryotes and had a higher taxonomic classification were anoxic phototrophs such as Halobacteria and Alphaproteobacteria.

To see if ASV assignment could be improved, taxonomic assignment was repeated with reverse complement matching for each sequence allowed. However, this resulted in 34 distinct ASVs identified as *Sphagnum* at 97% clustering, as well as far fewer ASVs being identified as non-plants. Therefore, the original taxonomic assignment without reverse complement checking was retained for all further analysis.

When sequences not assigned to either eukaryote or prokaryote at empire level are removed, relative abundance of prokaryote reads could be calculated. The 32N treatment had a higher relative abundance of prokaryotes than all other treatments (Fig. 4.24). However, this difference was not statistically significant (Kruskal-Wallis chi-squared = 4.299, df = 4, p = 0.367). This was likely due to the low proportion of prokaryote reads. Use of primers better suited to prokaryotes may have helped to correct this, as the rbcL primers were developed for use with barcoding of eukaryotes. Certainly, these barcodes seem to be unsuitable for identification of cyanobacteria, as these would be expected to be present at detectable levels in the DNA sample from this site had they been successfully amplified.

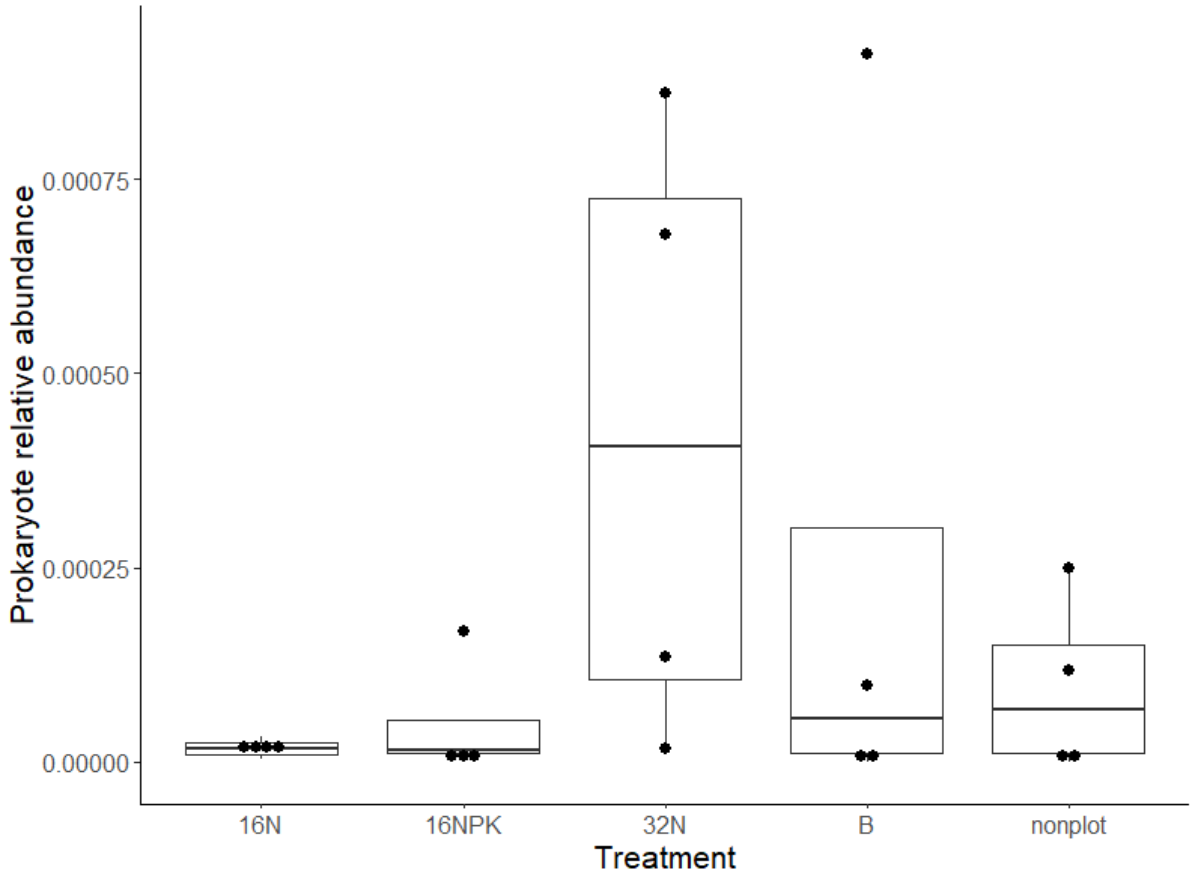



**Figure 4.24.** Proportion of prokaryote ASVs (of those assigned to empire level) from the first bioinformatics method. Boxes show median and interquartile range. Points show actual data. The treatment labelled B is background deposition control plots and nonplot is the experimental disturbance controls.

### *4.3.5.2 Kraken results*

There was no significant difference between proportion of microbial reads which were prokaryotic between nutrient treatments (Kruskall-Wallis: chi-squared = 5.99, df = 4, p = 0.2002). Whilst the 32N treatment (the highest N input) initially appeared to have a greater proportion of prokaryotes, the violin plot shows this to be due to only two samples (Fig. 4.25). The low read count for prokaryotes across all samples (likely due to poor primer matching for cyanobacteria – see Section 3.4) likely skews this data too much to draw conclusions on the proportion of prokaryotes. All treatments have at least one sample with no reads identified as prokaryotic (Fig. 4.26). It is worth noting that the Kraken2 method of taxonomic identification gives a different pattern for this than the

DADA2 custom database method. This is likely due to the high number of reads for the initial bioinformatics pipeline which were not classified to prokaryote/eukaryote level.

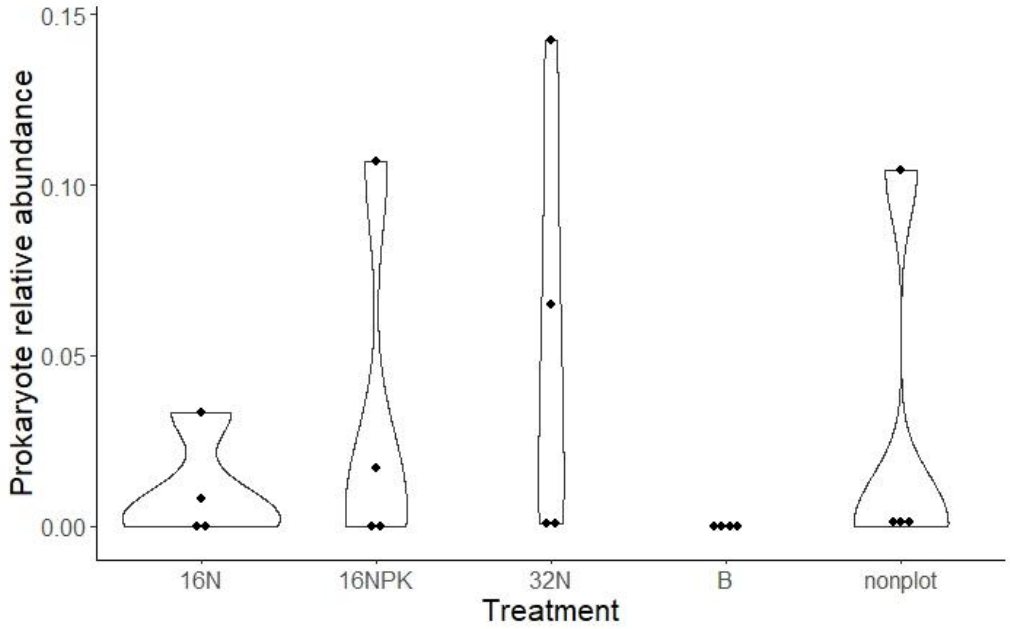



**Figure 4.25.** Violin plot showing proportion of microbial reads assigned to prokaryotes by nutrient addition treatment (second bioinformatics method). The treatment labelled B is background deposition control plots and nonplot is the experimental disturbance controls.

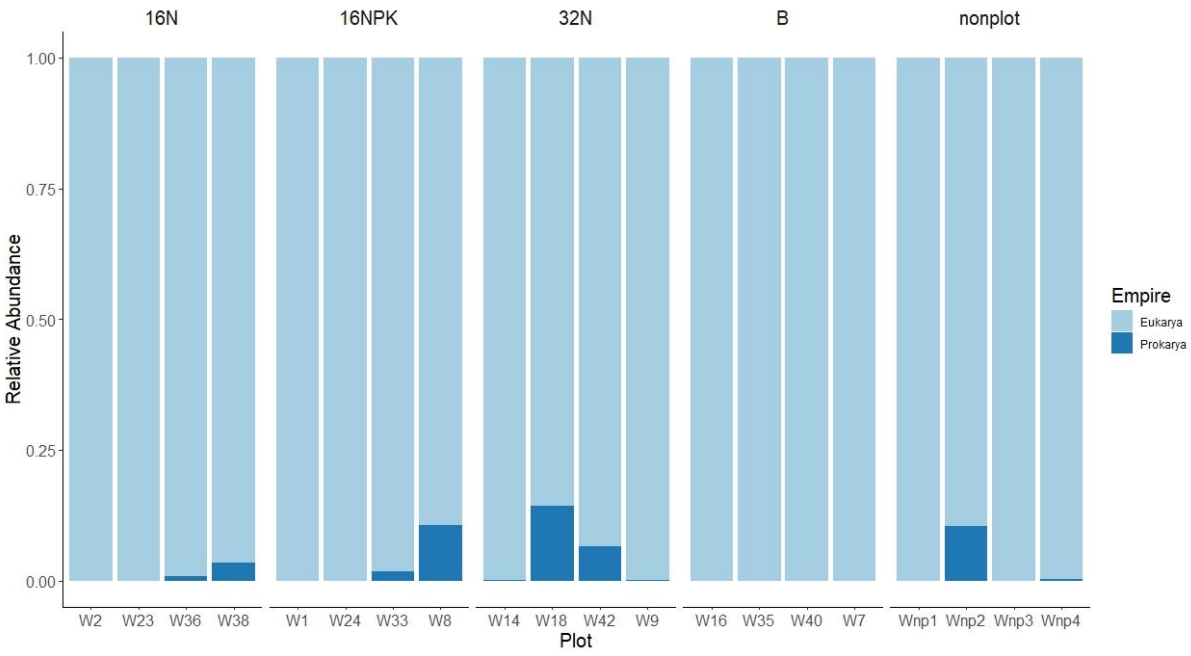



**Figure 4.26.** Relative abundance of microbial reads assigned to Eukaryota and Prokaryota by nutrient treatment (second bioinformatics method). Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

Of the prokaryotic reads identified, the most widespread was Pseudomonadata. Only the sample for plot W33 had reads assigned to cyanobacteria (Fig. 4.27).

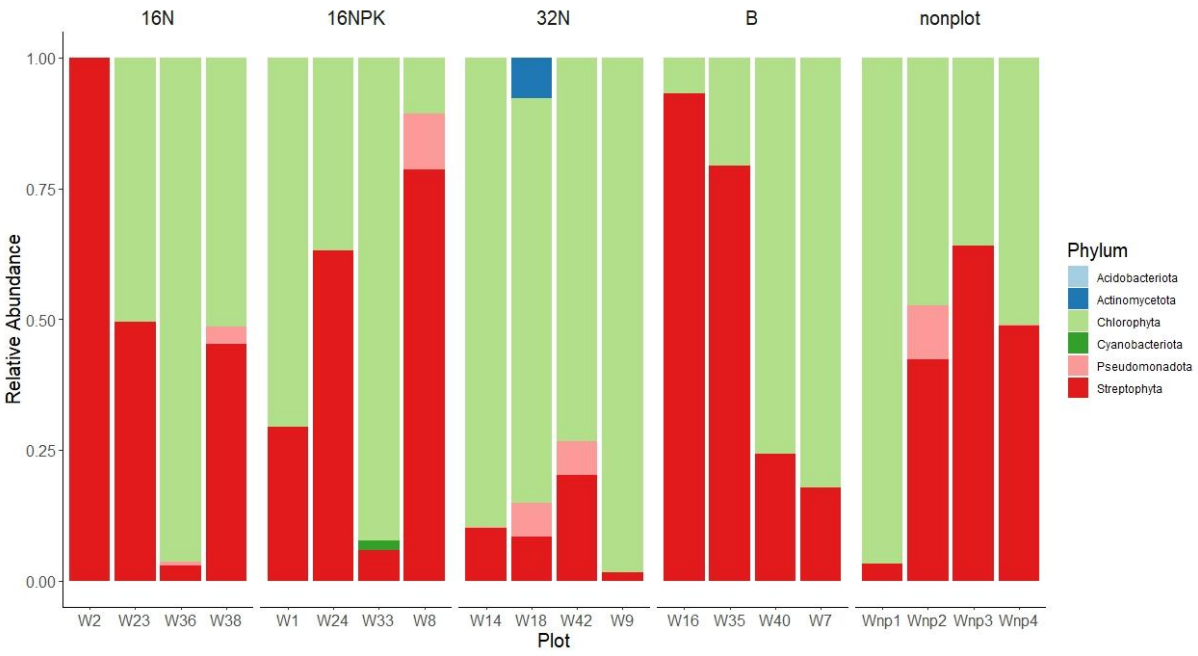



**Figure 4.27.** Composition plot of relative abundance of microbial reads agglomerated at the phylum level (second bioinformatics method). Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

### 4.3.6 Biodiversity indices

There are suggestions of a pattern in Simpson's Biodiversity between treatments, with plots where 16kg N and 16kg N plus PK have been added having higher biodiversity than either controls (background and non-plot) or plots with higher level of N addition (Fig. 4.28). However, these trends are not statistically significant (Kruskal-Wallis rank sum test: $p = 0.5845$, $H = 2.84$, $df = 4$).

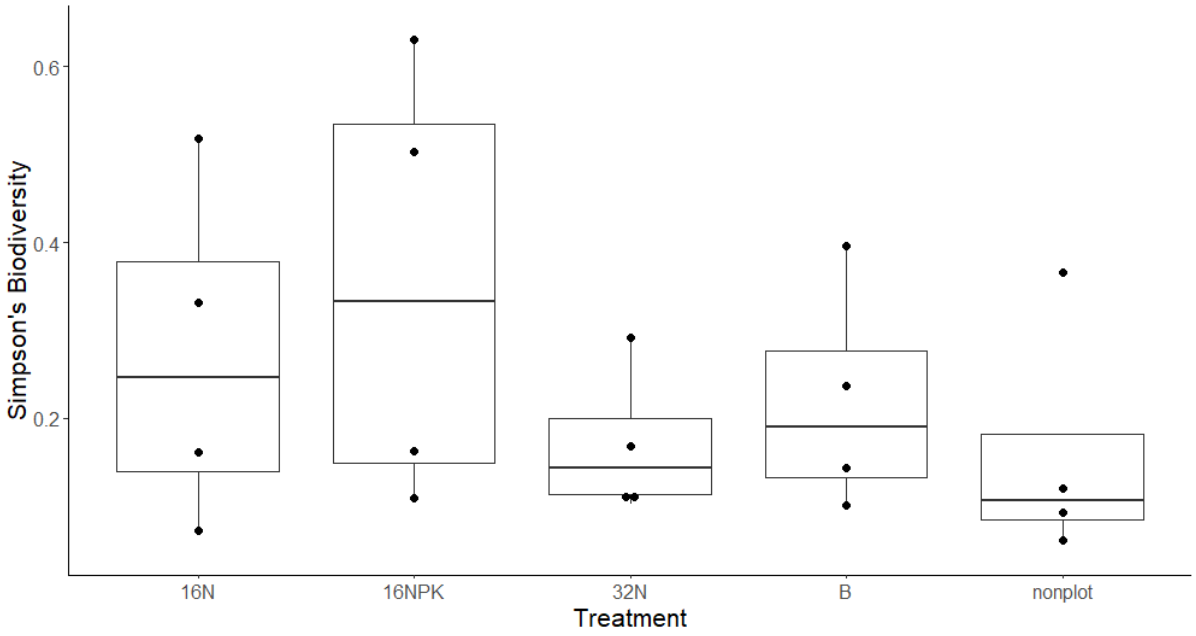



**Figure 4.28.** Simpsons biodiversity index by nutrient addition treatment and non-plot controls (first bioinformatics method). Treatment B is background deposition control plots and nonplot is experimental disturbance controls.

## 4.3.7 Plant community changes – Kraken2 bioinformatics

### *4.3.7.1 Proportion of Sphagnum reads*

For all samples, the number of reads was stable from the phylum level (Sphagnophytina) to the genus level (*Sphagnum*). Read counts and percentages then sharply dropped off at the species level (Table 4.2). Note that W40 also had reads assigned to *Ambuchanania leucobryoides*, a monophyletic Sphagnale endemic to Tasmania (Johnson *et al.*, 2008). This is likely a misassignment, with the extremely restricted range and habitat requirements of this species, alongside no recorded occurrences outside of its native range, making its presence as a neophyte at this site very unlikely. *Sphagnum ovatum*, which a low percentage of reads in a single sample were also assigned to, is a Himalayan species and similarly unlikely to be present (Eddy, 1977). Whilst *S. contortum* and *S. cuspidatum* are both UK species, it is unlikely that they were misclassified in the field as *S. capillifolium*. Therefore the taxonomic assignments at species level should be treated with caution.

**Table 4.2.** Percentage of reads assigned to the genus *Sphagnum* and species within it by sample.

| Sample | Percentage of reads assigned to genus *Sphagnum* | Percentage of reads assigned to species |
|---|---|---|
| W1 | 85.61 | *Sphagnum contortum* 0.27 |
| W2 | 97.05 | *S. contortum* 0.15 |
| W7 | 54.9 | *S. contortum* 0.17 |
| W8 | 98.64 | *S. contortum* 0.19<br>*S. cuspidatum* 0.1 |
| W9 | 92.64 | *S. contortum* 0.22<br>*S.* cuspidatum 0.1 |
| W14 | 90.21 | *S. contortum* 0.25 |
| W16 | 98.08 | *S. contortum* 0.27<br>*S. cuspidatum* 0.05 |
| W18 | 83.8 | *S. contortum* 0.22 |
| W23 | 93.4 | *S. contortum* 0.25 |
| W24 | 95.09 | *S. contortum* 0.27<br>*Sphagnum palustre* <0.01<br>*Sphagnum palustre var. imbricatum* <0.01 |
| W33 | 98.4 | *S. contortum* 0.2 |
| W35 | 63.17 | *S. contortum* 0.73<br>*S. cuspidatum* 0.11 |
| W36 | 89.39 | *S. contortum* 0.27<br>*S. cuspidatum* 0.16 |
| W38 | 97.45 | *S. contortum* 0.27<br>*S. cuspidatum* 0.01 |
| W40 | 94.09 | *S. contortum* 0.22<br>*S. cuspidatum* 0.05 |
| W42 | 88.13 | *S. contortum* 0.26<br>*S. ovatum* 0.01 |
| WNP1 | 94.41 | *S. contortum* 0.28<br>*S. cuspidatum* 0.14 |
| WNP2 | 84.88 | *S. contortum* 0.24 |
| WNP3 | 72.5 | *S. contortum* 0.39<br>*S. cuspidatum* 0.07 |
| WNP4 | 73.59 | *S. contortum* 0.21 |

### *4.3.7.2 Total plant community*

The majority of plant reads detected are unsurprisingly Spagnopsida (see above), as all samples were taken from *Sphagnum capillifolium* mounds as individual capitula. There would therefore have been limited opportunities for other macroscopic plant litter to be incorporated into the sample. However, the community identified shows that this method has also been successful in collecting DNA from other plants from the surrounding area. Whilst there is likely to be some misidentification at the species and genus level, at a lower taxonomic resolution this may still provide valuable insights.

Outside of Sphagnopsida, the most abundant plant classes were Magnoliopsida and Jungermanniopsida (Fig. 4.29). Polypodiopsida, Polytrichopsida, Bryopsida and Lycopodiopsida were also a major contributor to the non-Sphagnopsida relative abundance in some plots. Outside of *Sphagnum* (the host DNA), the majority of plant community reads appear to be vascular for nutrient treatment plots (16N, 16NPK and 32N) and non-vascular for the control plots (background and non-plot).

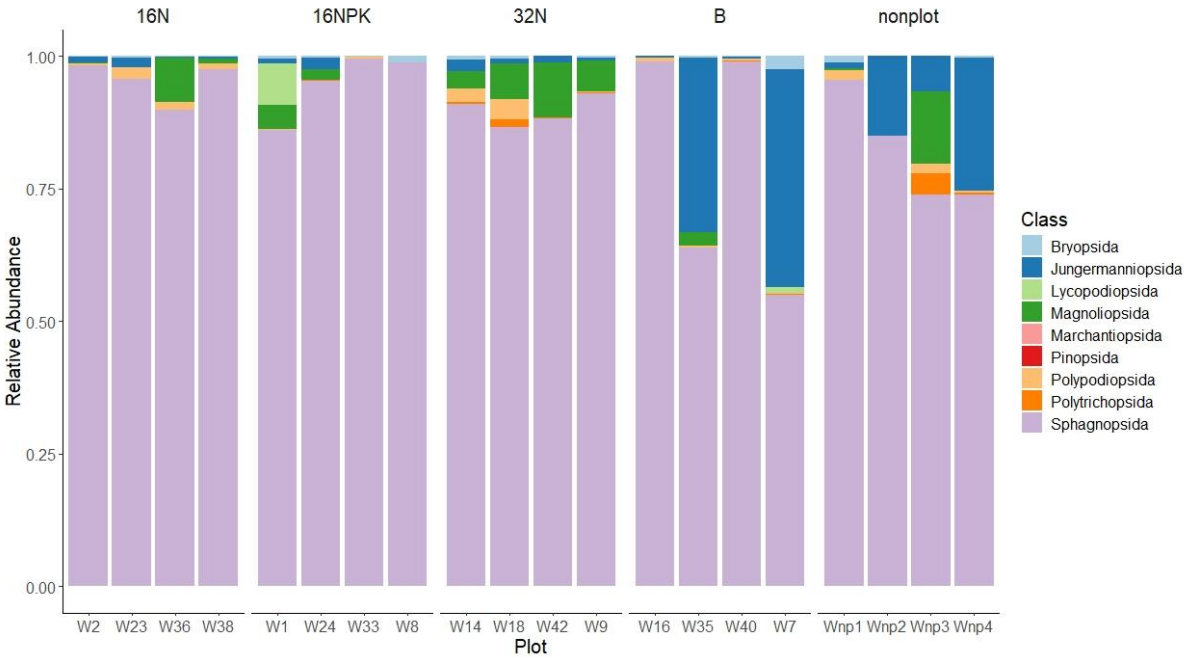



**Figure 4.29.** Relative abundance of the classes of vascular and non-vascular plants (second bioinformatics method). Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

There was no significant difference between plant communities in the different treatments (PERMANOVA: $R^2 = 0.37$, $p = 0.058$), but there was a significant difference in beta dispersal ($F = 10.42$ , $p = 0.001$), with all nutrient addition treatments clustering together (Fig. 4.30). There was no significant block effect (PERMANOVA: $R^2 = 0.23$, $p = 0.358$) or experimental plot effect (PERMANOVA: $R^2 = 0.075$, $p = 0.148$) for the combined plant community.

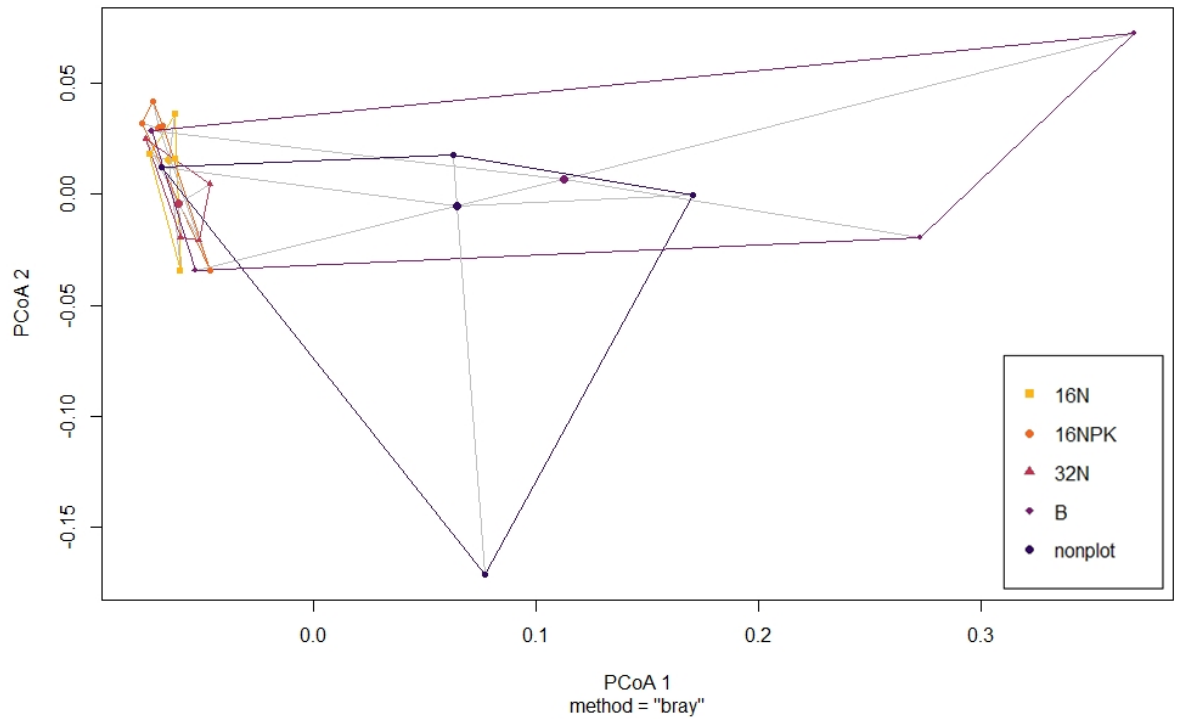



**Figure 4.30.** PCoA plot showing combined plant community beta dispersal by different nutrient addition treatments (second bioinformatics method). All nutrient addition treatments were clustered together, while background N deposition plots and non-plot controls had greater variance in the plant community. The treatment labelled B is background deposition control plots and nonplot is experimental disturbance controls.

### *4.3.7.3 Non-vascular plants*

For the non-vascular plant community, the treatments without nutrient abundance (background/ nutrient control, and non-plot/ experimental disturbance control) had higher relative abundance of Juggermaniopsida reads, although there was still some within treatment variation (Fig. 4.31). The nutrient addition treatments, however, saw

much lower within- treatment variation, and the vast majority of reads for all samples across all three nutrient addition treatments were Sphagnopsida.

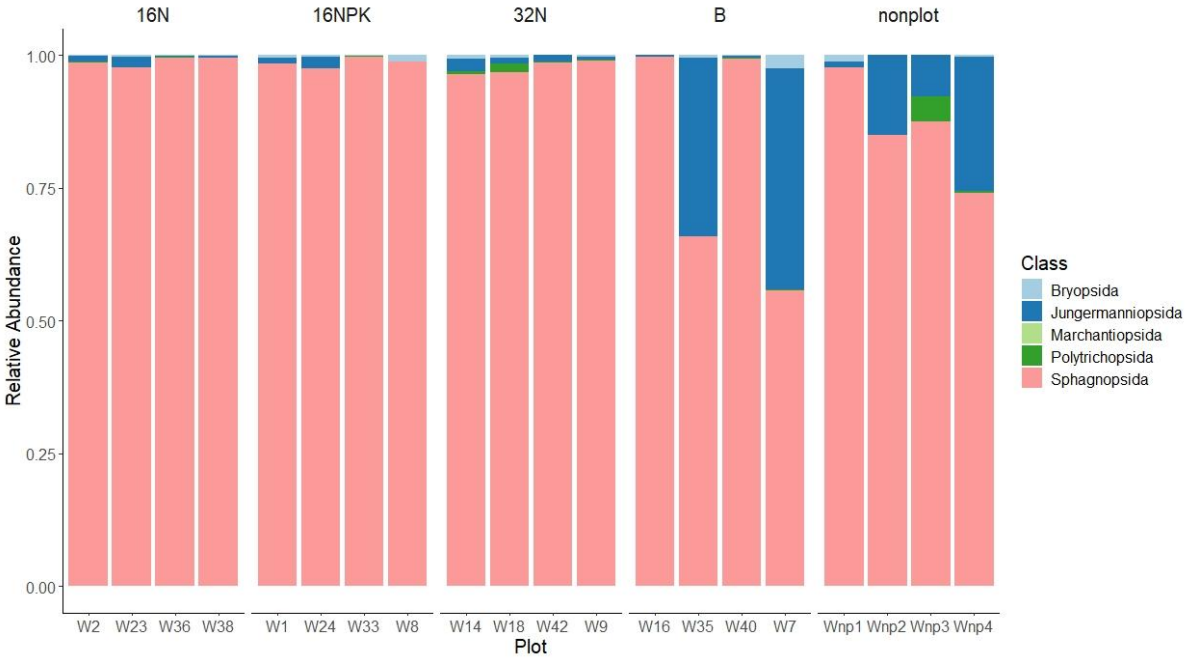



**Figure 4.31.** Relative abundane of the classes of non-vascular plants by nutrient addition treatment (second bioinformatics method). Labels above the plot show the N addition treatment: samples labelled B are background deposition control plots and nonplot are experimental disturbance controls.

There were significant differences between non-vascular plant communities in the different treatments (PERMANOVA: $R^2 = 0.405$, $p = 0.045$). There was also a significant difference in beta dispersal ($F= 6.10$, $p= 0.001$), with all nutrient addition treatments clustering together (Fig. 4.32). There was no significant block effect (PERMANOVA: $R^2 = 0.19$, $p= 0.461$) or experimental plot effect (PERMANOVA: $R^2 = 0.064$, $p= 0.178$) for the nonvascular plant community.

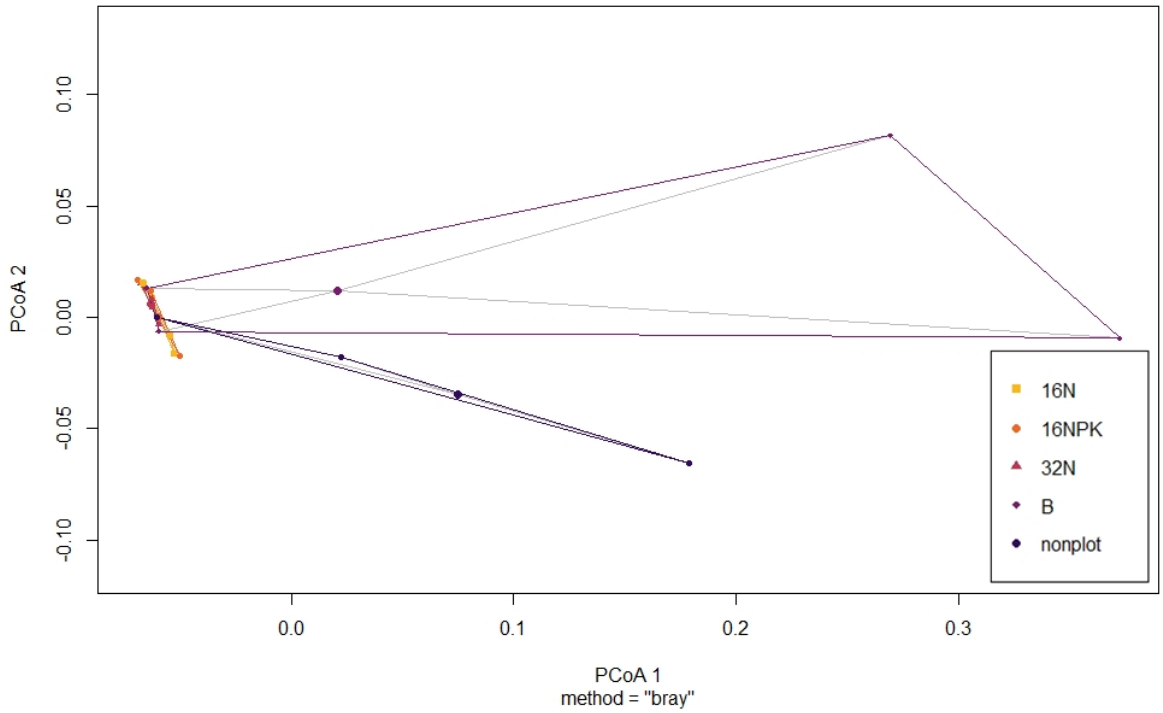



**Figure 4.32.** PCoA plot of nonvascular plant beta dispersal by nutrient addition treatment. Samples labelled B are background deposition control plots and nonplot are experimental disturbance controls. Data from the second bioinformatics method.

Detrended correspondent analysis shows clustering of all nutrient treatment plots (Fig. 4.33), with background and non-plot control varying along DCA1 (Fig. 4.34).

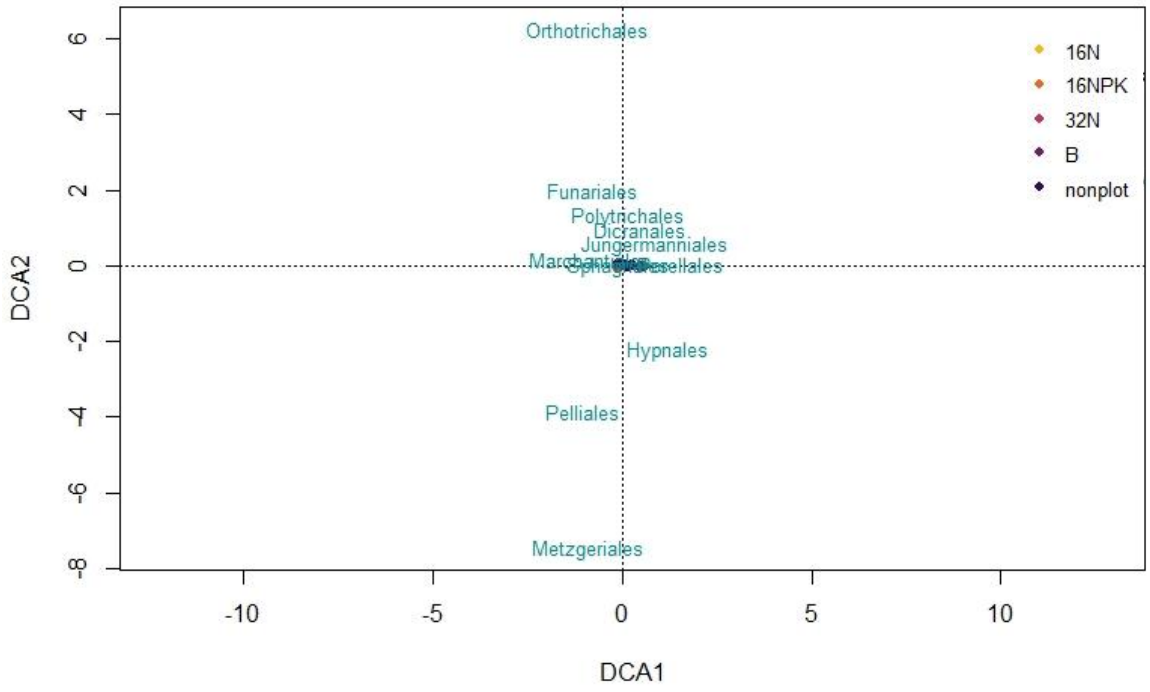



**Figure 4.33.** DCA plot of bryophyte community by treatment. Samples labelled B are background deposition control plots and nonplot are experimental disturbance controls. Data from the second bioinformatics method.

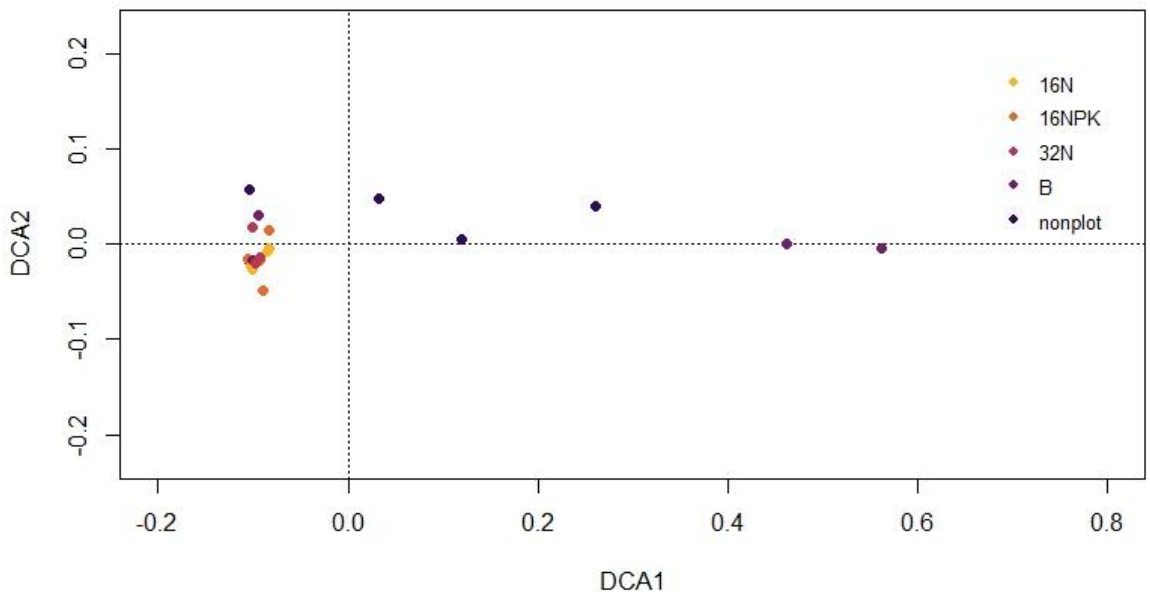



**Figure 4.34.** DCA plot of bryophyte community by nutrient addition treatment, zoomed to clearly show plot points. Samples labelled B are background deposition control plots and nonplot are experimental disturbance controls. Data from the second bioinformatics method.

### *4.3.7.4 Vascular plants*

There was a large amount of within – treatment variation in relative abundances of different vascular plant classes. The majority of reads were either Magnoliopsida or Polypodiopsida (Fig. 4.35). Magnoliopsida appears to be more dominant at the 32N treatment than the other treatments or controls.

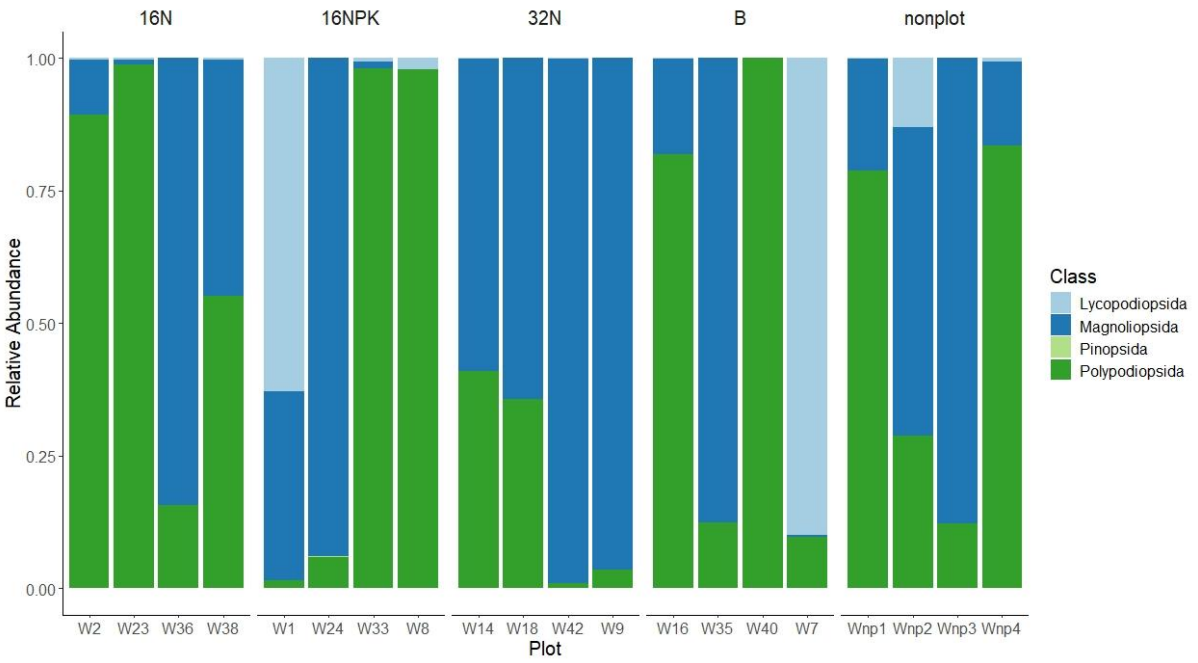



**Figure 4.35.** Relative abundance of the classes of vascular plants by treatment. Treatment B is the background deposition control plots and nonplot is the experimental disturbance controls. Data from the second bioinformatics method. Pinopsida was present but at very low relative abundance (<0.1% in sample W24), and so cannot be seen within the plot,

There were no significant differences between vascular plant communities in the different treatments (PERMANOVA: $R^2 = 0.17$, $p = 0.749$). There was no significant difference in beta dispersal ($F = 2.17$, $p = 0.126$) (Fig. 4.36). There was no significant block effect (PERMANOVA: $R^2 = 0.23$, $p = 0.325$) or experimental plot effect (PERMANOVA: $R^2 = 0.018$, $p = 0.947$) for the vascular plant community. Detrended correspondence analysis does not show any obvious clustering by treatment (Fig. 4.37). There appears to be possible clustering by class, however, some of the classes are likely to have been misassigned, so care should be taken not to over-interpret this. Zingiberales, for example, made up 0.02% of the reads in sample W2, with Kraken failing to assign these reads past the order level. The Zingiberales (which contains the ginger, canna and banana families, amongst others (Stevens, 2017; Baker *et al.*, 2002; Kress, Prince and Williams, 2002) are listed as a neophyte in the Botanical Society of Britain and Ireland's taxon list (BSBI, 2025a), with *Zingiber mioga* mentioned in particular, but

there are no records of this species or any other species in the Zingiberales in the UK on the BSBI taxon search record (BSBI, 2025b), iNaturalist (2025) or GBIF (2025), so its presence as a widespread neophyte is unlikely. *Zingiber mioga* has never been recorded in Scotland, and was not found on previous surveys of the site by the BSBI County recorder (Luke Gaskell, 2025, pers. comm., August 7). Instead, longer distance pollen dispersal from cultivation is the most likely source of Zingiberales DNA, with the alternative being incorrect taxonomic assignment by Kraken. Either way, the presence of the order at the sampling site should be treated with caution.

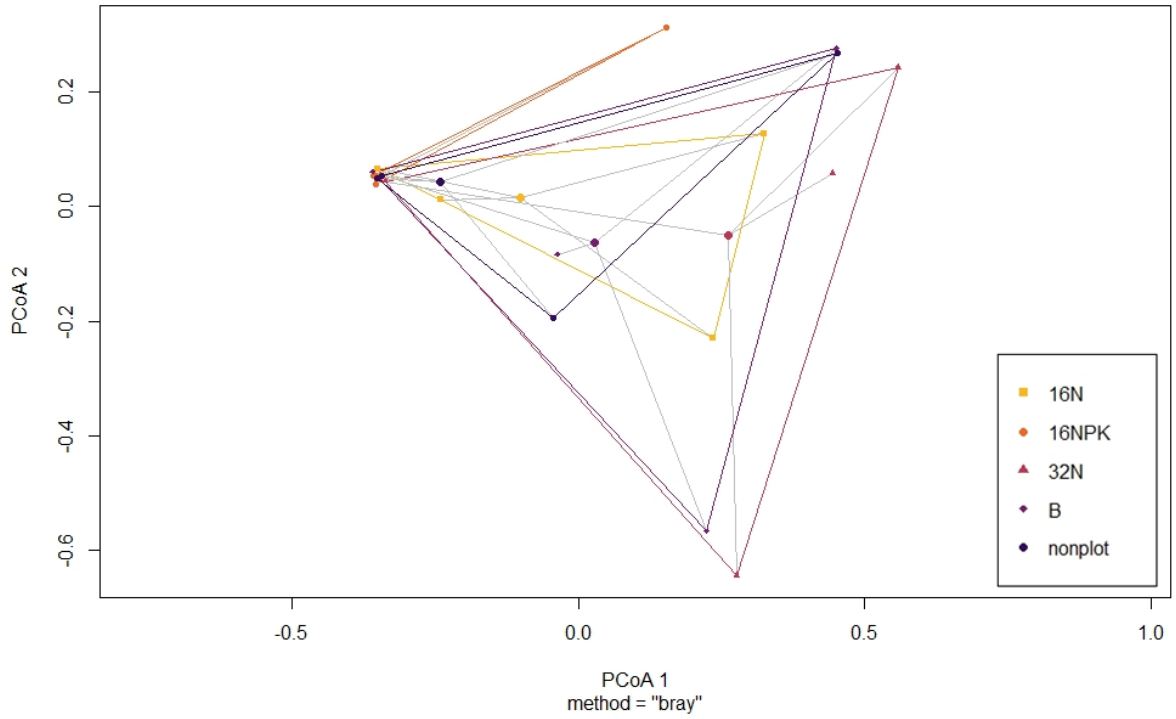



**Figure 4.36.** PCoA plot of vascular plant community beta dispersal by treatment. Treatment B is the background deposition control plots and nonplot is the experimental disturbance controls. Data from the second bioinformatics method.

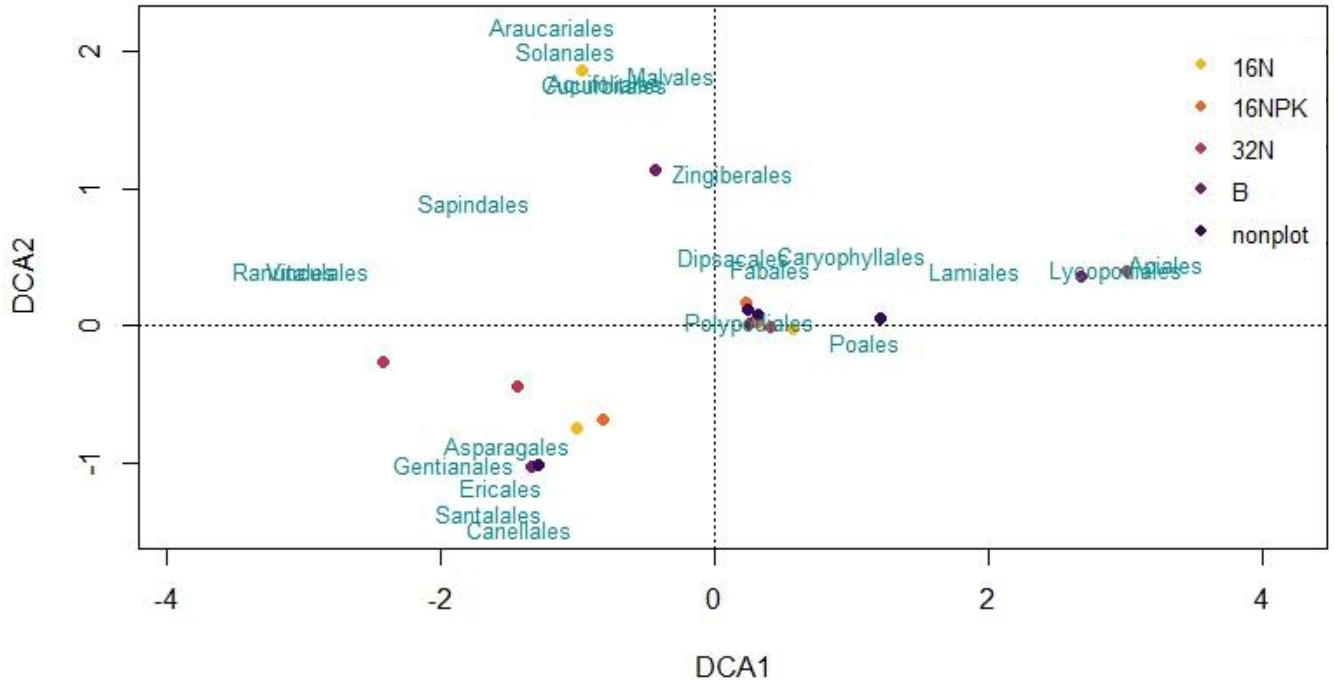



**Figure 4.37.** DCA plot of vascular plant community by treatment. Treatment B is the background deposition control plots and nonplot is the experimental disturbance controls. Data from the second bioinformatics method.

## 4.4 Discussion

### 4.4.1 Algal community

The lack of significant differences between nutrient treatments is surprising, running counter to that expected based on previous studies (e.g. Payne *et al.*, 2013), suggesting an unexpected resistance to disturbance from N, P and K deposition for peatland algae at this background deposition level where there is *Sphagnum capillifolium* present. It is possible that any nutrient effects on photosynthetic microbial communities (at least at this background deposition level) are likely indirect, and driven by vegetation changes (and, most probably, *Sphagnum* survival). Care should be taken extrapolating these results to areas where background deposition levels are lower (for example the Flow

Country of Caithness and Sutherland), as they may have algal community comprising of species with greater N and P sensitivity.

The high number of unclassified ASVs seen when taxonomic assignment was carried out using the custom database in DADA2 was not unexpected, especially as there are fewer databases for rbcL than for example 16S primers. This would have provided another explanation for the initial results contradicting the literature. Using a phylogenetically informed distance matrix such as Weighted Unifrac, rather than Bray-Curtis distance, would allow the distance matrix to consider with much greater accuracy those ASVs which could not be assigned to a taxonomic rank, as relatedness between unidentified ASVs could still be calculated and taken into account. However, the decision was made not to construct phylogenetic trees for this data (see Section 3.4). The low proportion of unidentified reads from the Kraken method suggest that a phylogenetic approach was not necessary, and suggest the results are an accurate portrayal of lack of change in the communities of algae which can be detected with this primer set.

Questions of abundance changes in cyanobacteria cannot be answered due to lack of ASVs identified as cyanobacteria. Considering their abundance and symbiotic relationship with *Sphagnum* (Carrell *et al*., 2022), it seems unlikely that they were not present. Issues with the reference database are unlikely to be the cause of this, as if the reference database used did not allow the algorithm to identify cyanobacterial reads, it would be expected that the Kraken2 bioinformatics pipeline, with the much larger reference database, would have had improved results. A literature search did not identify any papers which discussed the effectiveness of rbcL primers for cyanobacteria, and *in silico* PCR showed poor performance of this primer set for cyanobacteria (see Section 3.6). Different primer sets, or a multiple primer set approach, would be more suitable for investigating combined eukaryotic and prokaryotic algal communities. This would also allow the inclusion of missing eukaryotic algae taxa, e.g. diatoms and other algal groups with chloroplasts derived from Rhodophyta.

Whilst there were no overall significant differences in Bray-Curtis distances between communities, the composition plots highlight patterns for certain taxonomic groups which are of interest. The Streptophyta to Chlorophyta ratio appears to shift with nutrient treatment. However, there are no clear generalisable traits described in the literature belonging to these two large and extremely varied groups of green algae. Therefore, we must look at the taxa which contributed to this shift.

The reduction in relative abundance of Zygnemophyceae is unsurprising: desmids have long been recognised as characteristic of low nutrient environments (Prescott, 1948), and many desmids have been recorded as being sensitive to nutrient deposition (Olrik, 2003; Coesel, 1982). The difference in community between 32N and the other treatments suggests that at these addition levels nitrogen is more important than P or K, and that, at least in the presence of host *Sphagnum*, these classes as a whole may be less sensitive to lower levels of N deposition than expected. Of course, sensitivity is likely to vary at the species level, and so looking at higher taxonomic levels means that changes at the species level are missed. Nevertheless, detecting shifts at higher taxonomic levels suggests possible opportunities for monitoring, and also allows for better understanding of community dynamics to extrapolate into understanding and modelling of functions.

The increased relative abundance of Trebouxiales in the highest nitrogen addition treatments is more surprising. Kraken assigned many reads to the Trebouxiophyceae, and the genus *Trebouxia*. Species identified included *Trebouxia* sp. I19, a lichen photobiont (Kosecka *et al.*, 2022). However, care should be taken with species level taxonomic assignments; another species identified was Trebouxiophyceae sp. MX-AZ01, an extremophilic strain isolated from a geothermal spring in Mexico (Servín-Garcidueñas and Martínez-Romero, 2012). Whilst it is possible that this species has a wide and varied range, it seems more likely that it is instead misidentification of a related species. However, this change could be indicative of a shift in lichens from *Cladonia* spp. to more nitrophilous species with high N deposition recorded in the long-term vegetation monitoring from the site (Levy *et al.*, 2019).

Trebouxiophyceae are also notable as the symbiotic algae associated with mixotrophic testate amoeba. For plot W16, 64/81 reads identified as Trebouxiophyceae were identified as *Chlorella*, the genus containing the symbiont for mixotrophic testate amoeba (Gomaa *et al.*, 2014). They could therefore be indicative of wider food web changes as well as direct responses of free-living algae. Higher resolution data would be required to determine if this was the case. Ideally, the use of protist-specific primers alongside the algal community would allow a correlation between Trebouxiophyceae, *Chlorella* and testate amoeba abundances to be tested for. Traditional microscopy techniques for studying testate amoeba would be extremely difficult for this site, as they would require more material to be sampled than would be possible, especially from the higher nutrient treatment plots where *Sphagnum* abundance has declined.

### 4.4.2 Non-plot effect

In the original bioinformatics pipeline, the difference between plot and non-plot controls suggests an impact either from general disturbance of experimental plots, or perhaps a pattern reflecting spatial distance at a medium scale. Non-plot controls were taken from different hummocks, but were, by necessity due to limitations of possible sampling areas, located more closely together than the experimental plots, which had to be far enough away from each other to prevent the nutrient addition spray hitting the wrong plot. Furthermore, the Whim experimental site borders an area of commercial peat extraction. The non-plot samples were taken from the point furthest from this site, on the boundary with a neighbouring SSSI, and so may have had less hydrological effect from the drainage channels built for peat extraction.

Whilst care is taken to minimise disturbance from trampling and vegetation damage in the Whim experiment, the experimental plots have a higher intensity of disturbance than the area where the non-plot controls were taken from. There could also be possible differences in the surrounding non-*Sphagnum* vegetation for the different sites. This was not surveyed due to restricted sampling time. Differing vegetation could impact the algal communities via allelopathic interactions, which are important in determining the *Sphagnum* microbiome (Hamard *et al.*, 2019). Differences in hydrology or nutrient cycling are also possible, as are shading effects from overhanging

vegetation. Even differences in hummock height could affect the hydrological conditions experienced by the photosynthetic microbial community. Whatever is the cause, these results would suggest that, in this site at least, nutrient addition has less impact on peatland algal community than other, unexplained heterogeneity.

However, the second bioinformatics pipeline using Kraken2 found no significant differences between plot and non-plot samples, though there were, predictably, significant differences in beta diversity (beta-dispersal). As the results from the Kraken pipeline are likely to be more reliable (see above), this suggests that there is not, in fact, a measurable disturbance effect on the algal community composition, and that the variance between communities is due to a different, unexplained variable, or simply due to innately high variance. The lower beta diversity of non-plot groups may not be ecologically significant, as there were fewer non-plot than plot samples.

The Mantel test did not show spatial autocorrelation, with no statistically significant correlation between geospatial distance and Bray-Curtis distance of the algal community. This suggests that the lower beta diversity of non-plot samples is unlikely to be because of their closer relative spatial distance, despite non-plot samples being situated notably closer together due to lack of undisturbed *Sphagnum capillifolium* hummocks on the site.

### 4.4.3 Plant community changes

The pattern of the majority of non-host reads in nutrient addition plots being assigned to vascular plants, and in control plots being assigned to non-vascular plants, reflects vegetation patterns seen in other studies from the site (Levy *et al*., 2019). Nutrient addition treatment plots had less variance in the plant community as detected by metabarcoding. Previous studies have shown that P deposition results in decreasing *Sphagnum* cover and an increase in vascular plants (Lu *et al*., 2023).

Perhaps the most striking change in the vegetation community was the significant difference in non-vascular plants with all nutrient addition levels. Juggermaniopsida had notably higher relative abundance in plots with background nutrient addition levels. Whilst long distance dispersal may be possible for liverworts (Sotiaux, Sotiaux and

Vanderpoorten, 2006), it is generally assumed that the majority of bryophyte spores are deposited close to the parent plant (Stoneburner, Lane and Anderson, 1992; Söderström and Jonsson, 1989). As the treatment plots were 13m wide it is therefore likely that such a large proportion of liverwort reads are a genuine signal from treatment effects, rather than caused by other dispersal factors. These results suggest that any threshold in nutrient effects on bryophytes for this system is between the background level (8-11 kg N/ha/yr) and the lowest treatment level (background + 8kg N $ha^{-1}$ $y^{-1}$). Addition of P and K does not seem to affect the non-vascular plant community. The increased N levels reduce the variance in the non-vascular plant community, and thus the β diversity between treatment plots.

Vascular plants showed no such significant differences in community, and had greater between-sample variation. Ferns (Polypodiopsida) have a soil spore bank, and fern spores are found on *Sphagnum* bogs surfaces even if they are not growing on the bog and spores do not successfully germinate on the *Sphagnum* surface (Dyer and Lindsay, 1992). Most fern spores are deposited close to the parent plant, but longer distance dispersal of spores is possible, albeit rare (Rose and Dassler, 2017). Other vascular plants present a more complex picture. Wind dispersed pollen, for example for the Scots pine or grasses, can travel for long distances (Frisk *et al.*, 2022; Robledo-Arnuncio, 2011). This raises questions about how local the source plants would be when there was no clear source material in the samples. If these issues can be taken into account, however, it provides an interesting example of bryophyte samples providing eDNA insights into the wider vascular plant community.

## 4.5 Conclusions

Kraken2 is preferable to DADA2 for this dataset; the lower proportion of unclassified reads makes it possible to judge if community shifts are occurring, whilst maintaining the robustness of the DADA2 pipeline filtering for error and chimera removal. While some of the species and genus assigned by Kraken2 are unlikely to be accurate, insights are still possible at higher taxonomic levels, even in the loss of specific species may not be detectable. In general, the low read numbers for microbial taxa mean there would

need to be caution in drawing any conclusions about the algal community at the species level from this dataset.

Whilst the algal communities as a whole have no significant differences between treatments, there are indications of smaller scale shifts in certain taxonomic groups. In particular, there were more Trebouxiales and fewer Zygnemophyceae in the highest nitrogen addition treatment. The method also detected a change in the bryophyte community, with clear clustering of all nutrient addition treatment plots, and greater relative abundance of Jungermanniopsida in control plots. This is suggestive of greater sensitivity to nutrient addition in the bryophyte communities than in the algal communities, when host bryophyte is controlled for.

# 5 The effect of tree cover on photosynthetic microbes in a near natural UK bog woodland system

## 5.1 Introduction

Photosynthetic microbes are abundant and biodiverse in UK peatlands; a gram of peatland soil may contain 120,000 microalgae across the different photosynthetic taxa (Gilbert and Mitchell, 2006). However, they are less studied than other key microbial groups. They can play functionally significant roles in peatland ecosystems through photosynthesis and nutrient cycling (Jassey *et al.*, 2022; Yin *et al.*, 2022; Hamard *et al.*, 2021). If they show consistent responses to environmental variables, they may also have potential as bioindicators for peatland condition. Photosynthetic microbial communities are spatially variable at small scales and likely to differ with changes in tree cover and associated hydrology. This variance may be compounded by heterogeneity in the impact of different individual trees; higher leaf area in trees is associated with greater hydrological demand (Stovall, Shugart and Yang, 2019), and would also be expected to lead to a greater shading effect. However, in the UK this has not been investigated in stable or near-natural bog woodland systems.

Testate amoeba are protists commonly used for peatland environmental indication (Payne, 2013) and their responses to forested peatland have previously been investigated (Creevy *et al.,* 2023; Creevy *et al.,* 2018; Payne *et al.,* 2016). Mixotrophic testate amoebae (testate amoeba which obtain some or all of their carbon from photosynthesis via symbiotic algae) are less abundant in forested bogs (Payne *et al.,* 2016), and this effect can continue many years after tree felling (Creevy *et al.*, 2023). Photosynthetic algae are highly unlikely to be completely absent with tree cover, but this suggests that they may show similarly clear patterns based on tree presence.

The semi-natural and natural bog woodland found at Abernethy is extremely unusual in the UK, and previous work on microbes in UK bog woodlands has largely focused on forestry or forest to bog restoration. However, it is much more representative of peatlands in the wider European and North American context. In Canada, there is some

indication that forested bogs are more common in continental peatlands, with oceanic peatlands being more open, likely due to climatic effects (Thompson, Simpson and Beaudoin, 2016). Overall, forested and treed peatlands may make up around 50% of all Canadian peatlands (Beaulne *et al.*, 2021; Thompson *et al.*, 2016). Palaeobotanical evidence suggests that treed peatlands were historically found in areas of the UK where they are no longer present, even if only as a transitional state when changing climate resulted in shifts between wet and dry which favoured peat growth and tree growth respectively (Bishop *et al.,* 2018). Therefore, it is not unreasonable to assume that, with climate change, treed peatlands will again become more common in the UK.

By understanding the functioning of microbial communities in bog woodland at Abernethy, we hope to better understand the mechanisms by which tree cover drives changes in photosynthetic microbial community composition in UK bogs and bog woodlands. If findings can be generalised to other sites this would allow better understanding of the microbial processes in this unusual but important habitat type in the UK. This may allow comparisons to be made between algal communities and ecology in near-natural systems and in bog forestry or forest to bog restoration, which could help to inform management and restoration practices. Understanding the dynamics of forested peatlands are also important for climate modelling - forested peatlands have similar recent carbon accumulation to non-forested peatlands, but long-term carbon stores are lower, possibly due to greater carbon loss through higher rates of decomposition in the presence of trees than in non-forested peatlands (Beaulne *et al.*, 2021). In forested peatlands, carbon storage is higher in the peat itself than in the trees, and tree carbon storage decreases with increasing peat depth (Beaulne *et al.*, 2021). Therefore, considering forests on peat as ‘forests’ in climate modelling is unlikely to be sufficient, and better understanding of them as a form of peatland is crucial for accurate climate modelling.

This study aims to investigate photosynthetic microbial changes across a forest to bog ecotone and test these key hypotheses:

- Tree presence/absence impacts algal communities.
- Tree health and height effects the impact, with healthier trees (i.e. those with greater leaf area, see Section 5.2.2, Fig. 5.3) causing more disturbance to the communities due to larger shading and hydrological effects.
- Shading from trees, measured photographically as canopy cover, has a significant impact on photosynthetic microbial communities.
- Other environmental drivers of tree impact on microbes, such as changes to hydrology and nutrient profiles, can be identified from field data.

## 5.2 Methods

### 5.2.1 Site description

Fieldwork was carried out in Abernethy Forest and RSPB Loch Garten NNR (57.237306, -3.689021), close to Aviemore in the Cairngorms National Park (Fig. 5.1).




**Figure 5.1.** The location of Abernethy Forest (shown at the white point) in Scotland, between the Cairngorm plateau and Inverness, and in the UK as a whole. Satellite imagery from ArcGIS World Imagery (Sources: Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community).

The site is characterised as a post fluvioglacial landscape with forested areas on better drained higher ground and peatland formation where drainage is poorer (Mchaffie *et al*., 2009). The specific sampling area was chosen across an ecotone of forest to bog, with an area of open bog with mostly solitary trees, surrounded by thicker forest (Fig. 5.2). All trees growing within the sampling area were *Pinus sylvestris* L. Sampling was carried out on the 19th and 21st October 2021, towards the end of the growing season.

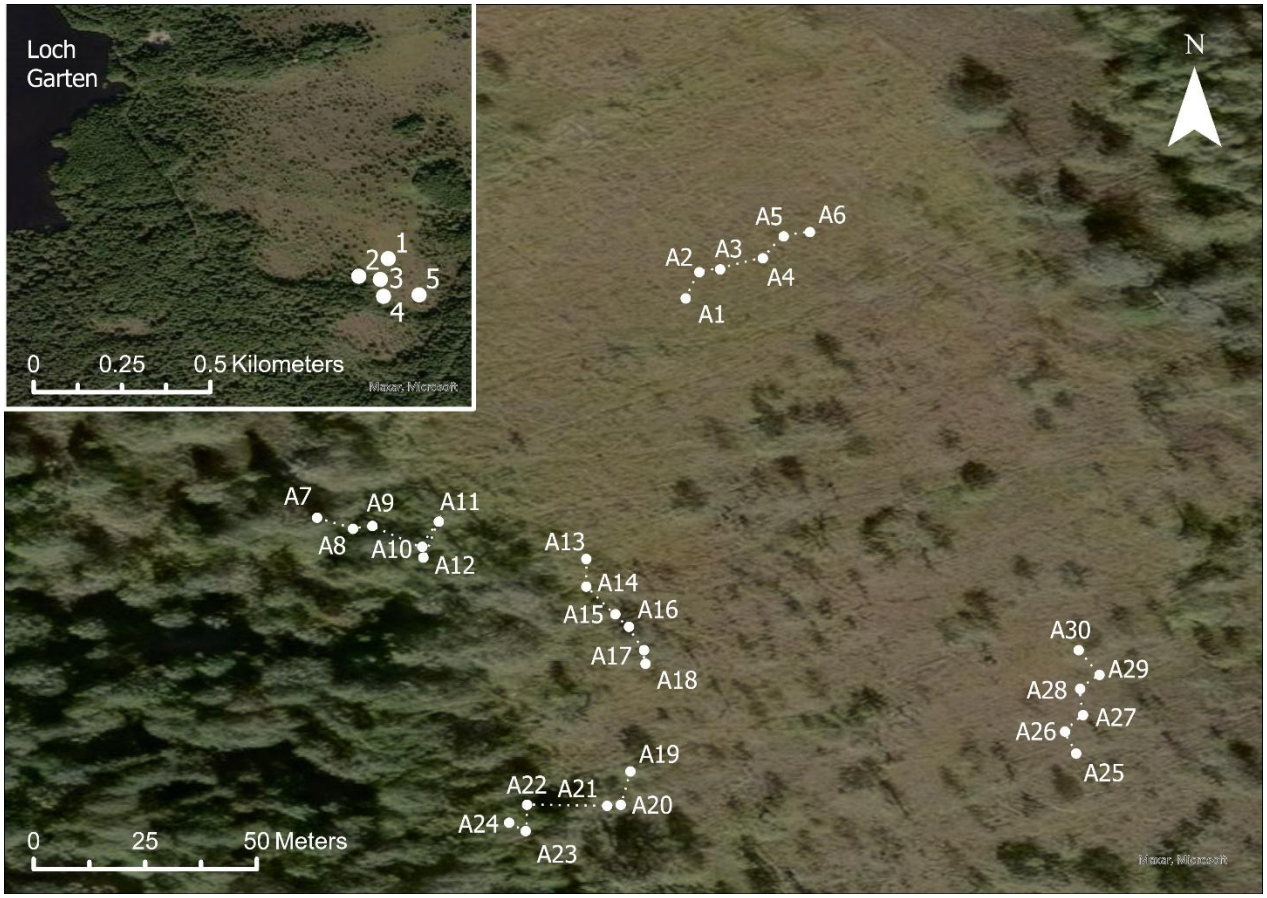



**Figure 5.2.** The location of sampling points within Abernethy Forest. These were chosen to cover the heterogeneity of the forest bog ecotone, with open bog, denser tree cover and the transition zone between. Sampling points were clustered into zigzag transects (shown in the inset image shows transects, located by the first point on each transect). Transects 1 and 5 were in the more open bog area, while transects 2-4 were in the forest bog transition zone. Satellite imagery from ArcGIS World Imagery (Sources: Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community).

### 5.2.2 Field methods

Sampling sites were chosen to encompass the heterogeneity of the forest to open bog ecotone. Six sampling points were located by walking a zigzag with 4m measured between each point (Fig. 5.2). At each sampling point, a 50cm$^2$ quadrat was dropped on an area defined visually as bryophyte lawn, and percentage cover of each vegetation species was recorded. A small sample with surface area of around 10cm$^2$ was then taken of the bryophyte present with the highest percentage cover within the quadrat

(*Sphagnum* spp. in all cases) for laboratory analysis. Small samples of other bryophytes were also taken for more detailed identification in the laboratory if confident field identification was not possible. Moss identification was carried out using Godfrey and Rogers (2021) and Atherton, Bosanquet and Lawley (eds, 2010), and lichen identification used resources from the British Lichen society, with reference to Coppins and Coppins (2008). Soil moisture, temperature, pH and oxidation-reduction potential (ORP) were measured with probes (Oyster 341350A-P, Extech Instruments, New Hampshire, USA and WET150 meter, Delta-T Devices, Cambridge, UK) at a depth of 15cm from the moss surface. Probes were calibrated and operated according to manufacturer's instructions.

Each tree stem within 2m of the sampling point was measured. Tree height was measured with a tape for trees below 215cm, and was calculated using a clinometer for trees above 215cm. Circumference at breast height (~1.5m) was measured with a tape, from which diameter at breast height (DBH) was calculated.

Trees shorter than breast height had no DBH measurement. DBH and clinometer angle were measured by the same researcher for consistency. Trees were also graded in terms of health, with 1 being an apparently dead tree, 2 being a tree with limited needles, discolouration or other signs of major stress, and 3 being a tree in apparently good condition (Fig. 5.3).

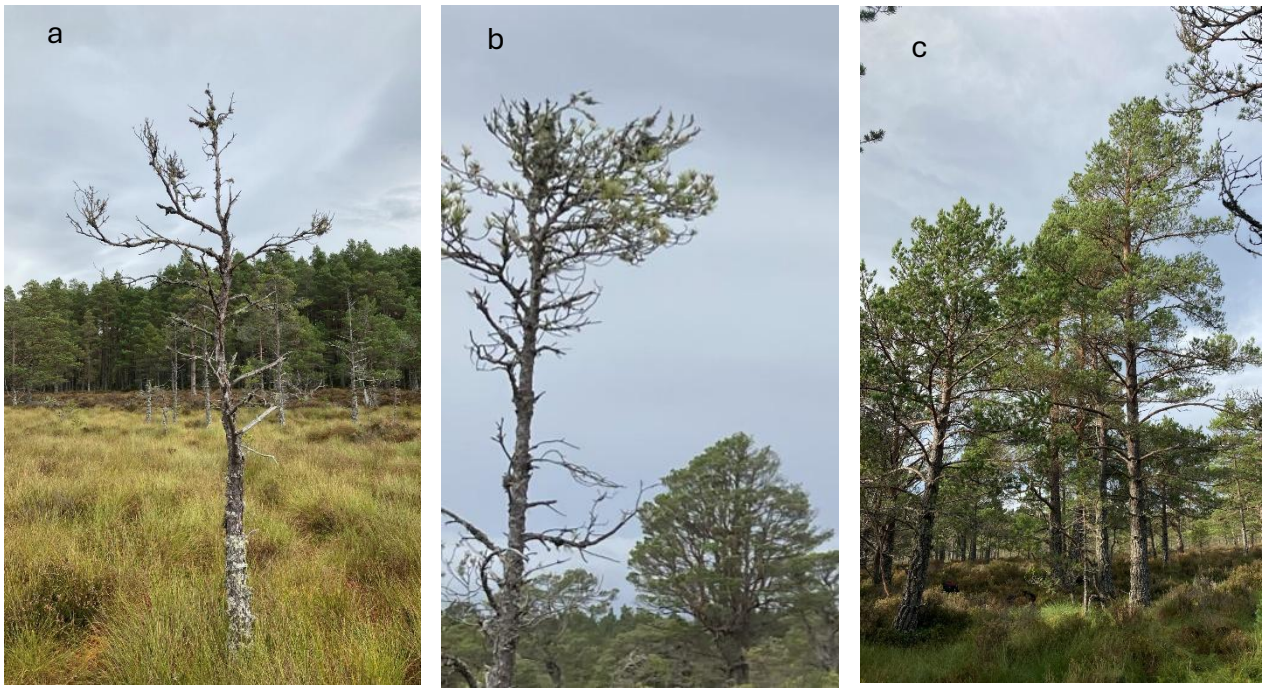



**Figure 5.3.** An illustration of the tree health grading used in Chapter 5. A) A tree graded 1, completely dead. What appears to be needles are actually lichen. B) A tree graded 2, still alive but with sparse needles. C) Trees graded 3, in apparently good condition with no obvious signs of major stress.

To calculate canopy cover, two "fish-eye" images were taken at each sampling point using a hemispherical lens (Apexel HB170 aspherical lens, Apexel, Shenzhen, China) and a smartphone camera (iPhone 12 back dual wide camera, Apple Inc., USA). One image was taken in a North-South orientation and the other in an east-west orientation to maximise field of view in combined images, following the protocol in Bianchi *et al*. (2017). GPS coordinates were extracted from the N_S oriented photo for each plot using the GeoTagged Photos to Points tool in ArcGIS Pro 3.4.2, except for site 1, where this image had been lost. Coordinates for this site were instead extracted from a reference image of the quadrat.

Three TMS-4 data loggers (TOMST, Prague, Czech Republic) were placed across one of the transects to provide longer term climatic data across the variation in tree cover. One was placed at each end of the transect, and the third centrally, to capture climatic conditions over a range of tree-cover states. Loggers were installed into the top 10cm of

the bryophyte layer, allowing measurement of surface peat (0-6cm) moisture and temperature, temperature of the *Sphagnum* capitula, and free-air temperature.

### 5.2.3 Sample storage

In the field, *Sphagnum* samples were cut so that each stem was 15cm long, to encompass the photic zone. As some *Sphagnum* colonies at this site had lower density than at other field sites, the decision was made in the field that the photic zone may extend further below the capitulum than in the samples for previous chapters. Samples were stored in the dark in an insulated bag at field temperatures for up to 2 days before storing at 5°C. In the laboratory, the *Sphagnum* samples were subdivided. One subsample was stored at -80°C for molecular analysis. The remaining sample remained at 5°C until nutrient analysis.

### 5.2.4 Laboratory analysis

#### *5.2.4.1 DNA extraction and sequencing*

To prevent PCR and sequencing being swamped by *Sphagnum* DNA (and DNA derived from other vascular and non-vascular plant litter), a pre-filtering step was carried out, as detailed in Chapter 4. Briefly, d2$H_2O$ was added to *Sphagnum*, shaken by hand 30 times and then filtered through a 195µm nylon microsieve. This was chosen to allow microbes to pass through into the filtrate, but retain *Sphagnum* fragments and other litter detritus, especially pine needles. The filtrate was then filtered through a pre-sterilised 0.2µm mixed cellulose ester (MCE) filter paper under vacuum (Fig. 5.4). DNA was extracted from this filter paper using Quiagen Powersoil DNA extraction kits following manufacturer's instructions. PCRs using rbcLa primers, purification and sequencing were carried out as detailed in Chapter 4.

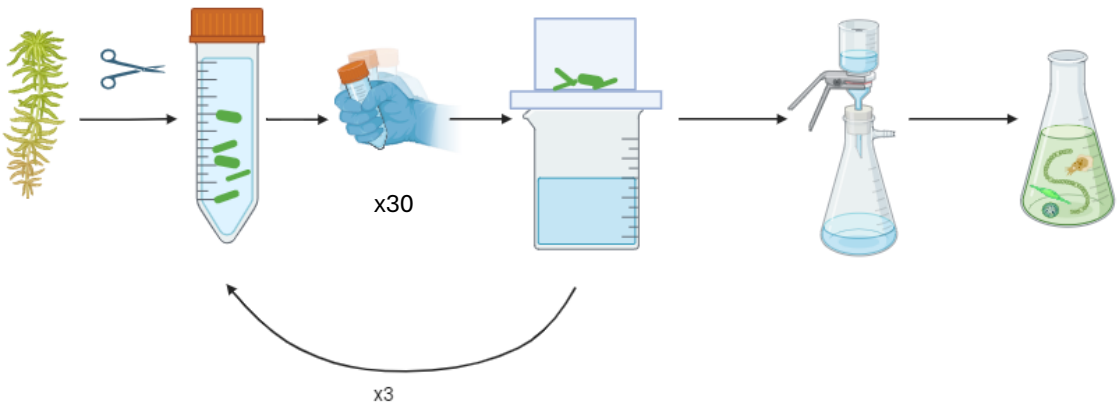



**Figure 5.4.** Diagram showing the microbial sample filtering process for Chapter 5.The Sphagnum capitulum was first cut into small pieces with sterilised scissors, shaken 30 times in d2$H_2O$ and filtered through a microsieve to remove plant material. The shaking and filtering was repeated twice (to a total of three passes through the microsieve) and the combined filtrate was then vacuum filtered to extract the microbes.

### *5.2.4.2 Environmental measurements*

To calculate bryophyte percentage moisture, 12-18g wet weight *Sphagnum* was dried at 40°C until at stable weight, then dry weight was subtracted from wet weight to give water content in g.

To measure nutrient levels, the pore water was extracted from 10g (wet weight) of moss by centrifuging over a 1.2µm GF filter (see Fig. 4.4). As there was too much material to complete this in one step, the moss sample was divided into two and centrifuged one after the other, leaving the liquid in the bottom of the centrifuge tube between runs. The sample was centrifuged at 4000RPM for 5 minutes for each half of the moss sample,

then the filter paper was centrifuged without moss for a further 2 minutes to remove residual water from the filter. These extracts were then vacuum filtered through 0.45µm Nylon filters (except for samples 28-30, which were filtered through 0.45µm syringe filters due to a shortage of membranes). Extracts were frozen prior to analysis for ammonium N, nitrate/ total organic N and phosphate P using a Seal Analytical auto-analyser 3 (AA3) (Seal Analytical, Wrexham, UK). This was operated according to the manufacturer's instructions, and calibrated as standard during the run with reference samples of known concentration. As these samples were larger than those collected from the experiment in Chapter 4, there was enough pore water collected to run them on the AA3 without dilution being required.

## 5.2.5 Data analysis

### *5.2.5.1 Canopy cover*

#### 5.2.5.1.1 Enhance HP and rCaiman

Images were converted from HEIC to JPEG format using an iPhone SE (the conversion was carried out automatically by downloading from a shared drive, saving as an image and then reuploading). Canopy cover calculations were carried out based on the protocol described in Bianchi *et al*. (2017). The N-S and E-W images for each sampling point were merged using the software Hugin (2009). Merged image dimensions were calculated using 'specify selection' in ImageJ2 via Fiji (Schindelin *et al*., 2012). Thresholding of images was carried out using the EnhanceHP algorithm via the package rCaiman (Díaz, 2023) in R. Both Bianchi *et al*. (2017) and Díaz (2023) showed that the EnhanceHP was more effective than the other main method, known as ISOdata. However, correct lens calibration is vital for this to be successful (Díaz, 2023).

Lens calibration was attempted following the simplified geometric method in Díaz *et al*. (2024), with further guidance in Díaz (2023). Briefly, a semicircular template (Diaz, 2017) with radius of 45cm was printed in A1 size and trimmed and lined up on a table according to the instructions in Diaz (2023). Pushpins were placed at the pre-set points marked on the template along the perimeter of the circle, with known angles between them. The camera was lined up with the pushpins and the lines marked on the template

following the instructions in Diaz (2023) (Fig. 5.5), and an image was taken using the smartphone and lens which had been used in the field (see Section 5.2.2). The pushpin points in the resulting image were digitised using ImageJ and lens calibration was calculated using the calibrate_lens() function in rCaiman (Díaz, 2023), which compares the known angles between the push-pin points to the angles calculated from the image.

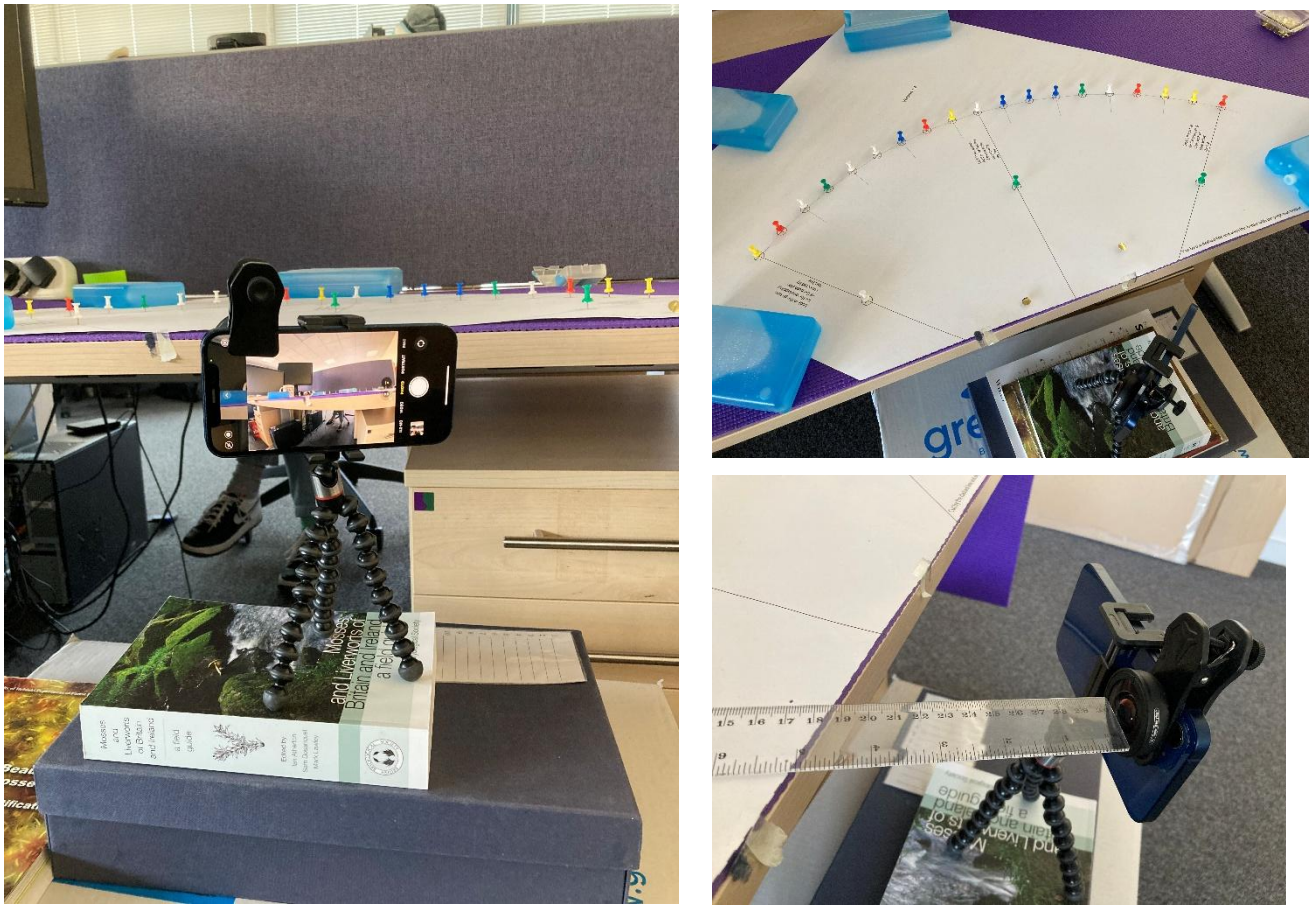


**Figure 5.5.** Images showing the setup for hemispherical lens calibration as described in Diaz (2023) and Diaz *et al.* (2024), including phone camera with clip on lens to be calibrated, the arrangement of pushpins on the template, and using a ruler to align the lens with the table surface.

Testing the lens coefficient (function: test_lens_coef) failed at the 'Test that lens projection function works between the 0-to-1 range' stage. The rCaiman handbook (Diaz, 2023) states that re-collecting data may be necessary if failure occurs at this point, but variations of point inputting did not stop this failing at this stage. Attempting

to progress with the pipeline with the lens calibration also resulted in failure of subsequent functions.

Repeated failure of hemispherical lens calibration may have been because the lens was a wide-angle lens, but not a true hemisphere. There was also less visual distortion of the images than would be expected from a hemispheric lens. Therefore, the decision was made to calculate the canopy cover manually using ImageJ via Fiji (Schindelin *et al.*, 2012).

5.2.5.1.2 Canopy cover estimates with ImageJ

As the merged images were different shapes and had large black boundaries (see Fig. 5.6) this would have interfered with thresholding. Therefore, the decision was made to calculate the canopy cover for the separate rectangular E-W and N-S oriented images for each site, and then calculate the average.



**Figure 5.6.** Merged canopy cover images created using Hugin. The thick black margins and variable shapes would make trimming these for use in ImageJ challenging.

Each image was converted into 8-bit before thresholding to separate the trees from the background sky. Thresholding with the default setting worked well for images with high canopy cover (Fig. 5.7), but struggled with images with lower cover and more sky (Fig. 5.8). Thresholding in Minimum or Intermodes settings had better differentiation; intermodes was selected as it picked out more of the tree branches (Fig. 5.9), and the vignetting could be more easily controlled for when thresholding.

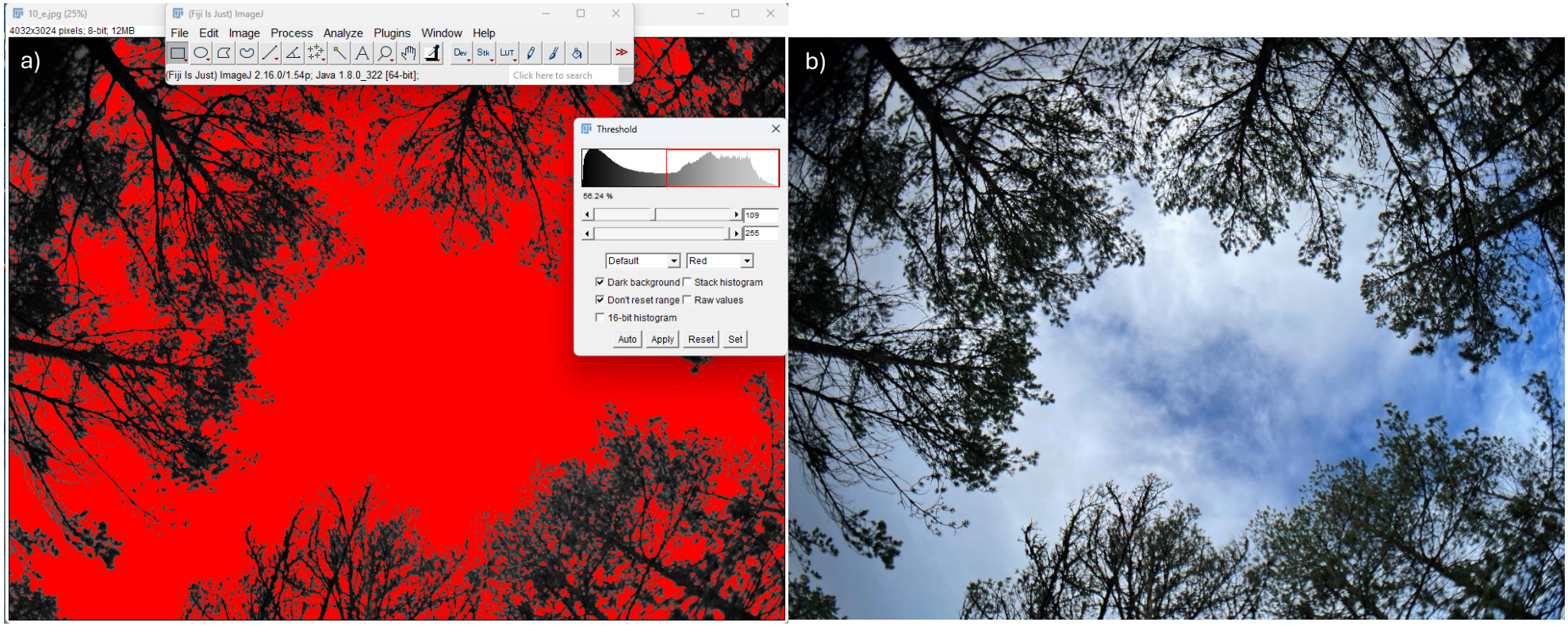



**Figure 5.7.** a) An image with high canopy cover with default thresholding in Fiji. Trees versus background are fairly well distinguished. B) The original image.

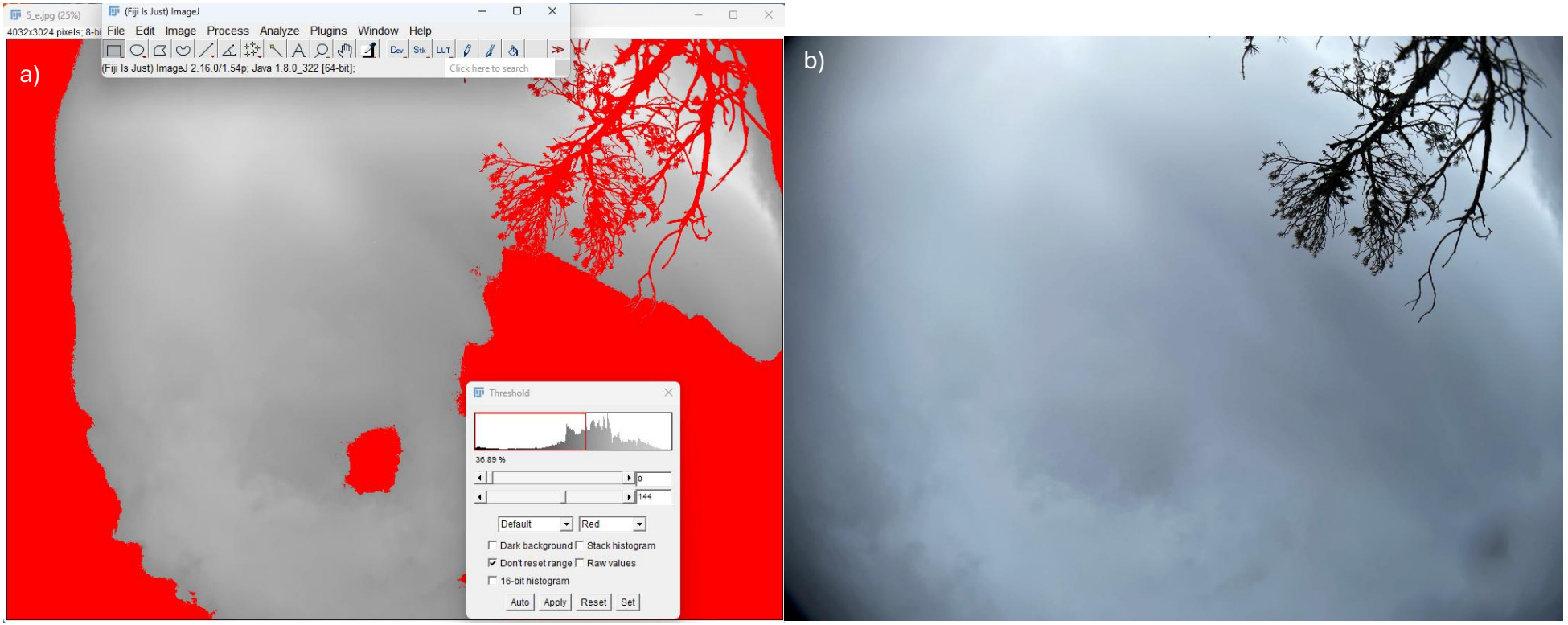



**Figure 5.8.** a) Default thresholding of an image with lower canopy cover in Fiji. Large areas of the sky are misclassified along with the tree. b) The original image.

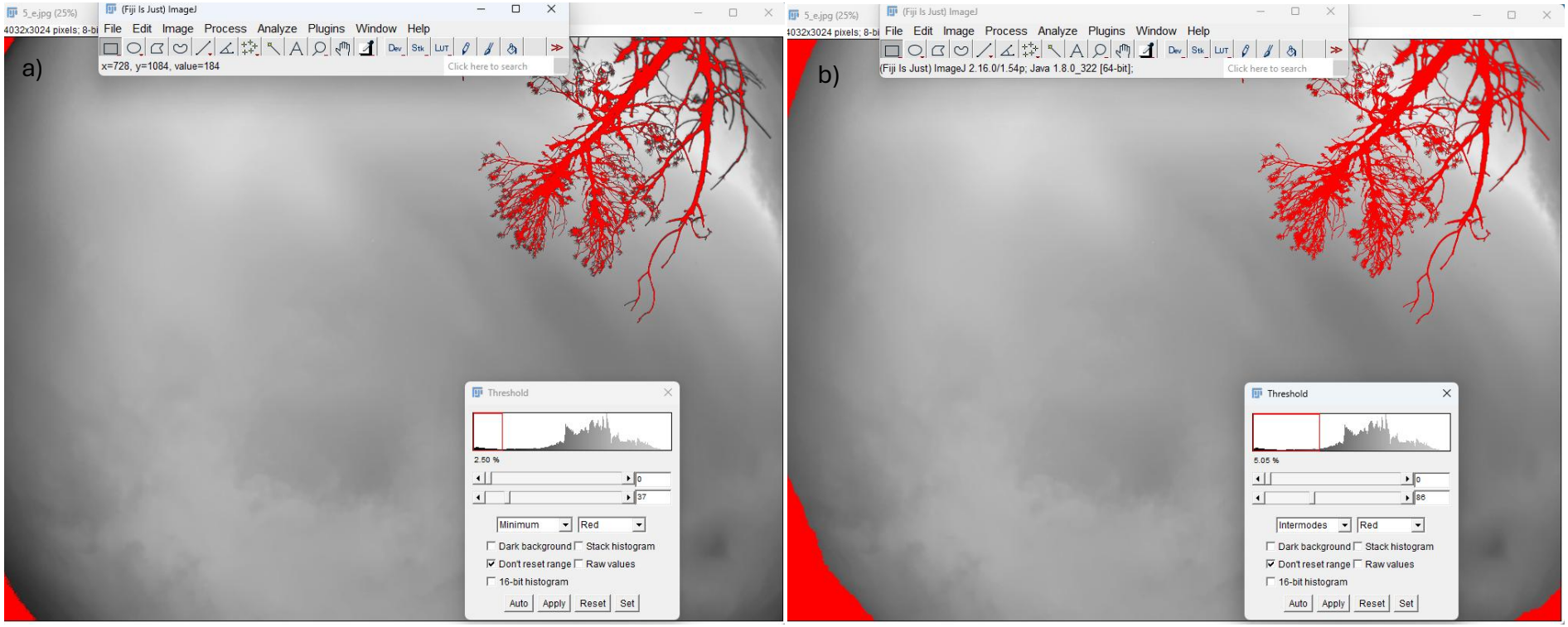



**Figure 5.9.** a) An image with low canopy cover thresholded in 'minimum' mode in Fiji. b) The same image thresholded in 'intermodes' mode. The same image thresholded in default mode and the original image can be seen in Fig. 5.8.

Intermodes performed well at the automatic threshold for most images. The automatic level, calculated for each individual image, was used rather than a pre-determined thresholding level because varying sky conditions during the fieldwork led to different brightnesses of the images. Attempts to standardise thresholding value for tree images led to over or under classification. Where the automatic thresholding level led to misclassification of background sky as trees (especially an issue for images taken with less cloud cover), the thresholding level was manually adjusted until no central sky pixels were classified as trees, vignetting was minimised, and tree classification was still equivalent to the original images.  9 out of 25 images with tree cover had to have the thresholding level manually decided in this way. The automatic (and manual when this was used) thresholding values used for each image were recorded (Supplementary Table 5.1). Following thresholding, percentage cover was calculated using the 'Analyze particles' function in Fiji.

Not all images displayed vignetting, but all images taken with no visible trees (open sites) did. These were therefore used to calculate an average vignetting percentage. As 5 of the first six images taken were all open, these had very similar sky conditions and brightnesses, and so a fixed thresholding value could be used. Open images were thresholded at level 40, as this gave a better estimate of the vignetting effect than a lower threshold (Fig. 5.10). The average vignetting value of 1.558 was calculated as the mean of the percentage covers for all open (treeless) images.

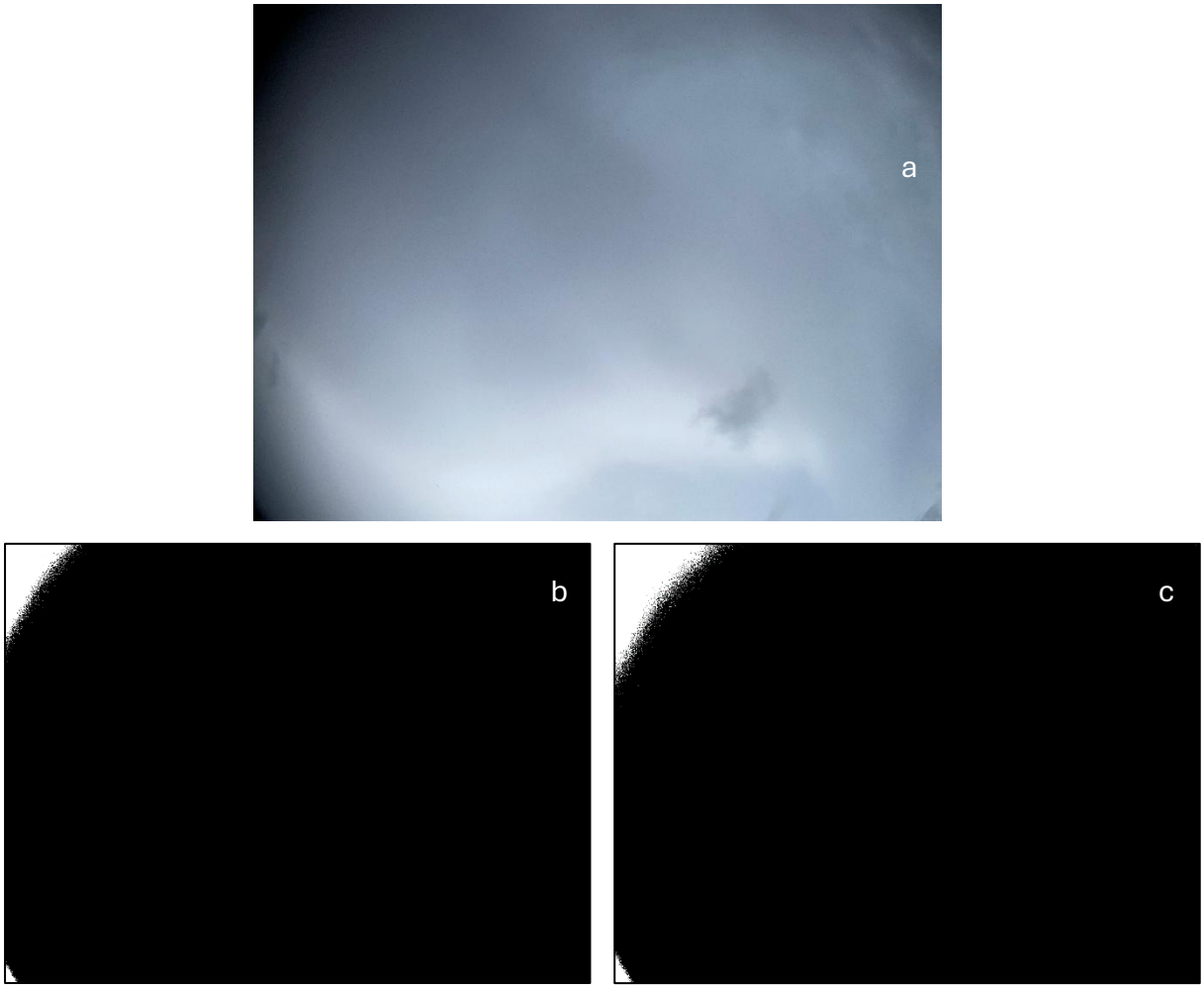



**Figure 5.10.** An image with canopy cover of zero (image 2e) in a) the original colour, b) the same image thresholded at level 20, and c) thresholded at level 40. The more cautious thresholding at 40 looks closer to the vignetting effect seen on the original image.

Images with vignetting (or physical obstructions picked up in the image, such as other field equipment, which blocked light and prevented seeing canopy cover) were classed as showing vignetting. For each of these images, the mean vignetting percentage cover was subtracted from the calculated percentage cover. Whilst not all vignetting effects were the same size, this was judged to be a cautious approach to prevent overestimation of canopy cover, especially in more open images. Following vignette removal, an average of the canopy cover for the E-W and N-S oriented images for each site was calculated; this was the final calculated canopy cover percentage. Images with no trees visible were classed as zero.

### *5.2.5.2 Bioinformatics and data processing*

Bioinformatics were carried out as described in Chapter 4. They were initially carried out using a DADA2 pipeline, with taxonomic assigned carried out using DADA2 with a custom rbcL database. However, this resulted in very low levels of taxonomic assignment. Taxonomic assignment was therefore carried out using Kraken2 on data following cleaning, filtering and sample inference in DADA2, which resulted in a much lower percentage of reads being unclassified (see Section 3.5.2). For a full explanation of the methods development, see Chapter 3. Because of concerns over the accuracy of the taxonomic assignment at the species level, a conservative approach was taken, and only higher taxonomic levels (order and above) were taken forward for subsequent analysis from this second bioinformatics method.

As in Chapter 4, data were filtered to remove non-target taxa likely to be erroneous. The taxa filtered out were Annelida, Arthropoda and Chordata, which included all animal taxa identified. Total read counts per sample were calculated before and after removal of these taxa. The removal did not affect the read counts of the six samples with the lowest number of reads (see Table 5.1). No samples needed to be filtered out because of a read count lower than 10,000 (Fig. 5.11).

**Table 5.1,** Total read counts for the six samples in Chapter 5 with the lowest number of reads.

| Sample | Total read count before removal of animal reads | Total read count after removal of animal reads |
|---|---:|---:|
| A28 | 204569 | 204569 |
| A29 | 351098 | 348221 |
| A6 | 371210 | 371210 |
| A18 | 384064 | 380139 |
| A10 | 466009 | 464898 |
| A27 | 489052 | 489052 |

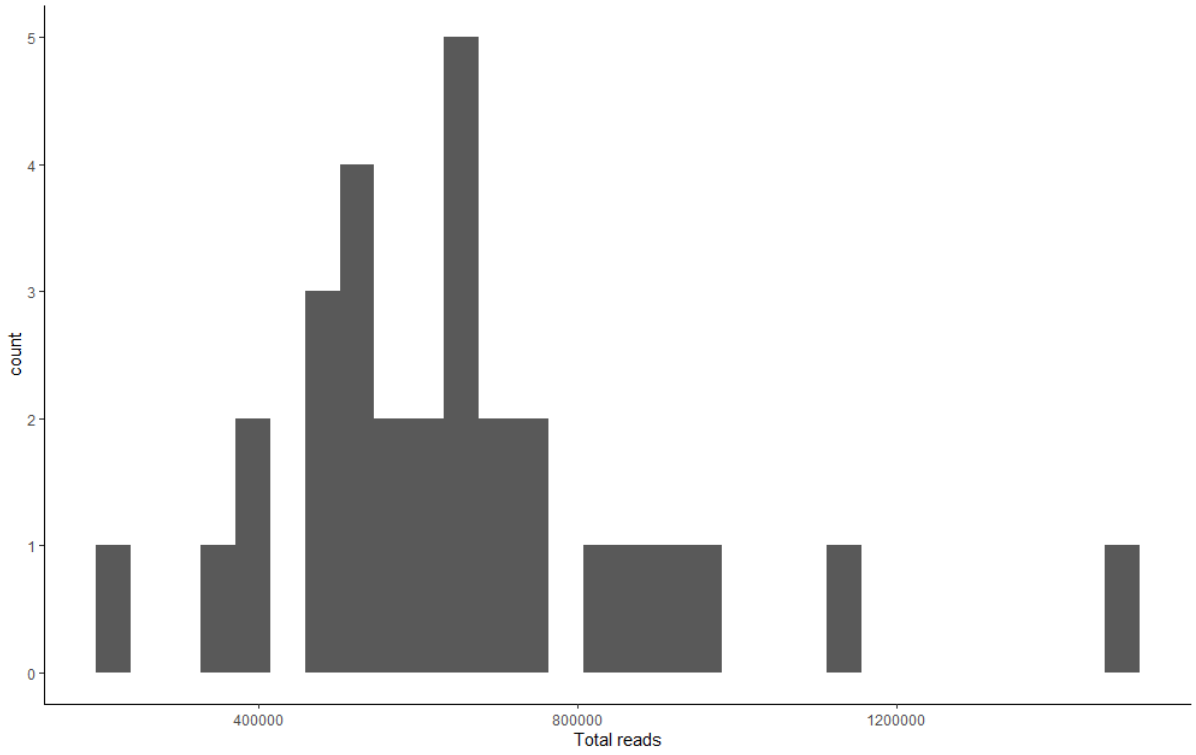



**Figure 5.11.** Read count histogram after removal of animal reads for all samples from Chapter 5.

Rarefaction curves were calculated using the rarecurve function in vegan (Oksanen *et al*., 2022) in R. A cumulative sum scaling (CSS) transformation was carried out on the abundance table using the metagenomeSeq package (Paulson *et al*., 2013). This CSS transformed data was then plotted to see if there was any correlation between observed features and sequencing depth. As the rarefaction curves appeared to all be levelling out or to have met their asymptote (Fig. 5.12) and there was no correlation between observed richness and read count (a proxy for sequencing depth) (Fig. 5.13), the decision was taken that data did not need to be normalised by sequencing depth. This is likely due to pooling the data to order level, reducing the number of ASVs involved.

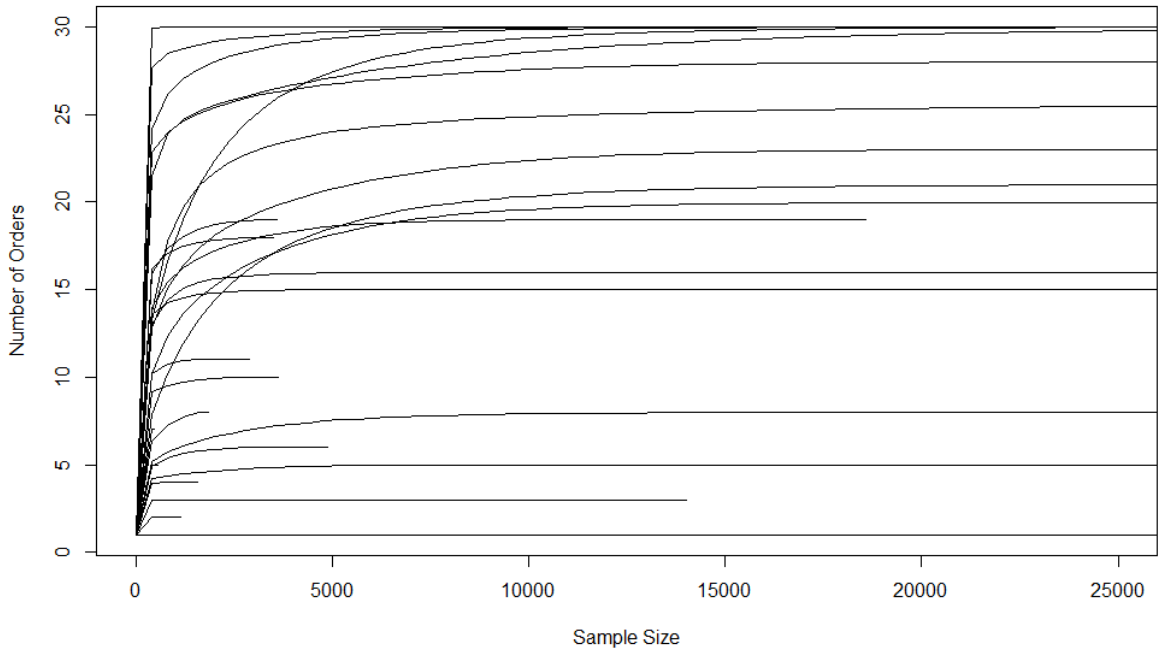



**Figure 5.12.** Rarefaction curves per sample after removal of animal taxa for all samples from Chapter 5. All rarefaction curves appear to be levelling out or to have met their asymptote. Note that as the data has been pooled at order level, the 'ASVs' are orders rather than species/ASVs proper.

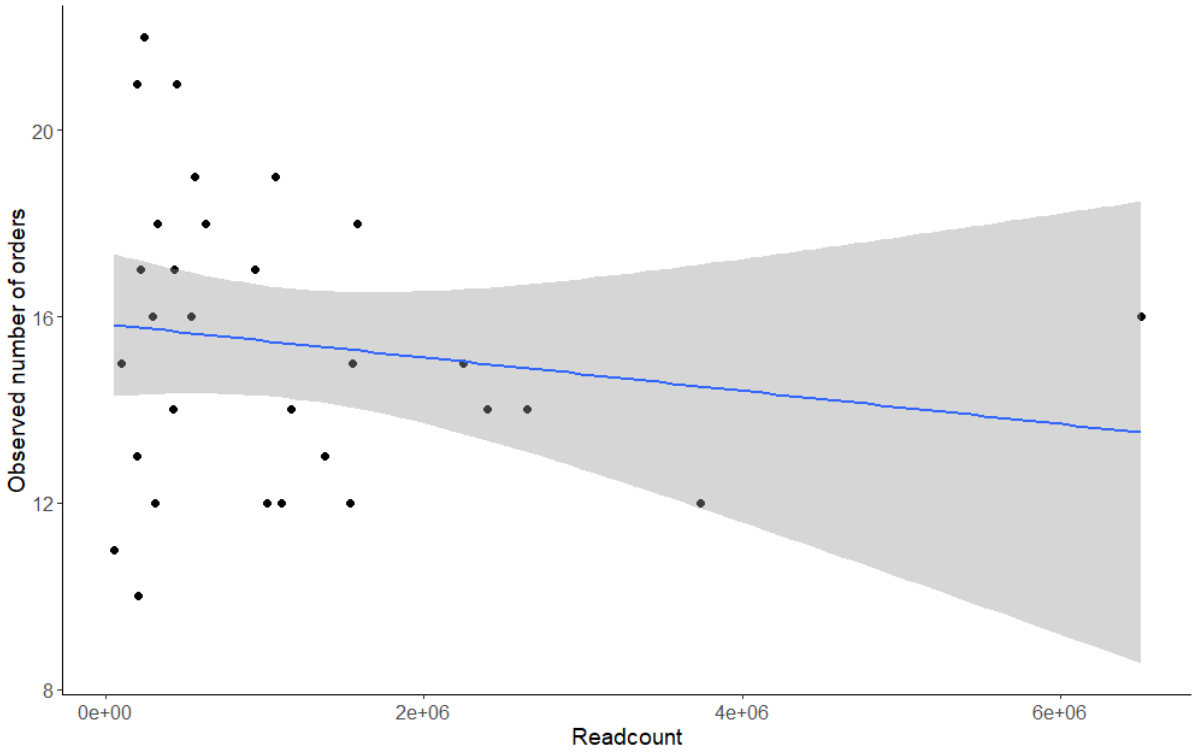



**Figure 5.13.** Correlation plot between observed taxonomic richness and read count (a proxy for sequencing depth) for Chapter 5 data (CSS transformed), showing no correlation between observed richness and read count.

### *5.2.5.3 Differences between transects*

As the data from the second bioinformatics method has been pooled at the order level, standard biodiversity indices such as species richness and evenness could not be calculated for this data. However, as the initial bioinformatics was processed in the same way prior to taxonomic assignment for both methods, the first estimates of diversity indices can still be used if the taxonomic uncertainty is taken into account.

As lines of samples were arranged as zigzags across the forest-bog ecotone to encompass a large range of tree cover conditions, rather than being true transects, issues with differences between them are less concerning than they would be for a traditional block-designed experiment. However, to better understand variation in the data, a Kruskal-Wallis test for significant differences in biodiversity indices in different transects was carried out. Based on this sampling layout, a Kruskal-Wallis test was also

used to test for significant differences in the number of tree stems within 2m of sampling between transects.

Differences between transects for the full algal community dataset were also investigated. Multivariate homogeneity of variance was tested with the permutest.betadisper function in vegan with 999 permutations. A Bray-Curtis dissimilarity matrix was constructed using the distance function in Phyloseq, then PERMANOVA carried out between transects using the adonis2 function with 9999 permutations in vegan (Oksanen *et al.*, 2022). Pairwise PERMANOVA was performed using the pairwise.adonis function from the pairwiseAdonis package (Martinez Arbizu, 2020). This was repeated for the data from the second bioinformatics method.

As there was variation in the relative abundance of Prokarya to Eukarya (Fig. 5.14), it was investigated if the ratio of prokaryotes to eukaryotes differed between transects. The percentage of all microbial reads (plant taxa excluded) which were prokaryotes of any taxa was calculated. As the data was not normally distributed (Shapiro-Wilk normality test $W = 0.69$, $p \leq 0.001$), a Bartlett test of homogeneity of variances and Kruskal-Wallis test was used.

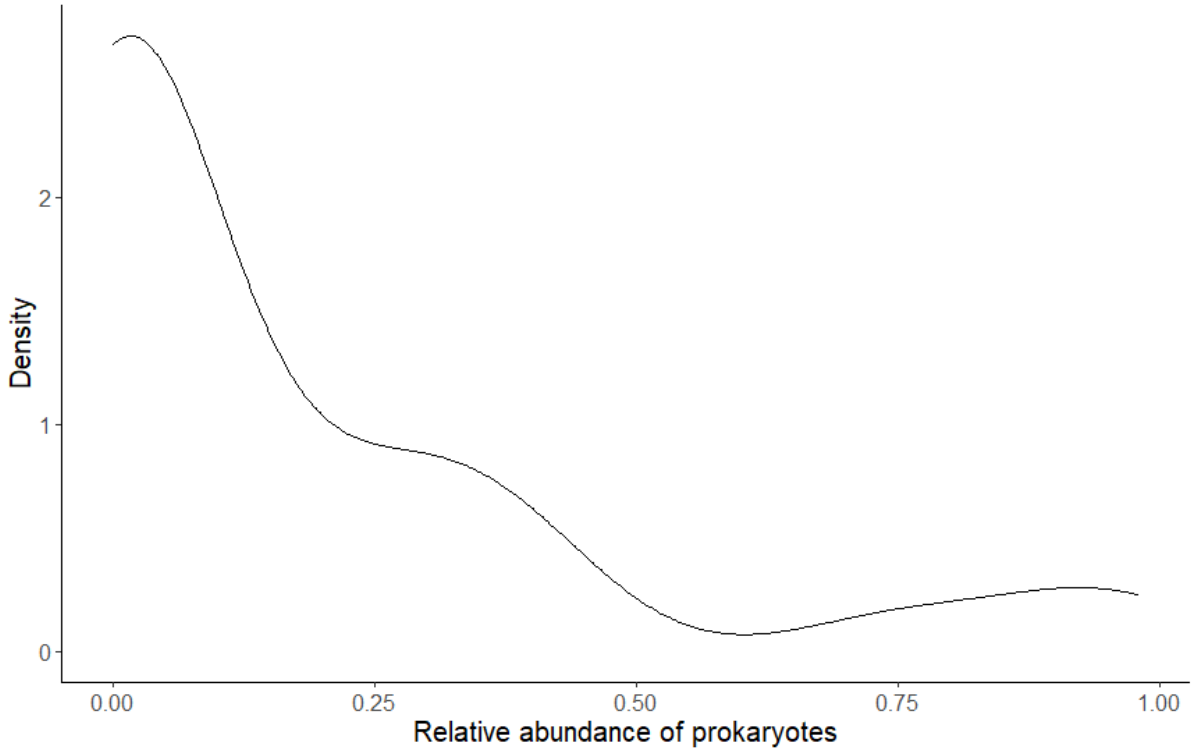



**Figure 5.14.** Density plot of relative abundance (proportional) of prokaryote reads for Chapter 5. While the majority of samples have low relative abundance of prokaryote reads, a few do have very high relative abundance.

### *5.2.5.4 Distance based redundancy analysis*

Distance based redundancy analysis (dbRDA) was selected for analysis, allowing the community matrices to be used as the response for a multiple regression analysis. Unlike other similar methods, dbRDA can handle non-Euclidean distances common to ecological presence matrices, which have high numbers of 0s where a taxon is absent. All numerical explanatory variables were scaled to convert to Z scores to handle heterogeneity of measurement scales. Variables were examined for collinearity by constructing a correlation matrix using cor.mtest and corrplot from the corrplot function in R (Wei and Simko, 2021) (See Fig. 5.44 in results).

However, as RDA explanatory variables should be chosen biologically, and whether relationships are due to conditions such as moisture or a different, inherent property of

tree proximity is a key research question, the decision was made to retain all variables in the first instance.

For analysing data outputted from the first bioinformatics method, distance-based RDA analysis was carried out on the Bray-Curtis dissimilarity matrix of algal communities using the capscale function in Vegan. The model was as follows:

*Model 1: dist ~ tree_stem_count + pH + ORP + soil_temperature + moisture/$m^3$/$m^3$ + moss_percent_water*

*Where dist is the Bray-Curtis dissimilarity of the algal community, tree_stem_count is the number of tree stems within 2m of the sampling point, ORP is redox potential, moisture/$m^3$/$m^3$ is the soil percent moisture as measured by the soil moisture probe, and moss_percent_water is the moss percentage water by weight*.

An Anova with 9999 permutations was used as a post-hoc test for statistical significance of the model and explanatory variables for all dbRDAs.

This was repeated for the second bioinformatics method with the addition of tree height and health and canopy cover into the RDA analysis. These had not been included in the analysis of data from the initial bioinformatics method due to time constraints. To help with variable selection and assess collinearity of the increased number of variables, variance inflation factors (VIFs) were calculated following Borcard, Gillet, and Legendre (2018) for dbRDA variables. During model selection, the model specific VIFs were also calculated following each dbRDA using the function vif.cca in vegan. Tree health measurements were not included as the function could not handle NAs for plots with no trees within 2m.

Including both maximum and mean tree height and DBH unsurprisingly resulted in very high VIFs for tree variables, as they would be expected to be highly colinear. Maximum tree height and mean tree DBH were chosen as the combination of height and DBH variables which minimised tree height and DBH VIFs when included in the model, and also resulted in the lowest VIFs for variables with VIF>10. The model was as follows:

*Model 2: dist ~ tree_stem_count + pH + ORP + soil_temperature + moisture/$m^3/m^3$ + moss_percent_water + mean_tree_dbh + max_tree_height + mean_canopy_cover + dominant_sphag_species + latitude + longitude*

*Where dist is the Bray-Curtis dissimilarity of the algal community, tree_stem_count is the number of tree stems within 2m of the sampling point, ORP is the redox potential, moisture/$m^3/m^3$ is the soil percent moisture as measured by the soil moisture probe, moss_percent_water is the moss percentage water by weight, mean_tree_dbh is the mean DBH for all trees within 2m of the sampling point, max_tree_height is the maximum tree height within 2m of the sampling point, mean_canopy_cover is the calculated mean canopy cover from hemispherical photography, and dominant_sphag_species is the dominant species of Sphagnum recorded in the vegetation quadrat at the sampling point.*

Although there were no significant correlations between the dissimilarities between vegetation community observed in the field and the algal community, vegetation community variables thought most likely to influence algal community were also included into a dbRDA model. These included percentage cover of *Sphagnum* (of all species), *Erica tetralix*, *Cladonia portentosa* and *Trichophorum germanicum* (deer grass). These were chosen as species which, if dominating a quadrat, might influence the microenvironmental conditions experienced by the algal community. Due to a missing datapoint for *Sphagnum* percentage cover for a single sample, the Bray-Curtis distance matrix was re-calculated with the corresponding algal community sample removed.

Stepwise backward selection was carried out to minimise VIFs greater than 10 (calculated with cca.vif function in vegan- see supplementary Table 5.2 for further details). As with the model without ground vegetation variables, maximum tree height and mean tree DBH were chosen as the combination of height and DBH variables which minimised tree height and DBH VIFs when included in the model, and also resulted in the lowest VIFs for variables with VIF>10. The model was as follows:

*Model 3: dist ~ tree_stem_count + pH + ORP + soil_temperature + moisture/$m^3/m^3$ + moss_percent_water + max_tree_height + mean_tree_dbh + mean_canopy_cover + dominant_sphag_species + latitude + longitude + Calluna_vulgaris + Erica_tetralix + Trichophorum_germanicum + Cladonia_portentosa + percent_Sphagnum*

*Where dist is the Bray-Curtis dissimilarity of the algal community, tree_stem_count is the number of tree stems within 2m of the sampling point, ORP is the redox potential, moisture/$m^3/m^3$ is the soil percent moisture as measured by the soil moisture probe, moss_percent_water is the moss percentage water by weight, max_tree_height is the maximum tree height within 2m of the sampling point, mean_tree_dbh is the mean DBH for all trees within 2m of the sampling point, mean_canopy_cover is the calculated mean canopy cover from hemispherical photography, dominant_sphag_species is the dominant species of Sphagnum recorded in the vegetation quadrat at the sampling point, Calluna_vulgaris, Erica_tetralix, Trichophorum_germanicum and Cladonia_portentosa are the percentage cover of each species in the vegetation quadrat at the sampling point, and percent_Sphagnum is the percentage cover of all Sphagnum species combined in the vegetation quadrat at the sampling point*.

Reducing the model down to only one height or DBH measure further reduced VIFs for the tree height/DBH variables and mean canopy cover. The model including mean tree height had the lowest VIF for mean canopy cover, and all other tree variable VIFs below 10.

*Model 4: dist ~ tree_stem_count + pH + ORP + soil_temperature + moisture/$m^3/m^3$ + moss_percent_water + mean_tree_height + mean_canopy_cover + dominant_sphag_species + latitude + longitude + Calluna_vulgaris + Erica_tetralix + Trichophorum_germanicum + Cladonia_portentosa + percent_Sphagnum*

*Where dist is the Bray-Curtis dissimilarity of the algal community, tree_stem_count is the number of tree stems within 2m of the sampling point, ORP is the redox potential, moisture/$m^3/m^3$ is the soil percent moisture as measured by the soil moisture probe, moss_percent_water is the moss percentage water by weight, mean_tree_height is the*

*mean height of all trees within 2m of the sampling point, mean_canopy_cover is the calculated mean canopy cover from hemispherical photography, dominant_sphag_species is the dominant species of Sphagnum recorded in the vegetation quadrat at the sampling point, Calluna_vulgaris, Erica_tetralix, Trichophorum_germanicum and Cladonia_portentosa are the percentage cover of each species in the vegetation quadrat at the sampling point, and percent_Sphagnum is the percentage cover of all Sphagnum species combined in the vegetation quadrat at the sampling point.*

Automated model building using ordiR2step in vegan was also attempted to find a simplified model with variables which contributed to explaining a large amount of the variance. This resulted in the model:

*Model 5: dist ~ moss_percent_water + percent_Sphagnum + moisture_.m3.m3. + ORP*

*Where dist is the Bray-Curtis dissimilarity of the algal community, moss_percent_water is the moss percentage water by weight, percent_Sphagnum is the percentage cover of all Sphagnum species combined in the vegetation quadrat at the sampling point, moisture/$m^3$/$m^3$ is the soil percent moisture as measured by the soil moisture probe, and ORP is the redox potential.*

Mean canopy percentage cover was added into this model to test the effect of adding a tree cover variable into the simplified model. Mean canopy percentage cover was chosen out of the different tree cover variables because it included an indication of the shading effect of trees as well as other proximity effects, and also had strong positive correlations with the other tree variables. This resulted in the model:

*Model 6: dist ~ moss_percent_water + percent_Sphagnum + moisture_.m3.m3. + ORP + mean_canopy_cover*

*Where dist is the Bray-Curtis dissimilarity of the algal community, moss_percent_water is the moss percentage water by weight, percent_Sphagnum is the percentage cover of all Sphagnum species combined in the vegetation quadrat at the sampling point, moisture/$m^3$/$m^3$ is the soil percent moisture as measured by the soil moisture probe,*

*ORP is the redox potential, and mean_canopy_cover is the calculated mean canopy cover from hemispherical photography.*

#### *5.2.5.5 Mantel tests*

Mantel tests were carried out using Vegan to test for correlations between matrices.

To test for spatial autocorrelation, a Mantel test was carried out between the Bray-Curtis distance matrix for algal communities vs. the geospatial distance. Haversine geospatial distances were calculated from latitude-longitude data using the distm function in the geosphere package (Hijmans, 2024).

A Mantel test was also used to test for a correlation between the plant communities identified through molecular work and identified via quadrat in the field. Correlations between the algal community Bray-Curtis distance and the Bray-Curtis distances for the molecular plant community and field recorded plant community were also tested for. As one plot had data missing for the field plant community, the corresponding sample was removed from the algal community and molecular plant community to ensure matrices aligned.

#### *5.2.5.6 Changes in community composition*

Composition plots were created in R using ggplot2 (Wickham, 2016) and the function plot_composition from the package microbiome (Lahti *et al*., 2017). Where aggregation of rare taxa was necessary to visualise the data, this was carried out at the minimum level possible to produce a legible graph and has been recorded in relevant figure legends.

## 5.3 Results

### 5.3.1 Trees and canopy cover

90 trees were counted in total across the 30 plots. The number of tree stems within 2m of the sampling point ranged between 0 and 11, encompassing a range of tree densities found at the site (Fig. 5.15). This broadly divides into 7 treeless sites, 10 sites with a

small number of stems (1-2), 7 sites with a medium number of stems (3-5) and 6 sites with a high number of stems (6 and above). There was no attempt to treat stem number categorically in the field, but instead to sample across the range of the ecotone, maximising heterogeneity. Tree density is therefore treated as a numerical rather than categorical variable.

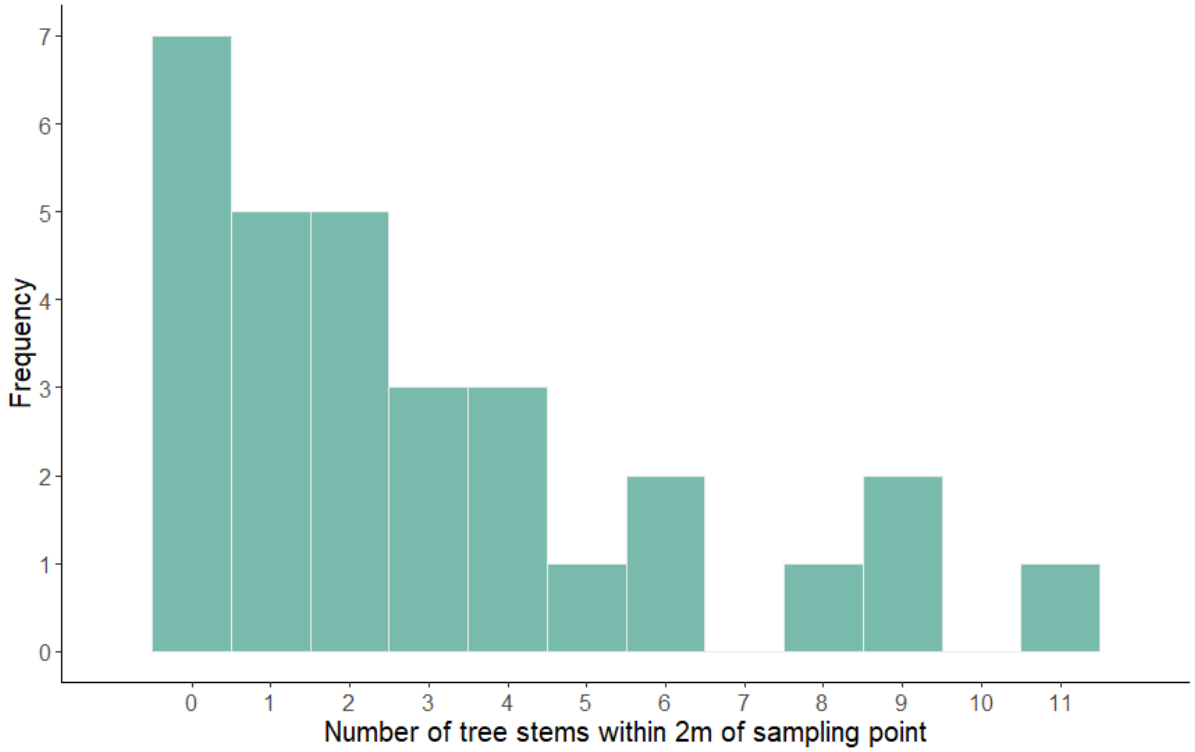



**Figure 5.15.** Histogram showing the frequency of plots with different numbers of tree stems within 2m of the sampling point.

Most trees were scored as healthy (n = 56), but many appeared to be unhealthy or visibly stressed (health score 2, n = 32). Only two dead trees (health score 1) were included in the experiment. Most sites with trees scoring 2 also included apparently healthy trees (Fig. 5.16).

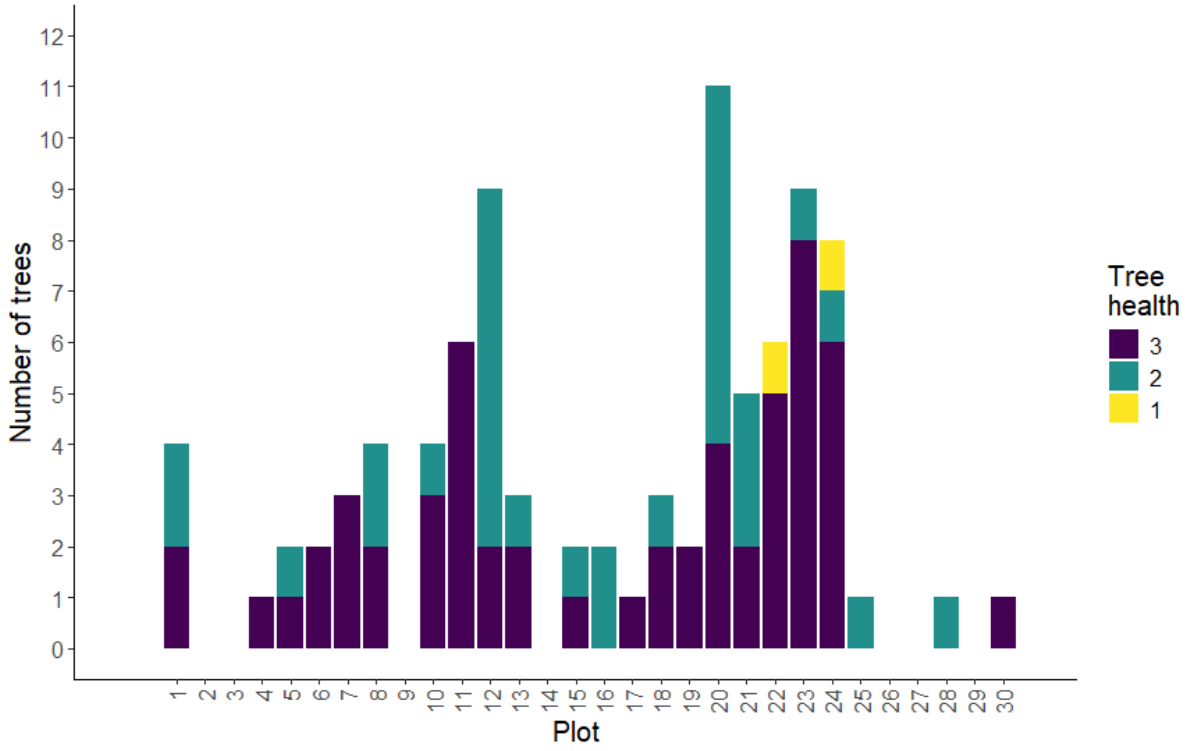



**Figure 5.16.** Number of trees of each health category by plot. Trees with a health score of 3 were considered healthy, those with a score of 2 were visibly unhealthy or stressed, and those with a score of 1 were dead. For more details, see Section 5.2.2.

Of the 90 trees measured in this study, 65 were tall enough for DBH to be measured. The shortest tree was 15cm tall and the tallest was 1240cm. The frequency of trees between 1 and 8m was relatively steady, with fewer trees above 8m high (Fig. 5.17).

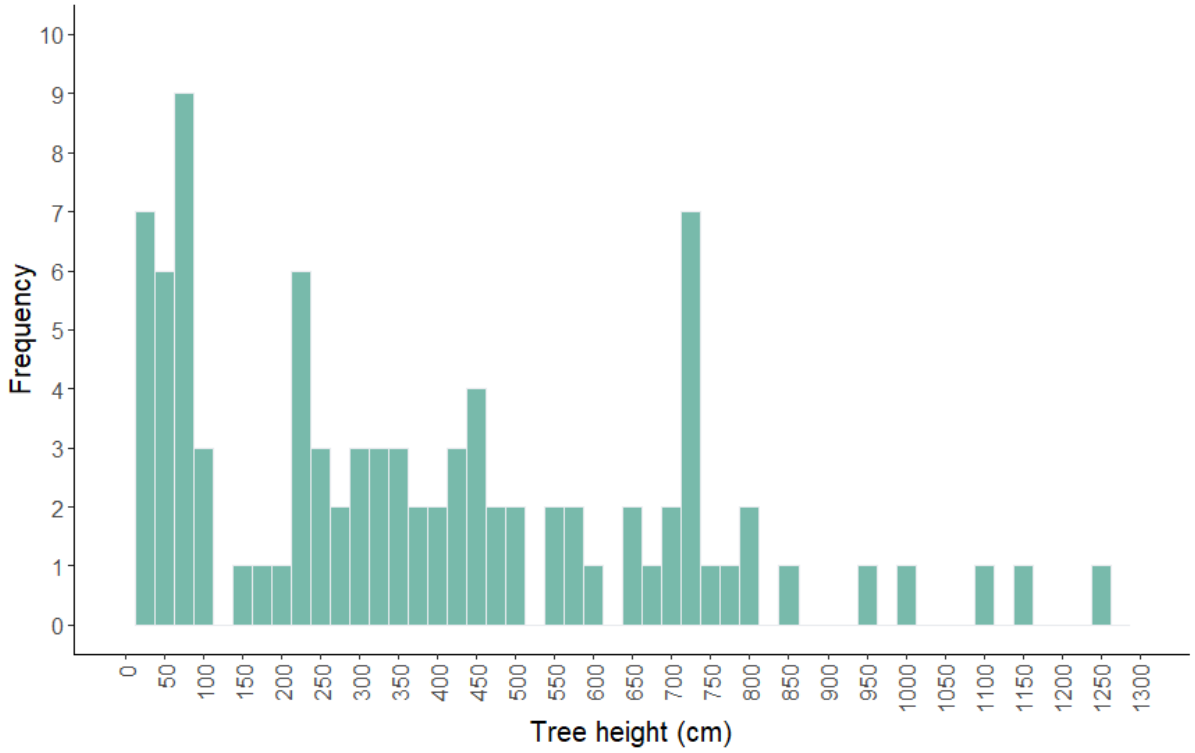



**Figure 5.17.** Histogram of tree heights sampled in Chapter 5.

The majority of trees had a DBH of below 20cm, with many having low DBH (Fig. 5.18).

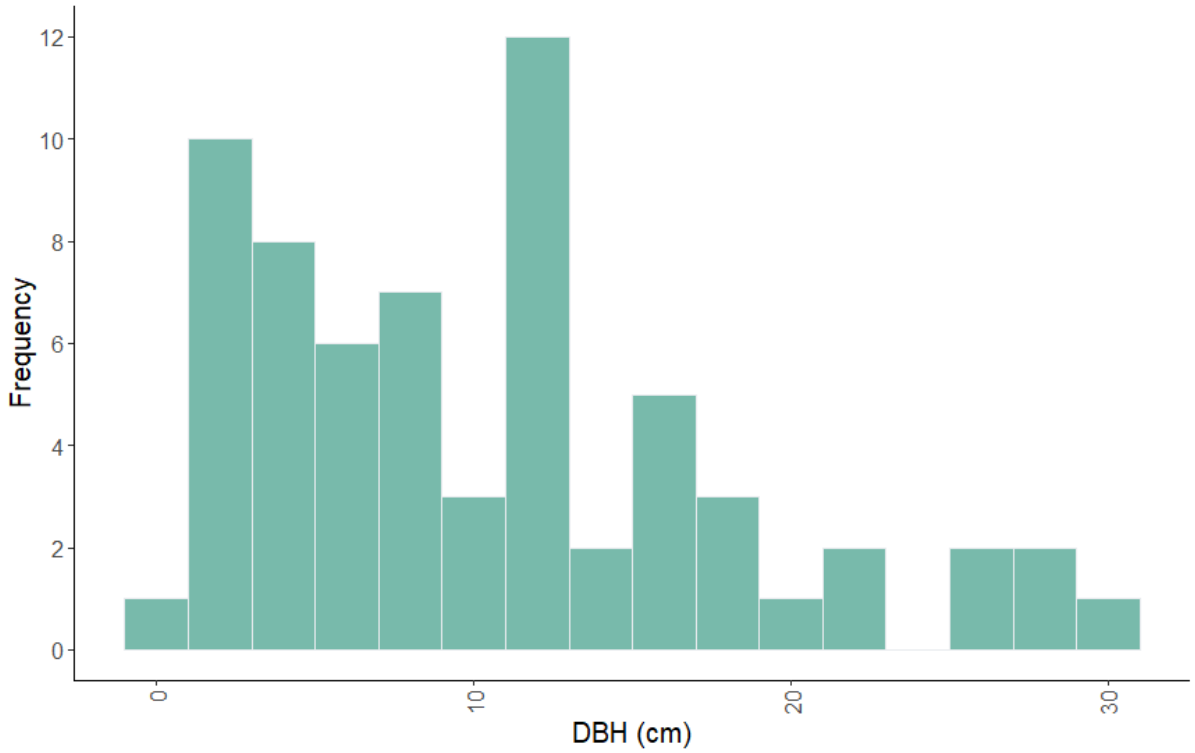



**Figure 5.18.** Histogram of DBH (diameter at breast height) measurements for trees surveyed in Chapter 5.

The highest calculated canopy cover was 61.74% cover. The lowest canopy cover calculated from an image was 0.35%. 5 plots had a canopy cover of 0, creating a skew towards lower canopy cover levels (Fig. 5.19). However, higher percentage canopy covers were still well represented in the dataset.

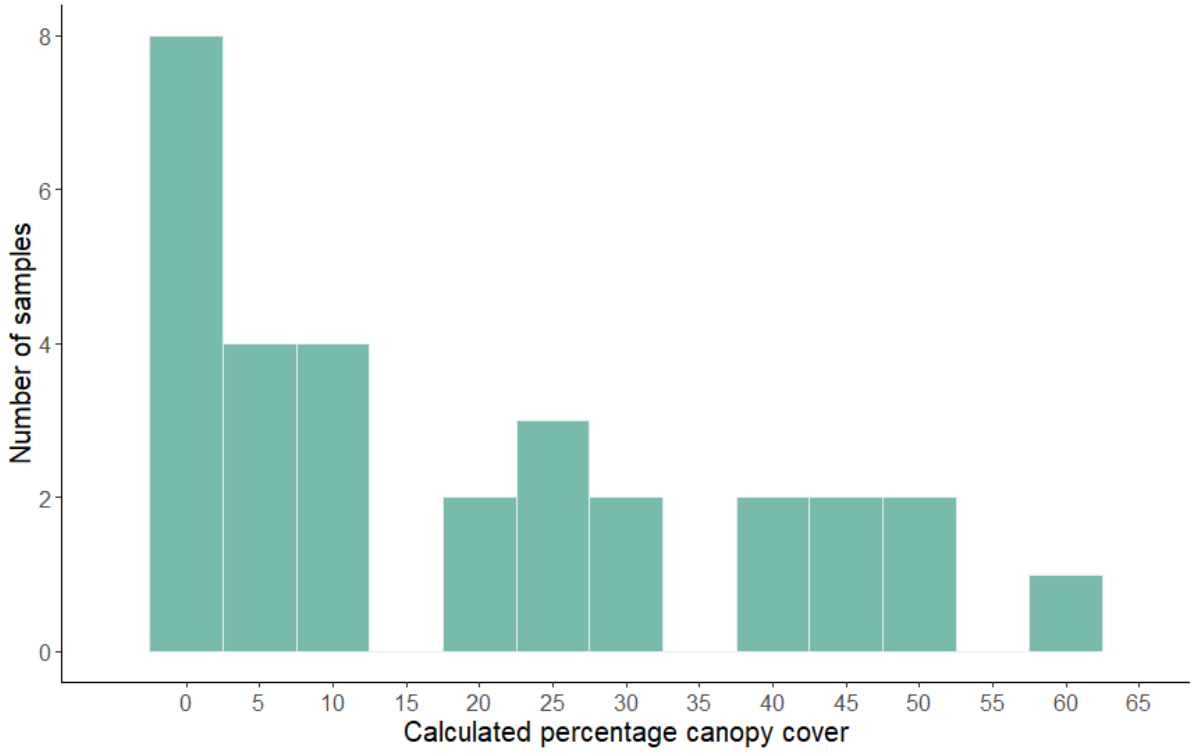



**Figure 5.19.** Histogram of calculated percentage canopy cover for plots in Chapter 5. Three samples with calculates canopy cover 0<X<2.5 were pooled with the treeless plots when creating this histogram, resulting in the higher count in the lowest category.

### 5.3.2 Environmental variables

Bryophyte percentage water was very high, between 93% and 100% for all samples (Fig. 5.20). This was much higher than the measured soil moisture at 15cm deep (Fig. 5.21).

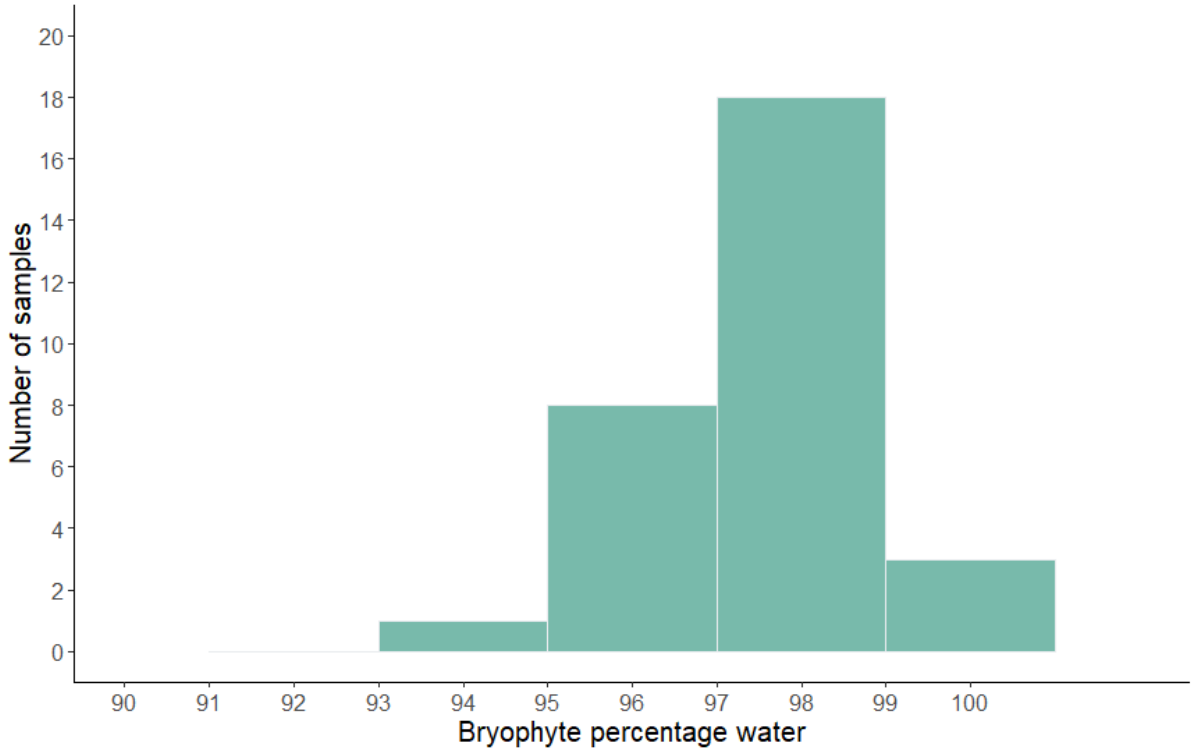



**Figure 5.20.** Histogram of bryophyte percentage water measurements by weight for samples in Chapter 5.

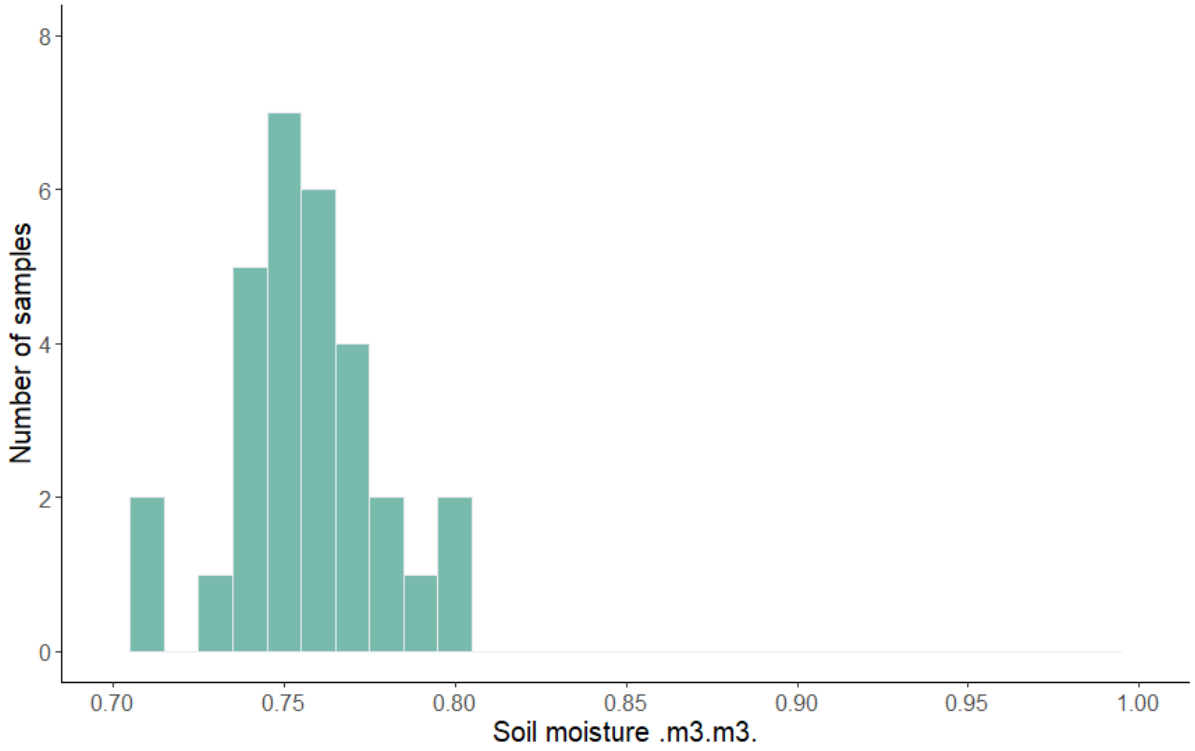



**Figure 5.21.** Histogram of soil moisture probe measurements at 15cm below the moss surface for samples in Chapter 5.

The majority of measured ammonium concentrations fell between 0.024 and 0.1mg $L^{-1}$, with one measurement (sample A24) of 0.509. The highest nitrate measurement was 0.117mg $L^{-1}$. Five nitrate measurements were below zero, likely due to fluctuations around the baseline. All nitrate measurements and all except one ammonia measurement were below the measured detection limits for the same machine reported in Khan (2022). Some nitrate measurements were above and some below the measured detection limit reported in Hodson *et al*. (2023) (Fig. 5.22). As these samples were not diluted prior to running on the AA3, this suggests extremely low pore water nutrient levels. Any differences in pore water nutrient concentrations between samples should be treated with caution in case of instrument error near or below minimum real world detection levels.

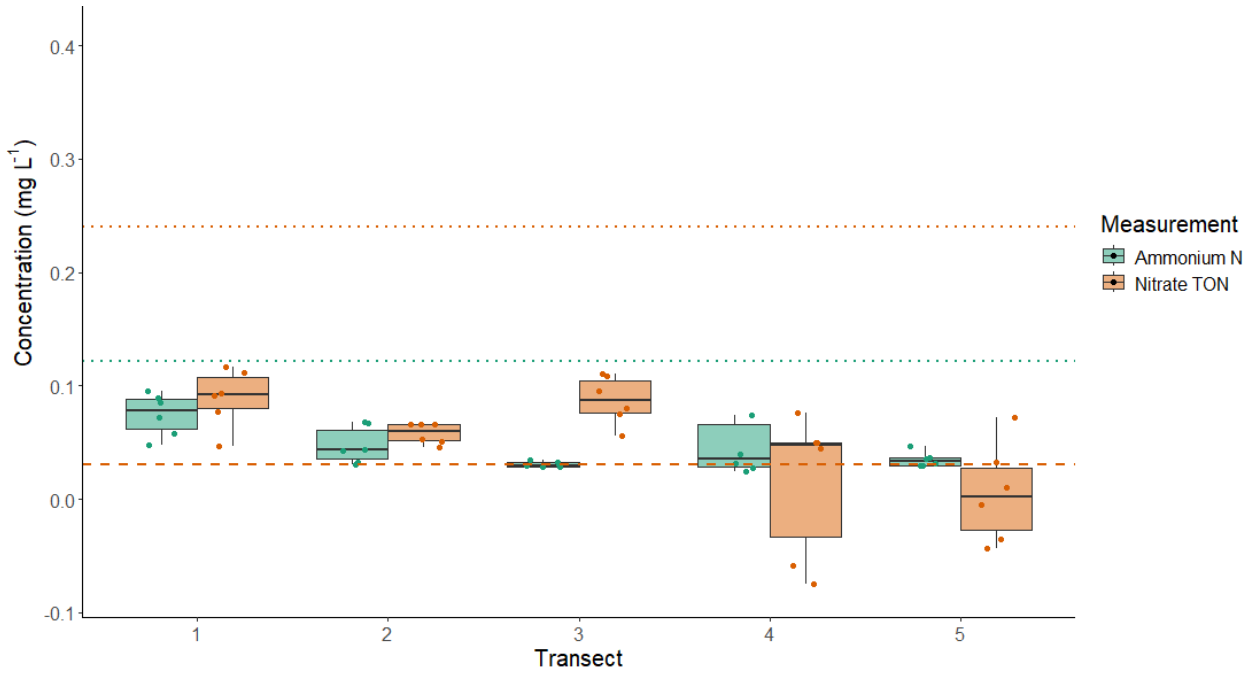



**Figure 5.22.** *Sphagnum* pore water nutrient concentrations from Abernethy, separated into different transects for ease of viewing. Points are the individual sample scores. The dashed line shows the measured real-world detection limits for nitrate on the same machine reported in Hodson et al. (2023). Dotted lines show measured detection limits for nitrate and ammonia reported in Khan (2022). All measurements except one ammonia measurement are below the detection limits from Khan (2022). Some nitrate measurements are above and some below the measured detection limit from Hodson et al. (2023).

The data loggers were in place from late October 2021 (when sampling took place) to July 2023. Moisture levels were relatively constant at all three sites between October 2021 and June 2022, with lower moisture levels and greater variability in summer and autumn (Fig. 5.23). The bog site had the highest moisture level when conditions were stable, and the forest site the lowest; however during the summer and autumn there were some points at which the ecotone site had the highest moisture levels. The open bog generally experienced the greatest extremes of temperature of the three sites, with higher surface temperatures in summer and lower in winter (Fig. 5.24). The logger in the forest had less extremes of temperature, and the ecotone logger was intermediate between the two.

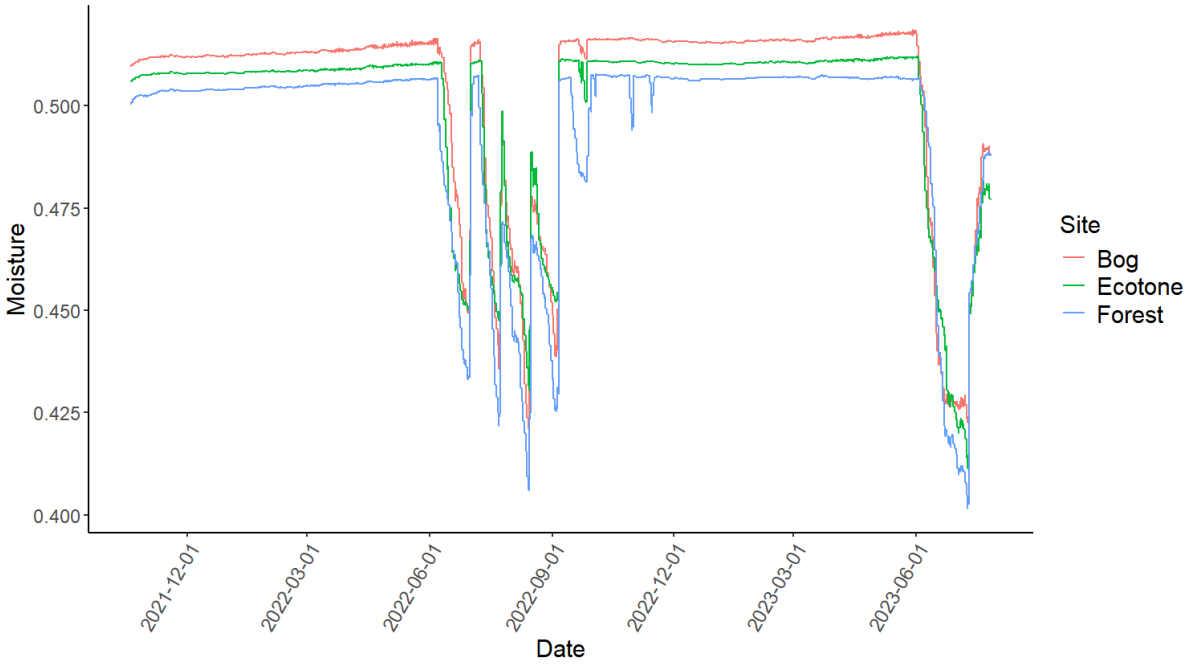



**Figure 5.23.** Long-term soil moisture measurements from data loggers for open bog, forest and forest-bog ecotone locations in the study area. Data recorded from Autumn 2021 (date of study) to Summer 2023.

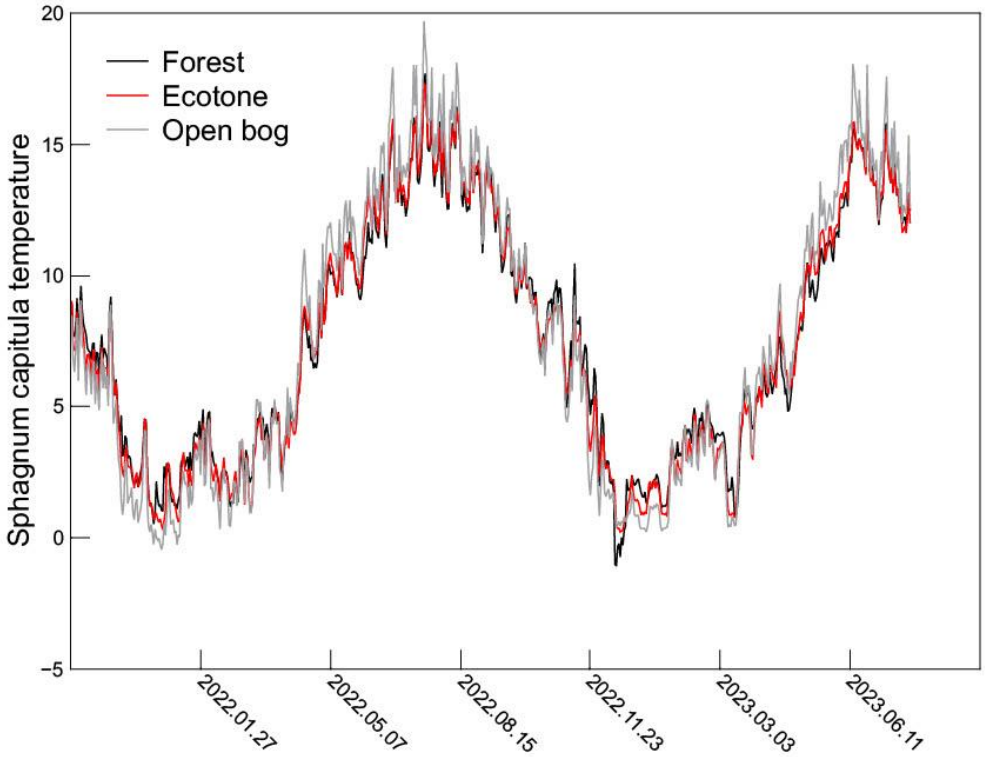



**Figure 5.24.** Long-term surface temperature measurements from data loggers for open bog, forest and forest-bog ecotone locations in the study area. Data recorded from Autumn 2021 (date of study) to Summer 2023. Graph by R. Mills.

### 5.3.3 Transect effects

Kruskal-Wallis tests carried out on transect effects on diversity indices calculated from the first bioinformatics method all showed no significant differences in algal diversity (Simpsons index $p = 0.2351$, Shannon $p = 0.172$).

Variance in algal community data (as identified with the first bioinformatics method) between different transects was not significantly heterogeneous ($F = 2.06$, $p = 0.108$). PERMANOVA showed that algal communities (as Bray-Curtis distance matrices) were statistically significantly different between transects ($p < 0.001$). Pairwise PERMANOVA showed significant differences between transects 1 vs 2 ($p = 0.03$) and transects 2 vs 3 ($p = 0.03$). All other transect pairs did not show a statistically significant difference.

Variance in algal communities as identified with the second bioinformatics method between different transects was not significantly heterogeneous (F = 0.87, p = 0.495). PERMANOVA showed that algal communities were not statistically significantly different between transects (p = 0.107) for the data from the second bioinformatics method.

For the ratio of prokaryotes to eukaryotes, there was no significant difference in homogeneity of variance between transects (Bartlett's K-squared = 8.3778, df = 4, p = 0.079). Despite transect 4 containing both of the samples with the highest percentage of prokaryote reads, a Kruskal-Wallis test showed no significant differences in the ratio of prokaryotes to eukaryotes between transects (p = 0.3822) (Fig. 5.25).

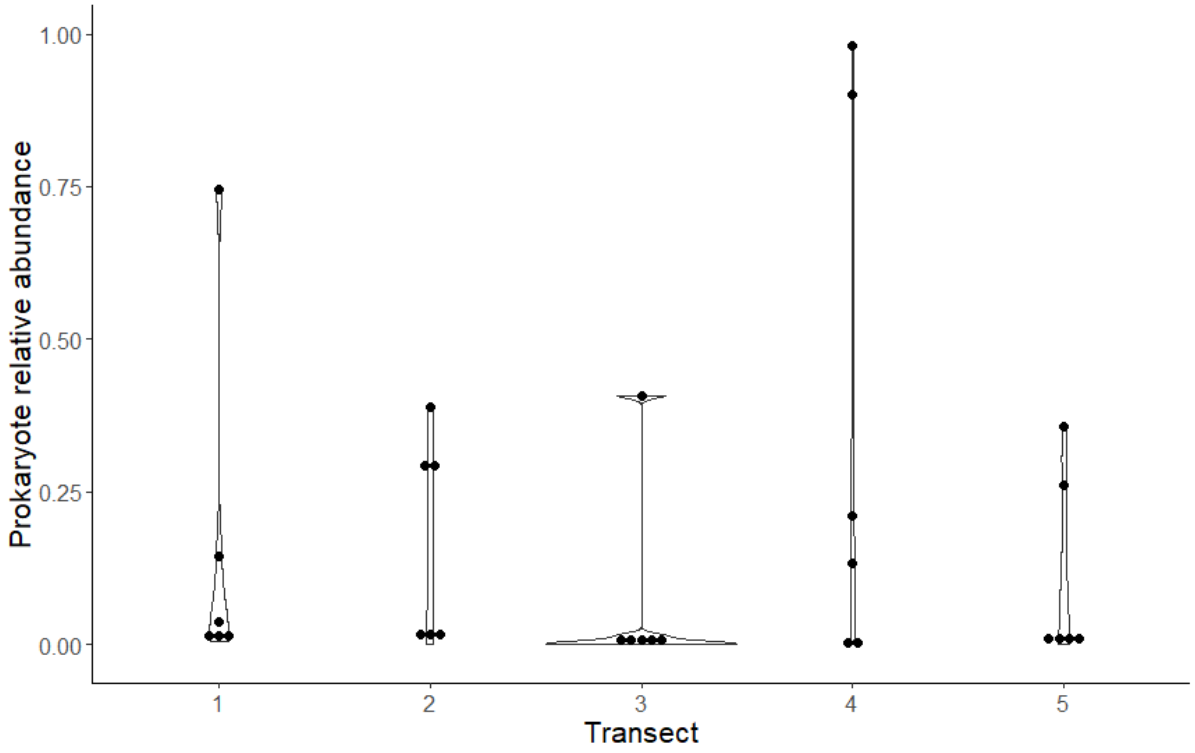



**Figure 5.25.** Relative abundance (as a proportion of total reads) of prokaryotes between transects for Chapter 5. Points show actual data.

There is a significant difference in the number of tree stems within 2m of sampling between transects (Kruskal-Wallis: p = 0.002962) (Fig. 5.26). As site selection was carried out based on canopy cover, the distribution of the number of tree stems within

2m of each plot was skewed, with more plots with no trees (open bog) than with high number of trees (Fig. 5.15). This reflected the makeup of the forested bog observed in the field, with most *Sphagnum* either in the open or with few nearby trees.

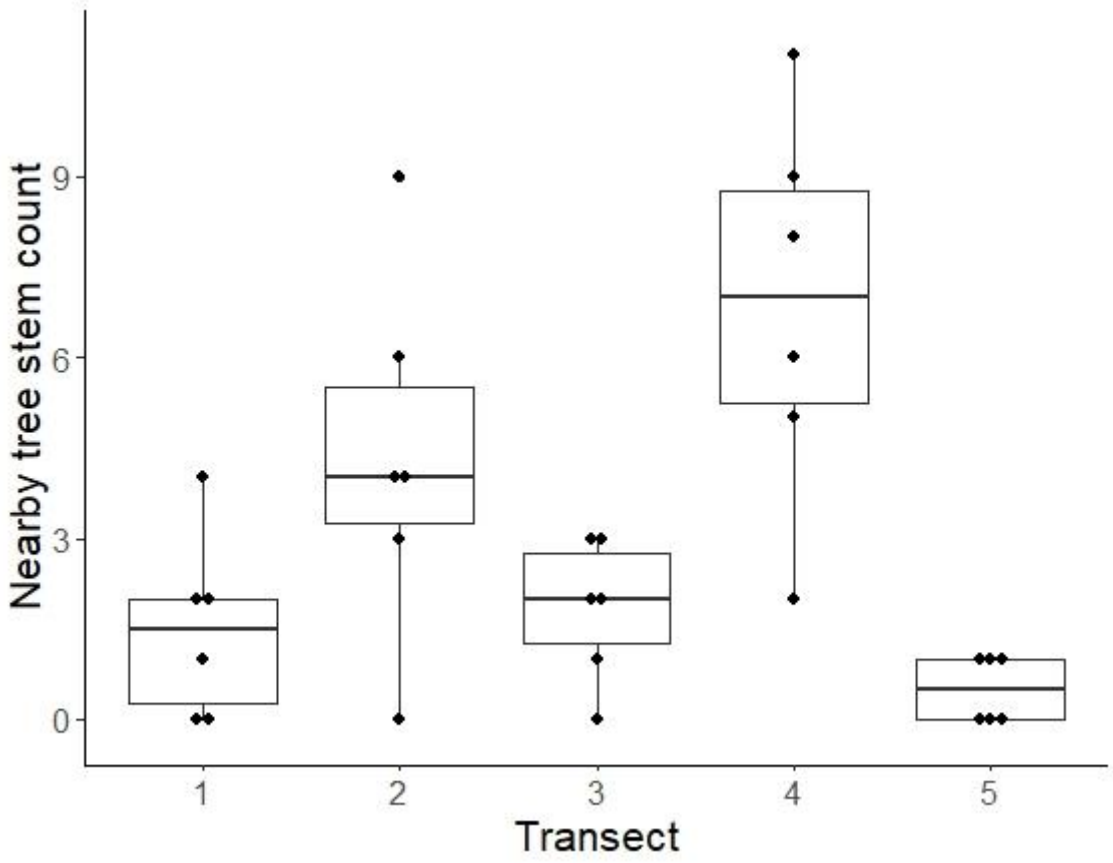



**Figure 5.26.** Variation in number of tree stems within 2m of sampling point by transect. Box shows median and interquartile range. Points show actual data points.

## 5.3.4 Community composition

### *5.3.4.1 First bioinformatics method*

As seen with the data from Whim (Chapter 4), the majority of ASVs were unclassified at the class and phylum levels (Figs. 5.27. and 5.28) with the first bioinformatics method.

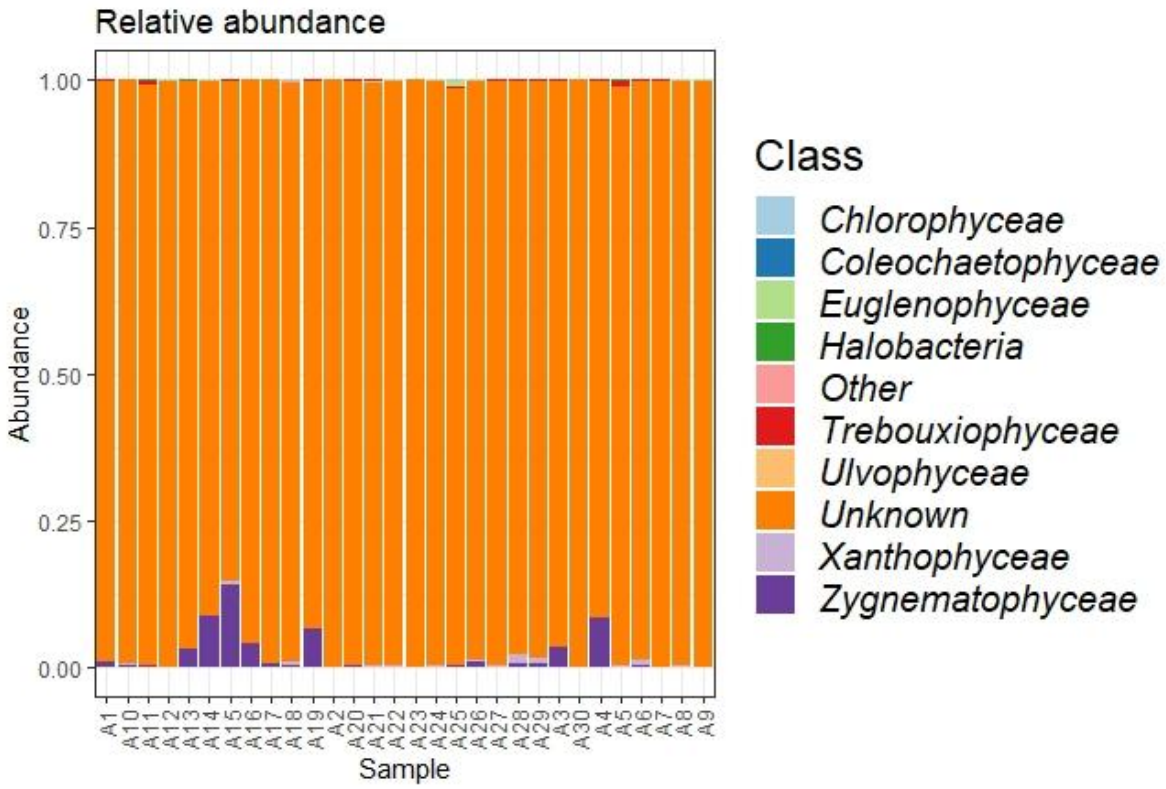



**Figure 5.27.** Relative abundance (proportional) of the most abundant microbial classes without filtering to remove unclassified ASVs (first bioinformatics method).

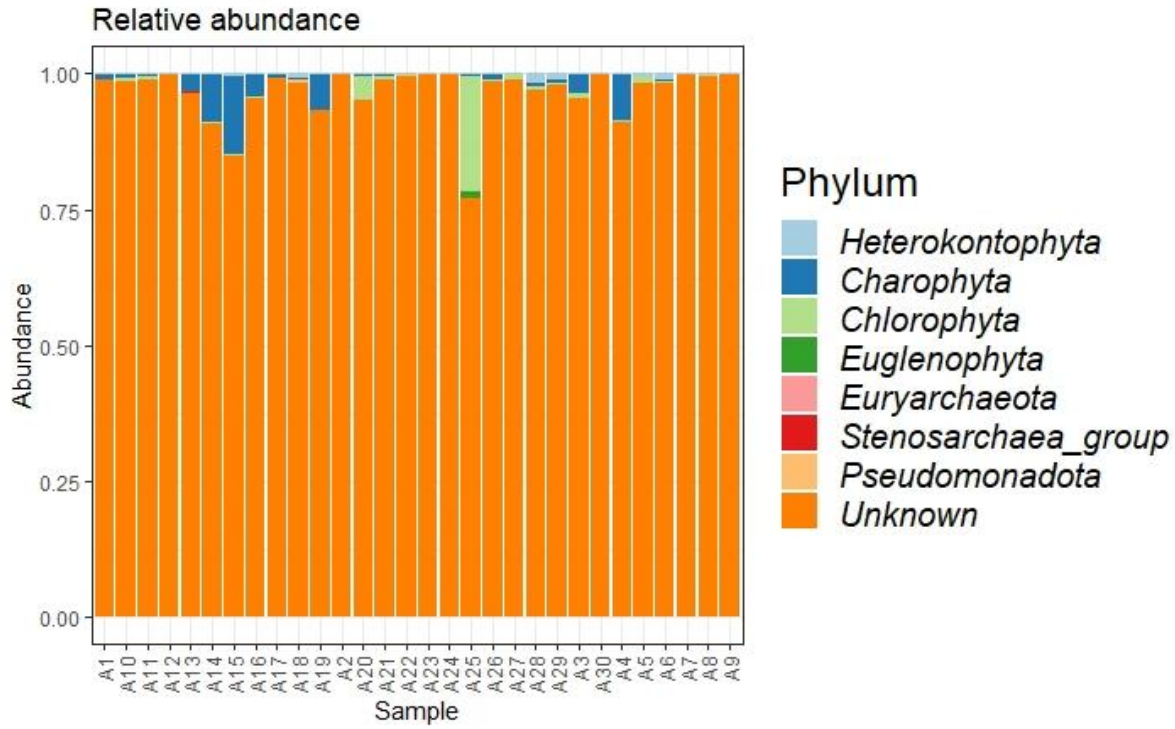



**Figure 5.28.** Relative abundance (proportional) of the most abundant microbial phyla without filtering to remove unclassified ASVs. Even at this taxonomic level, there are so many unknown ASVs that drawing conclusions about any patterns in abundance shifts from these figures is almost impossible. Data from the first bioinformatics method.

When ASVs unknown at the phylum level are filtered out, a much clearer indication of community variation can be seen. At the phylum level, communities in some samples are clearly dominated by Charophyta. However, other samples are dominated by Heterokontophyta or Chlorophyta (Fig. 5.29).

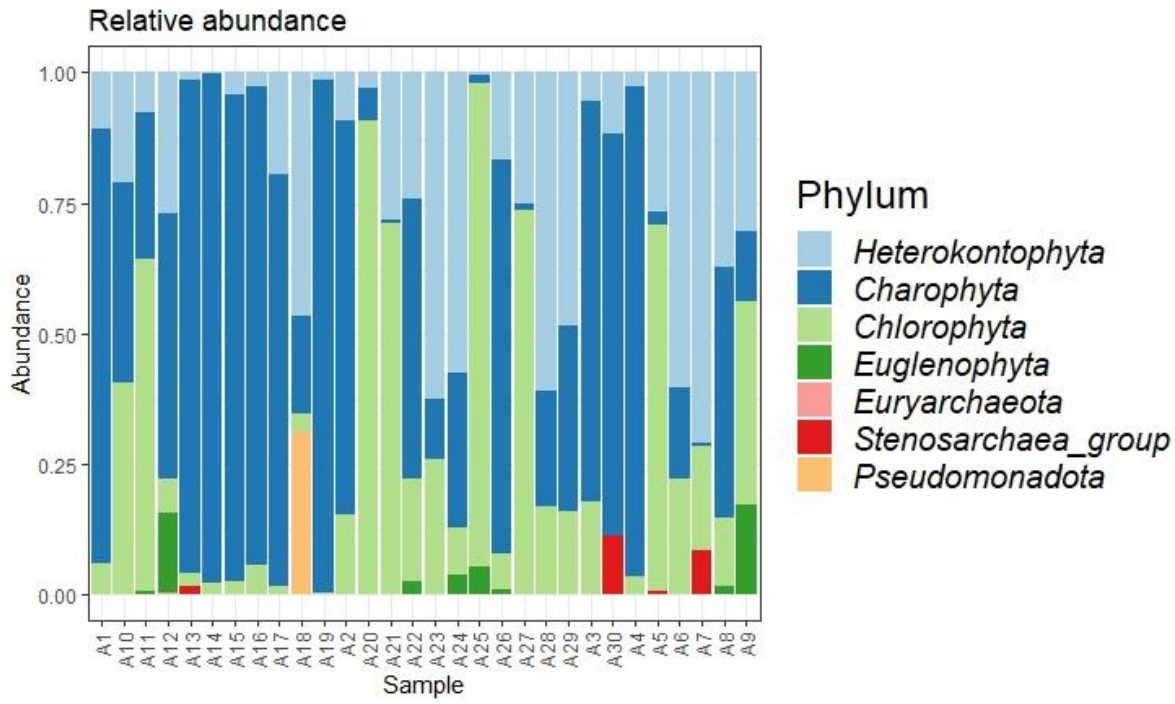



**Figure 5.29.** Relative abundance (proportional) of the most dominant microbial phyla after removal of ASVs not identified to phylum level. Data from the first bioinformatics method.

Of the ASVs classified at the phylum level, Zygnematophycea are the most dominant class for many samples (Fig. 5.30). Xanthophycea relative abundance is also high in some samples, but more variable in others. In some samples, the majority of ASVs classified to phylum level are unclassified to class level.

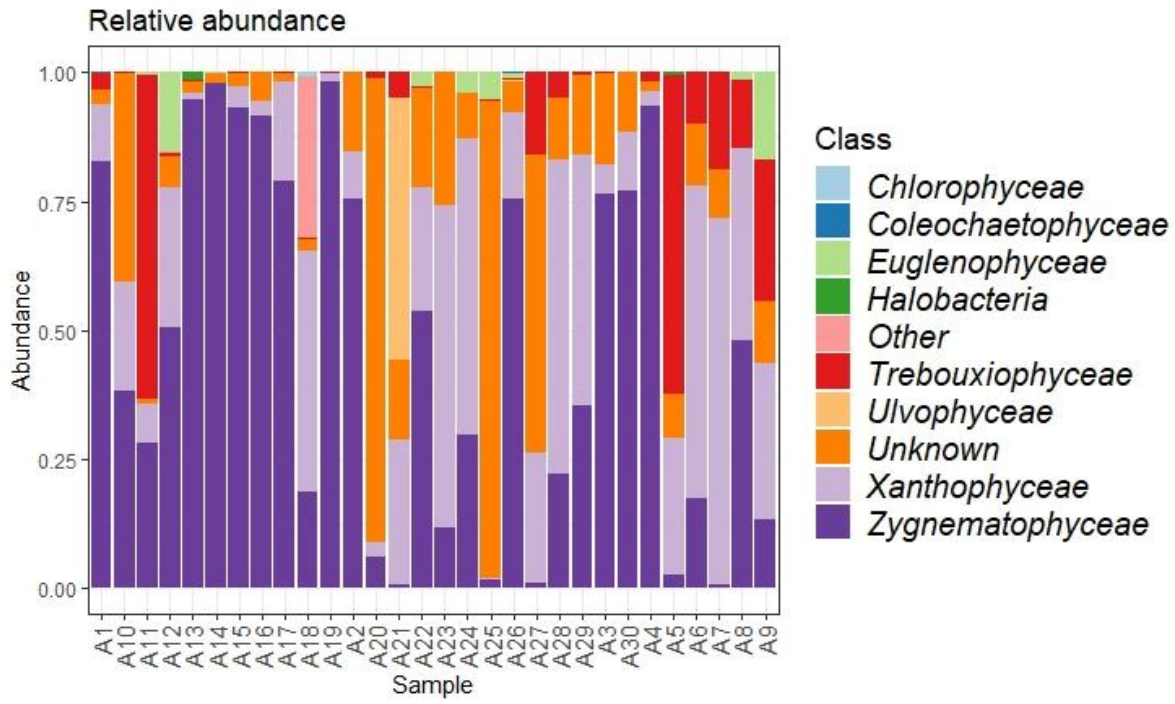



**Figure 5.30.** Relative abundance (proportional) of the most dominant microbial classes after removal of ASVs not identified to phylum level. Data from the first bioinformatics method.

### *5.3.4.2 Second bioinformatics method- taxonomic assignment with Kraken2*

For most samples, the majority of all reads (before filtering to separate microbial and plant reads) were assigned to *Sphagnum* spp. (Fig. 5.31). However, some samples had lower relative abundance of *Sphagnum* reads. The two samples with <50% of reads assigned to *Sphagnum* spp. (A5 and A27) were still dominated by Embryophyta (97.98% and 99.49% of identified reads, respectively). The Sphagnales were still the most abundant order in sample A27, and the Jungermanniales (of which the majority of reads were assigned to *Mylia anomala*) were the most abundant order in sample A5.

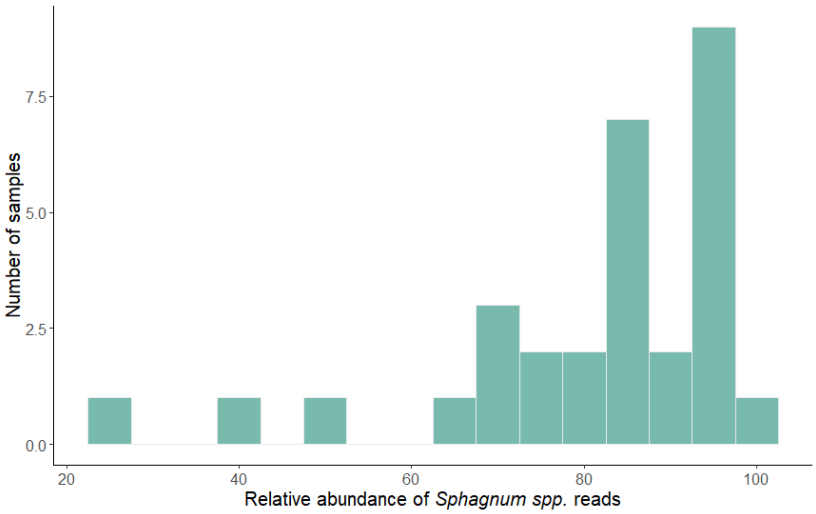



**Figure 5.31.** Histogram showing frequency of relative abundance of *Sphagnum* spp. reads between Chapter 5 samples. Data from second bioinformatics method.

#### 5.3.4.2.1 Microbial communities

Using Kraken2 for taxonomic assignment, the percentage of reads assigned to a taxonomy was much greater than with the original method (see Section 3.5.2).

For most samples, the majority of microbial reads were eukaryotic (Fig. 5.32). However, three samples had higher than 50% prokaryotic reads, and Kraken2 assigned the majority of reads for samples A23 and A24 to prokaryotes.

While some samples had very high relative abundances of prokaryotes, these were not cyanobacteria reads (Fig. 5.33). Cyanobacteria were only detected at very low abundances. The ratio of chlorophyte to streptophyte algae was variable, but most samples were dominated by streptophytes. However, chlorophytes did dominate some samples.

Where Streptophyta dominated, these appeared to be mostly Zygnemophyceae (Fig. 5.34). Trebouxiophycea were the most abundant chlorophyte class. The majority of Zygnemophyceae reads appeared to be Desmidiales, which dominated most of the samples (Fig. 5.35). Bryopsidales, Prasiolales and Chlorellales were the dominant Trebouxiophycea for most samples.

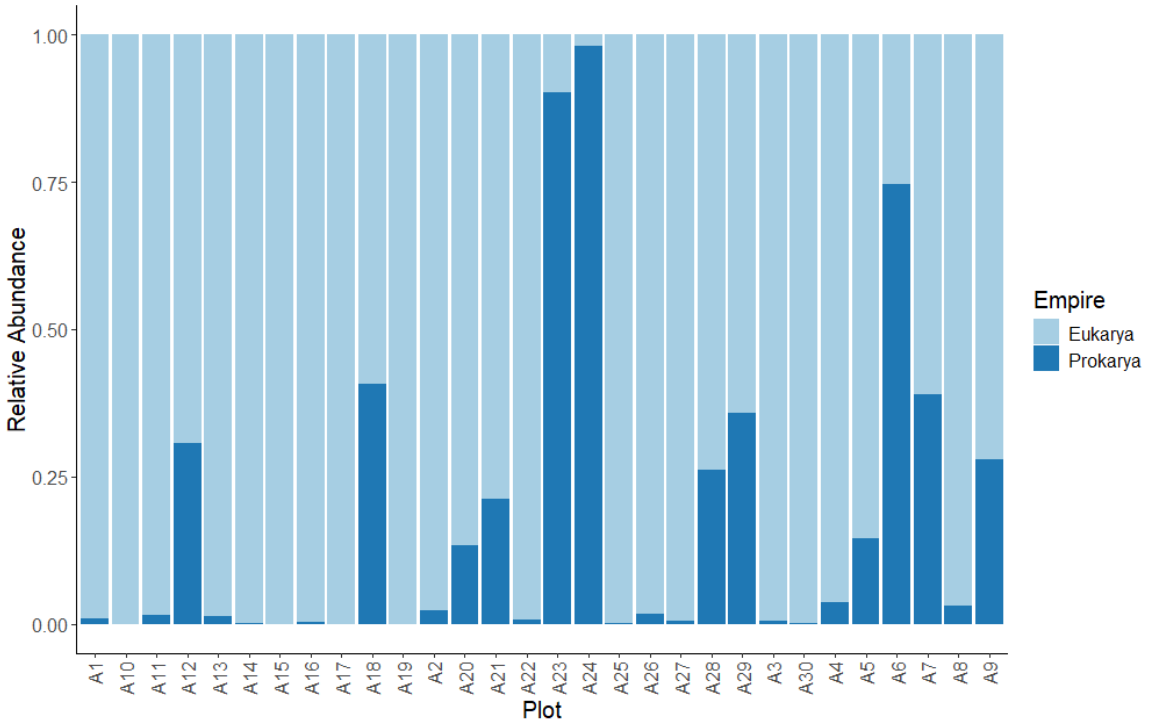



**Figure 5.32.** Composition plot of all microbes (full data set with plants removed) showing the relative abundances (proportional) of Eukarya and Prokarya for each sample, with data from the second bioinformatics method.

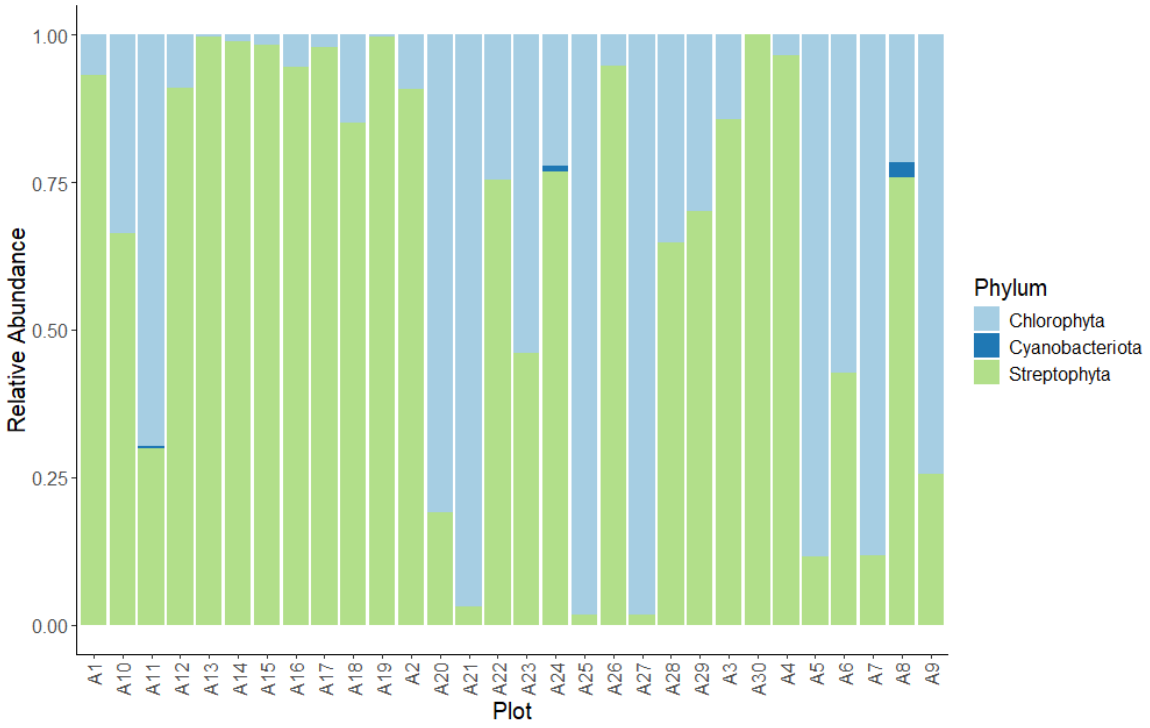



**Figure 5.33.** Relative abundance (proportional) of algal taxa at the phylum level with data from the second bioinformatics method. Note that Kraken2 uses different naming conventions to DADA2, and Charophyta and Streptophyta are synonymous.

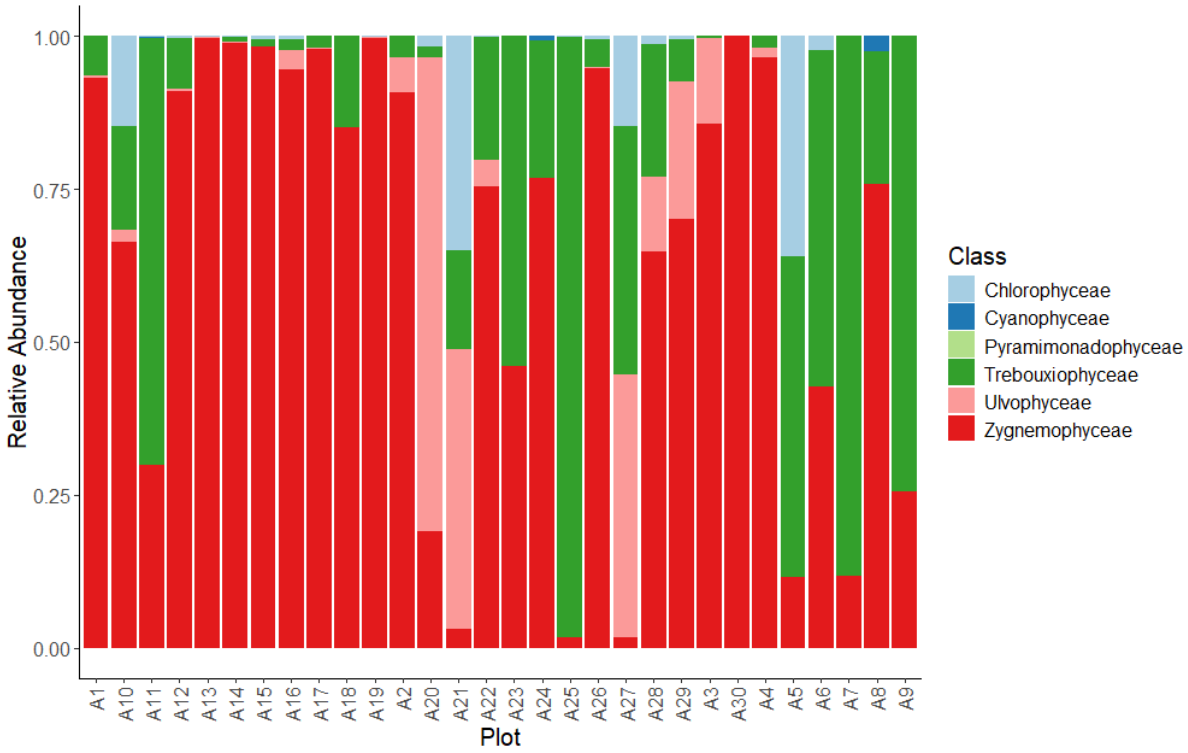



**Figure 5.34.** Proportional composition plot of algal taxa from Chapter 5 samples at the class level, with data from the second bioinformatics method.

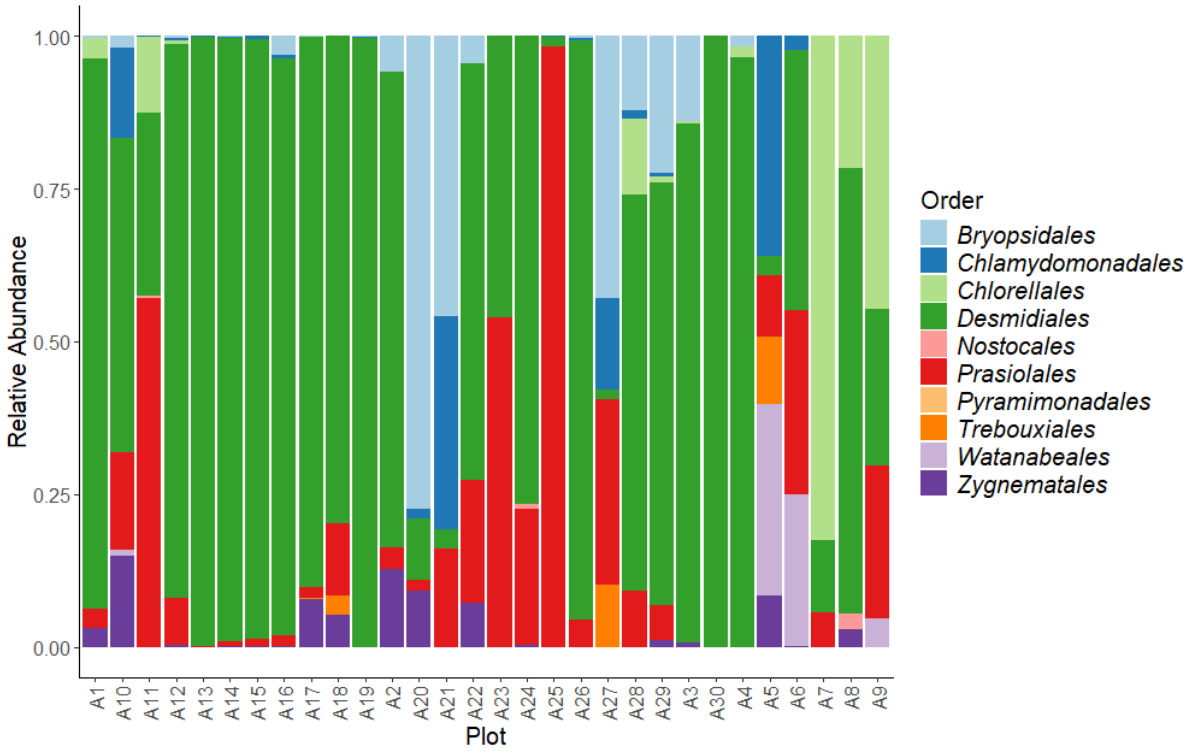



**Figure 5.35.** Proportional composition plot of algal taxa at the order level, with data from the second bioinformatics method.

When looking at all microbes (including the non-algal taxa) the algal taxa continue to dominate, except for the samples with more prokaryotes than eukaryotes. Where Prokarya were dominant, these appeared to be mostly Pseudomonadota (Fig. 5.36). At the class level, most Pseudomonadota reads were Alphaproteobacteria, with Actinomycetes and Terriglobia less abundant (Fig. 5.37).

Looking at all microbes, Desmidiales were still clearly dominant in most samples (Fig. 5.38). Caulobacterales were the most abundant prokaryote order. The dominance of Desmidiales varied between samples, with relative abundance in some samples being very low.

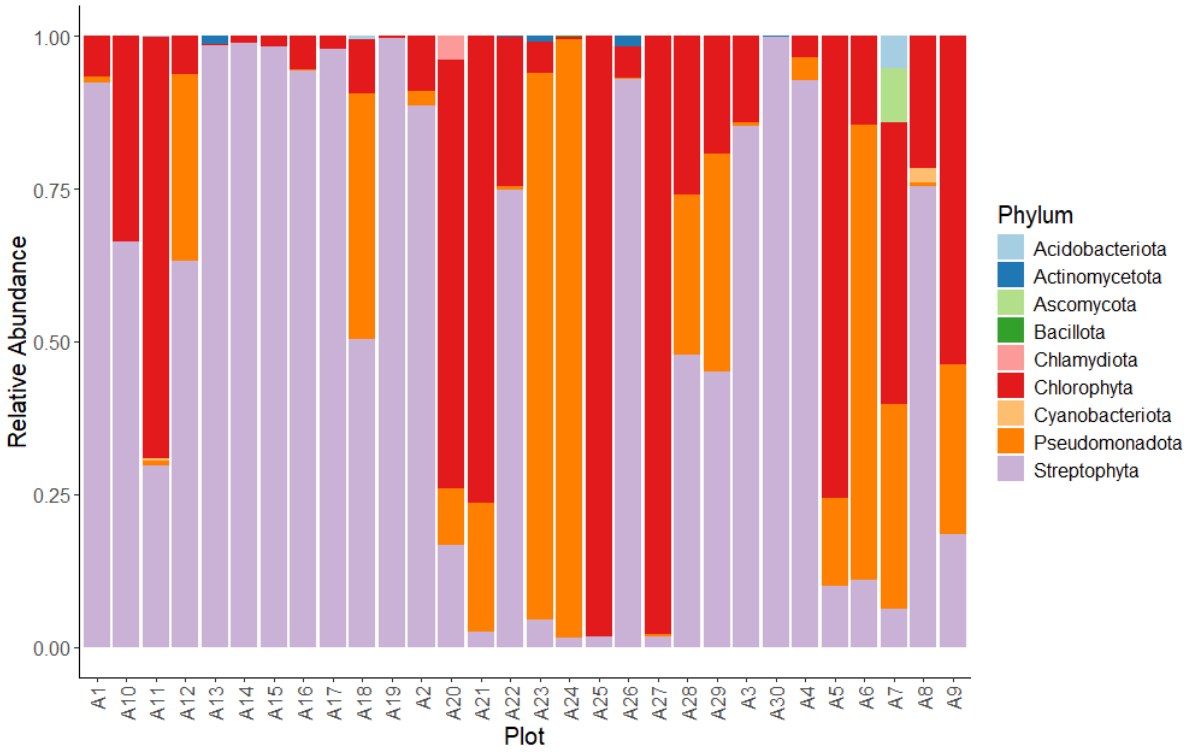



**Figure 5.36.** Proportional composition plot of all microbial reads at the phylum level, plotted with ggplot2. Data from the second bioinformatics method.

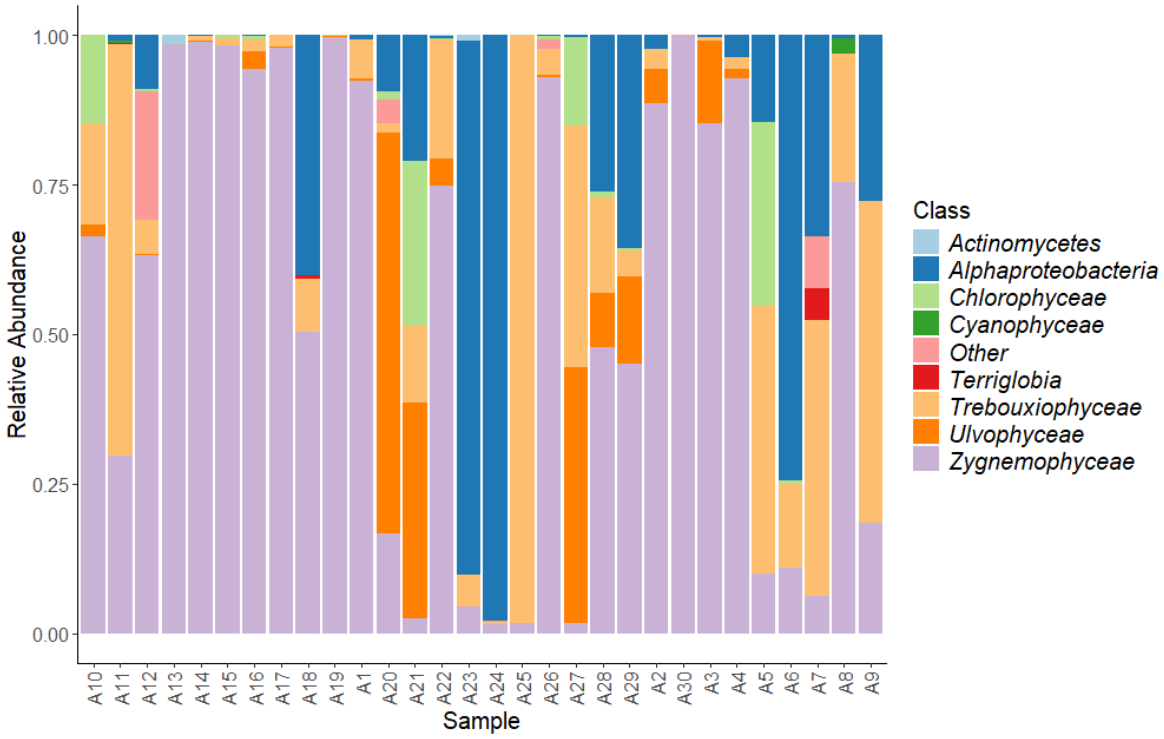



**Figure 5.37.** Proportional composition plot of all microbes at the class level, plotted in R using the function microbiome. Rare taxa were aggregated at prevalence = 5/100, detection = 0.1/100. Data from the second bioinformatics method.

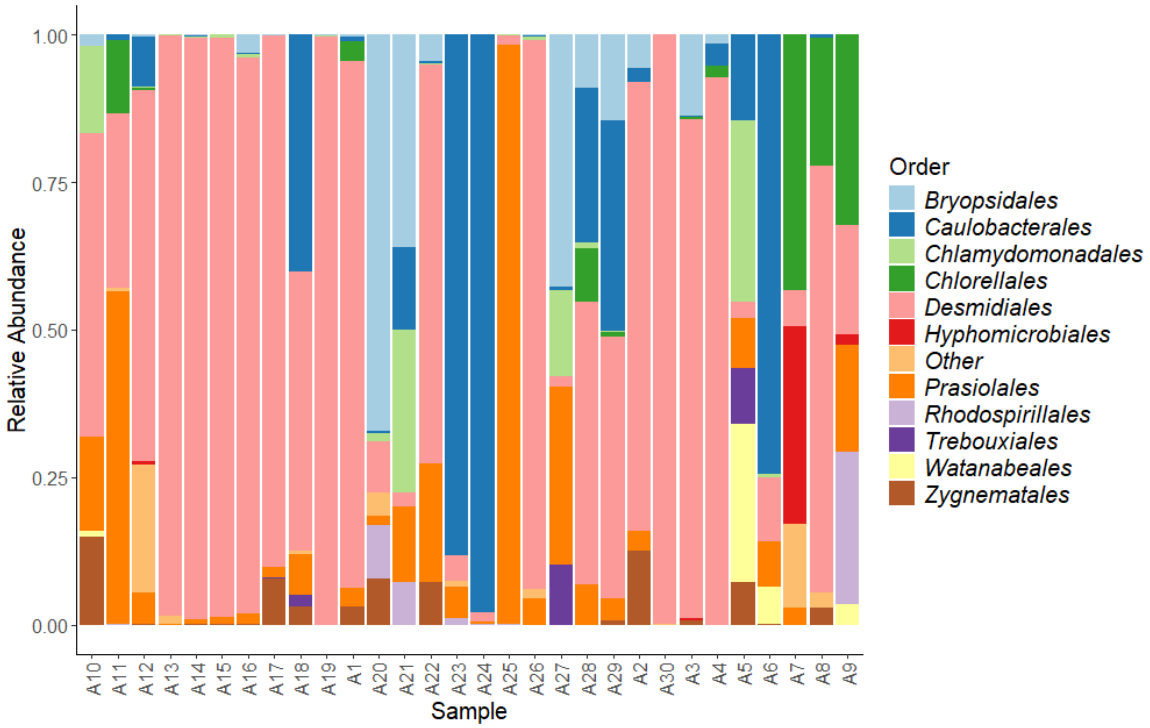



**Figure 5.38.** Proportional composition plot of all microbes at the order level, plotted in R using the function microbiome. Rare taxa were aggregated at prevalence = 10/100, detection = 0.1/100. Data from the second bioinformatics method.

#### 5.3.4.2.2 Molecular plant communities

Unsurprisingly, considering the proportion of total reads, Sphagnopsida dominated the plant community (Fig. 5.39). Other non-vascular plant classes, especially the Jungermanniopsida, were also abundant. The Magnoliopsida (flowering plants) were abundant in some sites but had very low relative abundance in others.

Of the 11 most abundant plant orders, five were vascular (Caryophyllales, Ericales, Lycopodiales, Poales, Polypodiales). However, these were, in most samples, much less abundant than the non-vascular plants (Fig. 5.40).

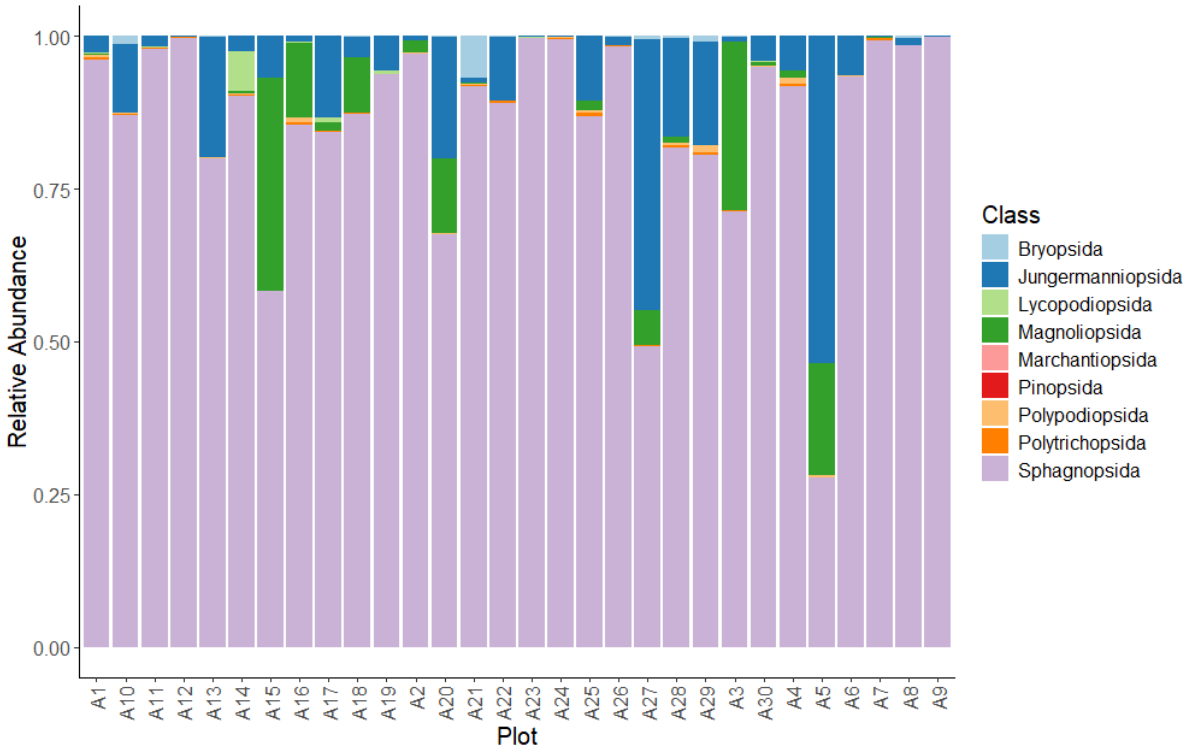



**Figure 5.39.** Proportional composition plot of all plants at the class level, with data from the second bioinformatics method.

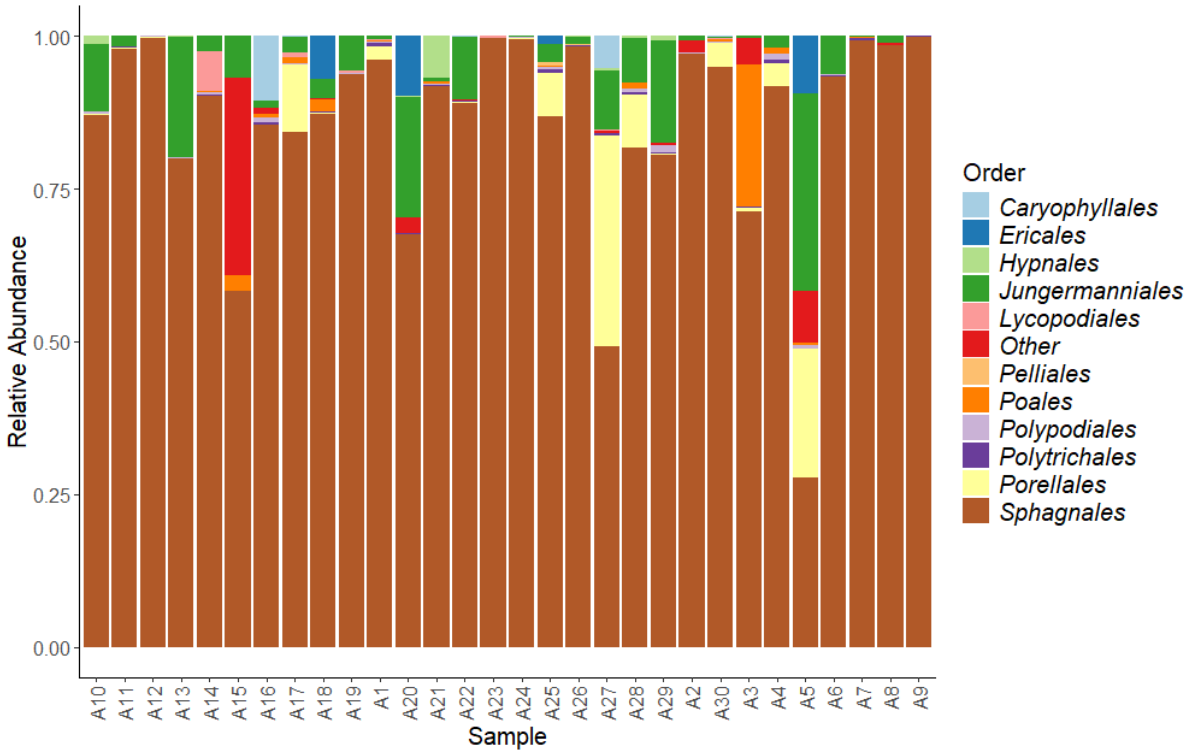



**Figure 5.40.** Proportional composition plot of all plants at the order level, plotted in R using the function microbiome. Rare taxa were aggregated at prevalence =2 5/100, detection = 1/100. Data from the second bioinformatics method.

Looking only at the bryophytes, Sphagnales were extremely dominant at the order level (Fig. 5.41). The Jungermanniopsida were mostly dominated by Jungermanniales and Porellales. The Hypnales were abundant in one sample, and all other orders had very low relative abundance.

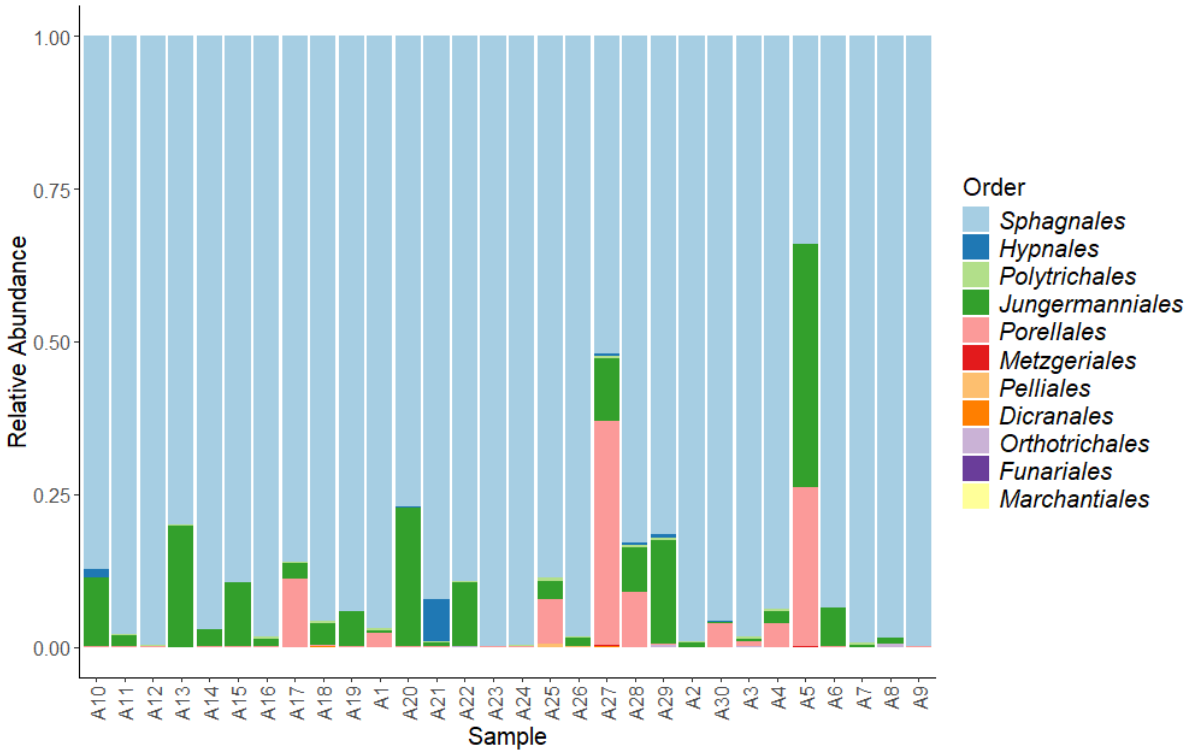



**Figure 5.41.** Proportional composition plot of bryophytes at the order level, plotted in R using the function microbiome. Data from the second bioinformatics method.

The vascular plant classes were varied between plots, with some samples dominated by the Lycopodiopsida, Magnoliopsida and Polypodiopsida (Fig. 5.42). Interestingly, despite noticeable amounts of litter from Pinus sylvestris, very low relative abundance of Pinopsida reads were detected.

At the order level, vascular plants communities continued to be varied (Fig. 5.43). Many of the orders were those which would be expected, such as the Ericales, Polypodiales and Poalles. However, some of the most abundant orders were more surprising, for example the Asparagales. Kraken assigned these reads to the Orchidacea, although the assigned species (*Angraecum leonis*) is not found in the UK.

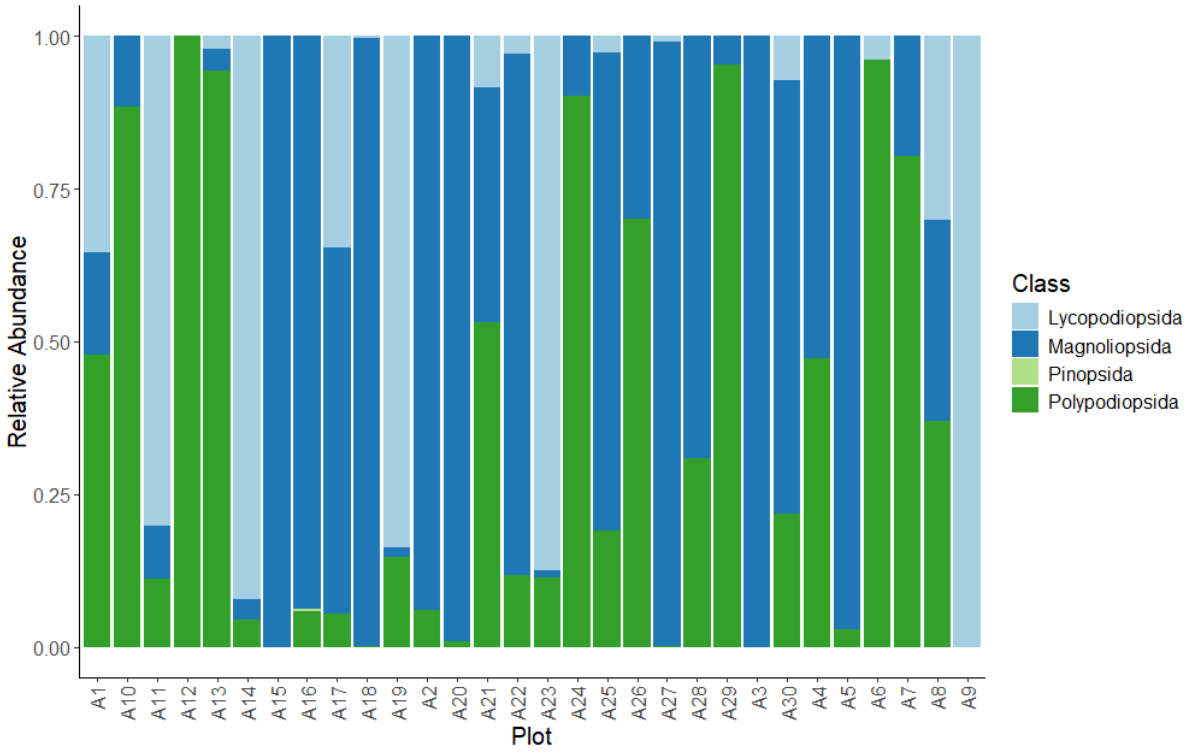



**Figure 5.42.** Proportional composition plot of vascular plants agglomerated at the class level. Data from the second bioinformatics method.

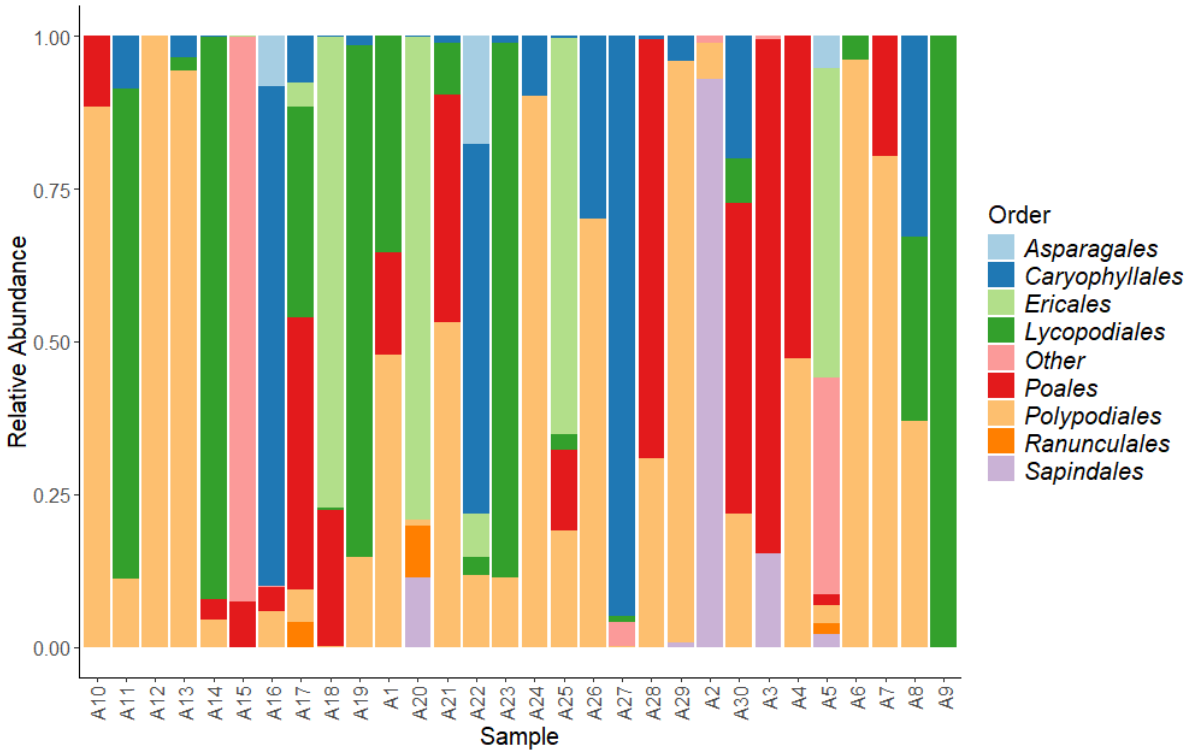



**Figure 5.43.** Proportional composition plot of vascular plants at the order level, plotted in R using the function microbiome. Rare taxa were aggregated at prevalence = 5/100, detection = 0.1/100. Data from the second bioinformatics method.

## 5.3.5 Correlations between variables

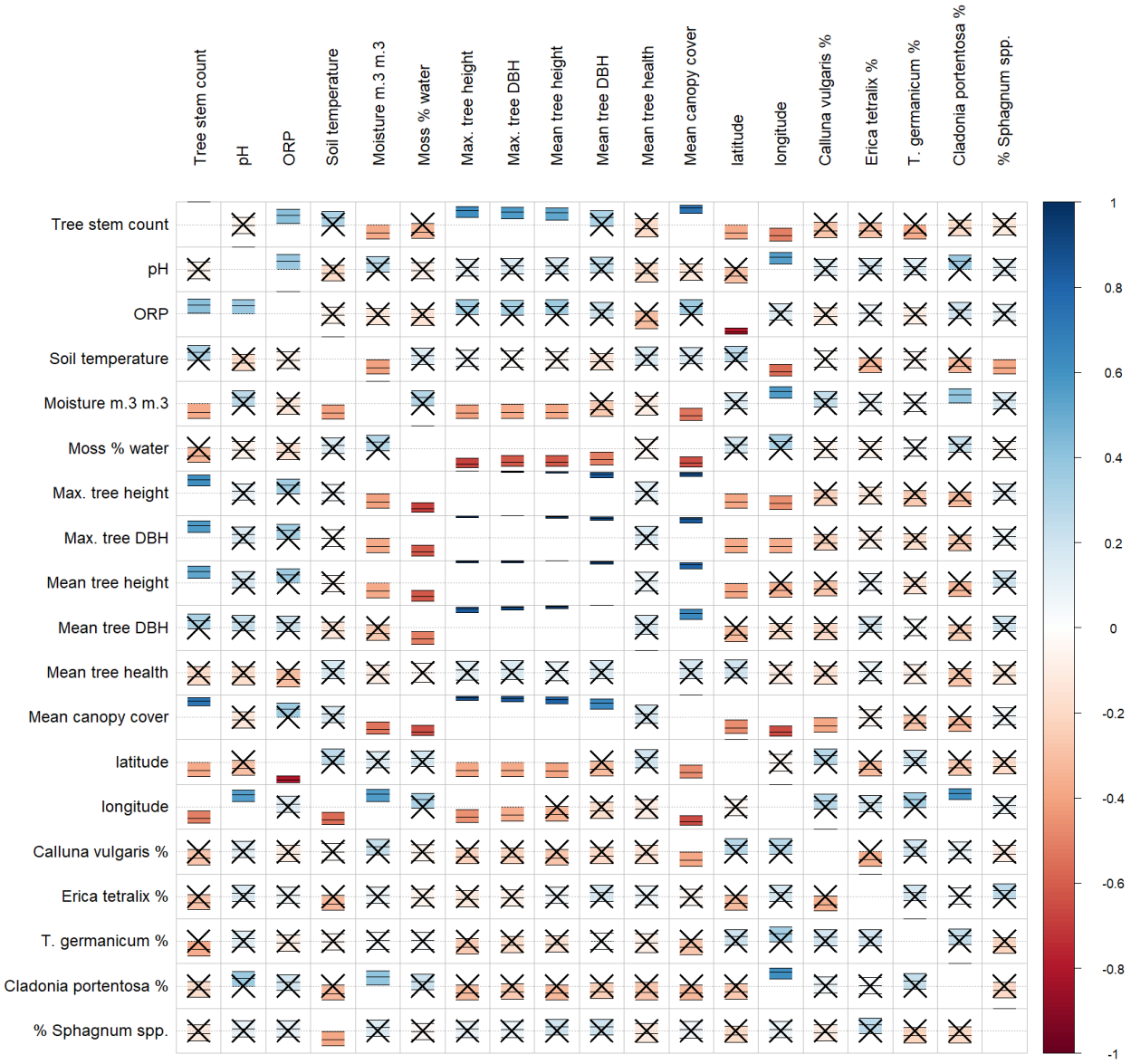



**Figure 5.44.** Correlation plot between all numeric variables without NA values. Blues show positive correlations, reds show negative correlation, with colour intensity and alignment within the box showing strength of relationship. Rectangles aligned closer to the top of the box have stronger positive correlations, and rectangles aligned closer to the bottom of the box have stronger negative correlations. The size of the rectangles shows the confidence intervals. Pairings with crosses do not have significant correlations.

There were few significant correlations between ground vegetation variables (Fig. 5.44). Percentage cover of *Sphagnum* was only significantly correlated with soil moisture, and percentage cover of *Cladonia* with soil moisture and latitude. *Erica tetralix* and *Tricophorum germanicum* were not significantly correlated with any other variables. *Calluna vulgaris* had a negative correlation with canopy cover.

Tree variables (height, DBH and canopy cover) unsurprisingly had strong positive correlations between them. Tree stem count was significantly correlated with all tree variables apart from mean DBH. The only exception to this was tree health, which was not significantly correlated with any other variables. Canopy cover had stronger and more positive correlations with maximum tree height and DBH than with other tree variables.

There were significant correlations between ORP and the number of tree stems within 2m, pH and ORP, and soil temperature and ground moisture (Fig. 5.44). Tree stem count, as expected, had a negative correlation with subsurface moisture. Interestingly, moss moisture (calculated in the lab after transport) and measured ground subsurface moisture were not significantly correlated. Moss moisture also had much stronger negative correlations with tree variables than soil moisture.

VIFs for non-tree variables were generally low, suggestive of collinearity being unlikely to cause problems (Table 5.2). Tree stem count also had a low VIF. Tree height and DBH (both mean and maximum) had extremely high VIFs, suggesting very strong collinearity. Canopy cover had a VIF above 10, which Borcard, Gillet, and Legendre (2018) suggest should be examined, and potentially avoided. However, this was much lower than the VIFs for the other tree measurement variables, and while obviously related to other tree variables (where tree height is below breast height, and there was no DBH measured, the measured canopy cover will also be extremely low due to lack of trees) it would be expected to be more indicative of the amount of shading experienced at the sampling point, rather than the impact of tree biomass on environmental conditions.

**Table 5.2**. Variance inflation factors (VIFs) for numerical variables for Chapter 5.

| Variable | VIF |
|---|---|
| Tree stem count | 3.376 |
| pH | 2.094 |
| ORP | 2.116 |
| Soil temperature | 1.746 |
| Moisture .m3.m3. | 2.057 |
| Moss percentage water | 2.940 |
| Maximum tree height | 92.088 |
| Maximum tree DBH | 54.555 |
| Mean tree height | 72.459 |
| Mean tree DBH | 35.930 |
| Mean percentage canopy cover | 10.448 |

There were no significant differences in moss percentage moisture for samples with different dominant *Sphagnum* species (Kruskal-Wallis: p = 0.092), and no obvious patterns between species (Fig. 5.45).

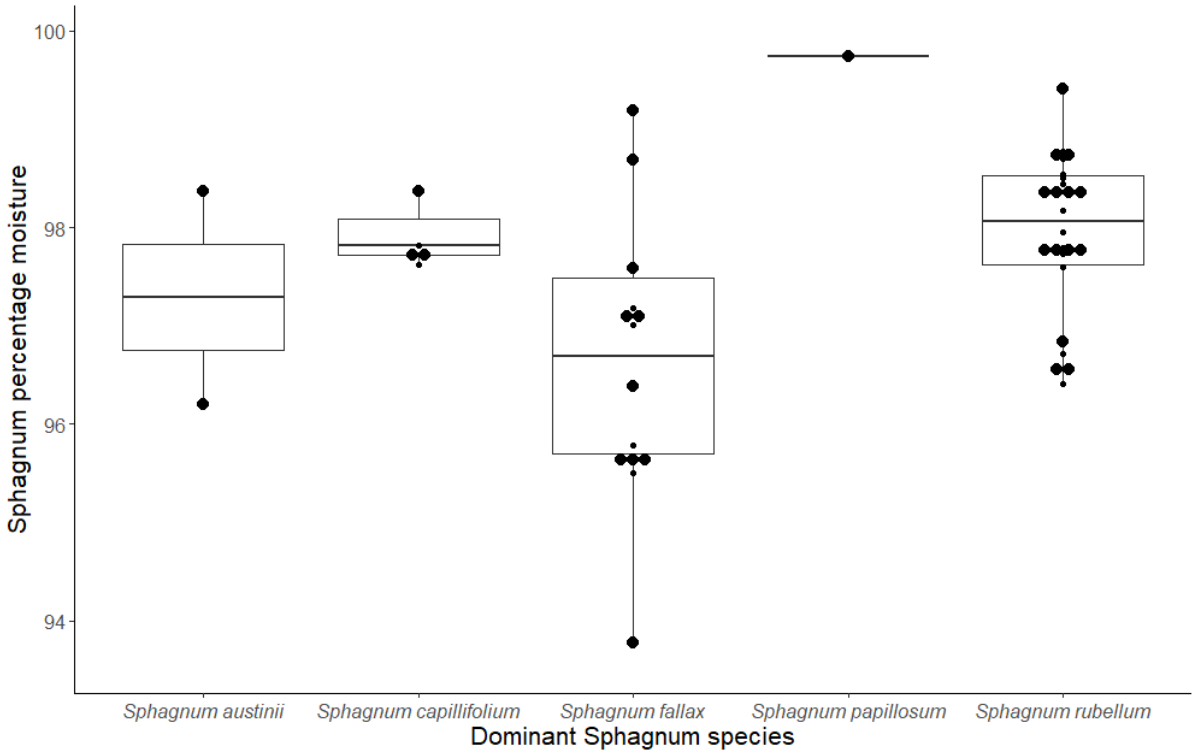



**Figure 5.45.** Sphagnum percentage moisture by weight for different dominant *Sphagnum* species. Box shows median and interquartile range. Points show actual data points.

### 5.3.6 Distance based RDA

Distance based RDA performance with data from the first bioinformatics method (model 1) was relatively poor, with selected variables explaining 27% of variance in algal community composition across sites. Of the variables included, only the number of tree stems within 2m of sampling was statistically significant (p = 0.022).

Distance based RDA from the second bioinformatics method and additional variables was more successful. Model 2 (without ground vegetation variables) explained 55.09% of variance in the algal community, although it was not statistically significant (p = 0.504). Moss percentage water was the only significant explanatory variable (p = 0.017).

When ground vegetation variables were added into the model (model 3, Fig. 5.46), constrained axis from the dbRDA explained 84.23% of variance in the algal community composition across sampling sites. The model had an overall p-value approaching

significance (p = 0.073). The number of tree stems was not significant, but pH (p = 0.042), moss percentage water (p = 0.006), percentage cover *Sphagnum* spp. (p = 0.037) and percentage cover of *Trichophorum germanicum* (p = 0.022) were all statistically significant.

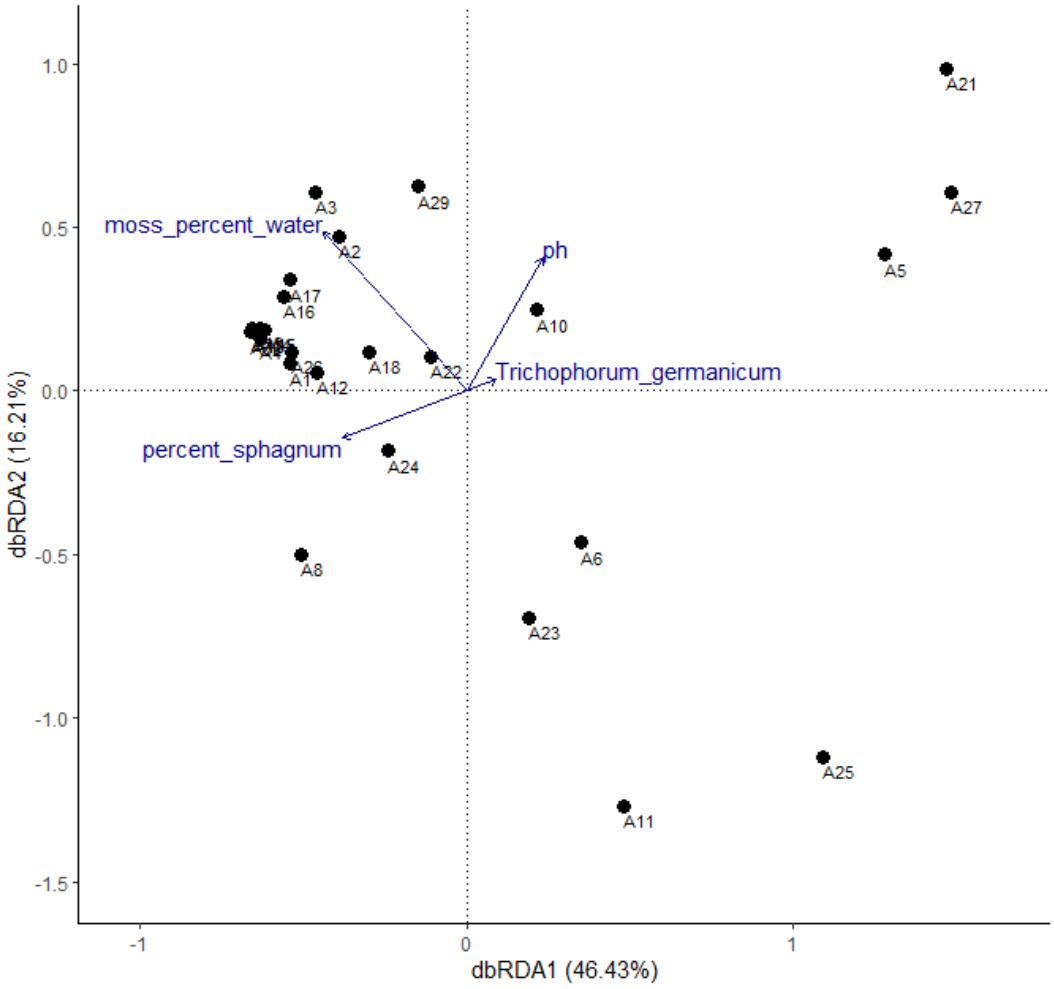



**Figure 5.46.** Biplot for dbRDA model 3, showing only statistically significant explanatory variables for the algal community ($p < 0.05$).

Model 4 (with mean tree height as the only tree height/DBH variable) explained 75.88% of variance but was not statistically significant (p=0.257). Only moss percentage water (p=0.019) and percentage cover *Sphagnum* spp. (p=0.033) were significant variables.

The simplified model produced by the ordiR2step function (model 5) explained 40.53% of algal community variance and was statistically significant (p = 0.001). All terms were

significant (moss percent water p = 0.007; percent *Sphagnum* cover p = 0.020; moisture /$m^3/m^3$ p = 0.013; ORP p = 0.034). Adding mean canopy percentage cover into this model (model 6, Fig. 5.47) increased the explained variance to 41.29%, and the model was still statistically significant, but mean canopy cover was not a statistically significant variable (p = 0.852).

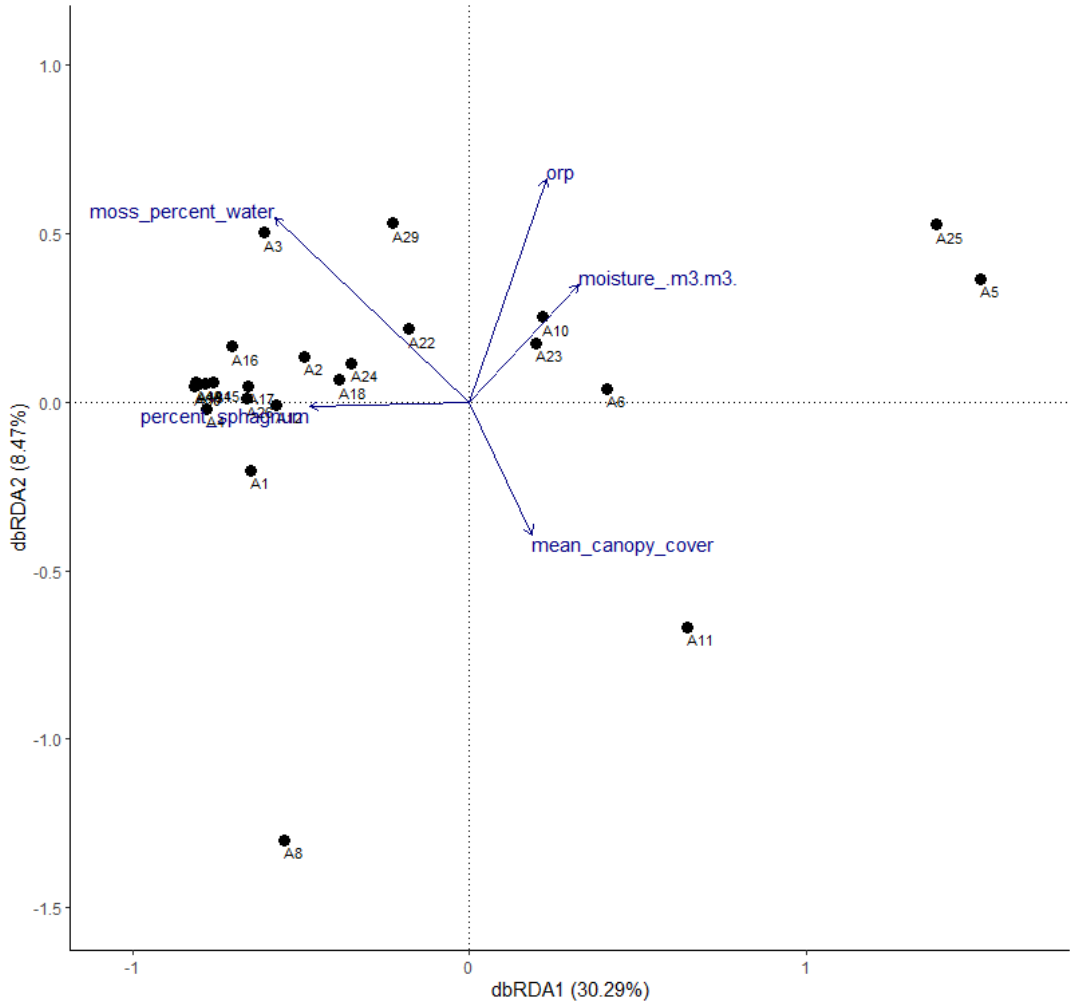



**Figure 5.47.** Biplot for dbRDA model 6. All explanatory variables plotted except mean canopy cover are statistically significant (p < 0.05).

### 5.3.7 Mantel tests

There was no statistically significant correlation between the distance matrices between the plant community observed in the field and the plant community observed through molecular data (Mantel r = 0.013, p = 0.417).

There was no significant correlation between algal dissimilarity and geospatial distance (Mantel r = 0.108, p = 0.114).

There was no statistically significant correlation between algae community distances and observed plant community distances (Mantel r = -0.05, p = 0.702). However, there was a statistically significant but weak positive correlation between algal community distances and molecular plant community distances (Mantel r = 0.332, p = 0.018).

## 5.4 Discussion

### 5.4.1 Changes in the algal community

Large amounts of variability were observed in the algal community even at the phylum level, suggesting that, while some taxonomic information is lost by agglomeration of molecular data and the subsequent exclusion of species and genera, there is still enough heterogeneity within the community to investigate explanatory variables.

Most samples were dominated by the Desmidiales, which is unsurprising as these are extremely abundant in other ombrotrophic peatlands in Scotland (Goodyer, 2014). In the few samples which were dominated by Chlorophyta, Trebouxiophycea were the dominant class, with most taxa belonging to either the Chlorelales or Prasiolales. Some samples were also dominated by Ulvophyceae in the Bryopsidales. These are a mostly marine group, with only the genus Dichotomosiphon found in freshwater environments (Cremen *et al.*, 2019). While this genus is often found in lakes, it has also been recorded in wet terrestrial environments (Guiry, 2018), and so its presence in this system is plausible. Chlorellales are well known in peatlands as the symbionts to the functionally significant mixotrophic testate amoeba (Gomaa *et al.*, 2014), and were a relatively abundant Trebouxiophycea order. However, there are many other species in this order

(Algaebase lists 525 species in the Chlorellales- Guiry, 2018), and so samples having a high relative abundance of Chlorellales does not necessarily indicate a high relative abundance of mixotrophic testate amoebae.

Samples A23 and A24 are notable as the two samples with microbial communities dominated by prokaryotes. Both have a high proportion of overall reads assigned to *Sphagnum* (98.98% and 92.56% respectively), and for both the dominant bacteria is a member of the Caulobacterales, in the Pseudomonadota. When looking into the taxonomy assigned by Kraken2, all Caulobacterales reads across both samples are assigned to *Phenylobacterium* sp. LH3H17, which was isolated from glacial till in Tibet (NCBI: GCF_024298925.) Interestingly, the genus *Phenylobacterium* contains species which specialise in using phenol groups from heterocyclic compounds, which are abundant in peatlands, and a sister taxon had high acid phosphatase activity, which would be useful for degrading the complex organic compounds in an acidic peatland environment (Baek, Shin and Yi, 2019). Low relative abundance of cyanobacteria is likely to be due to poor primer complementarity (see Section 3.6) rather than their absence from the system, as they are generally abundant in peatlands, albeit at lower densities than eukaryotic algae (Gilbert *et al*., 1998a).

#### *5.4.1.1 Differences between the bioinformatics methods*

Despite only explaining a small amount of the total variance, the distance-based RDA for the first bioinformatics method indicated that the number of trees in close proximity has a significant effect on the total algal community, but other environmental variables correlated and theoretically linked to tree presence (soil moisture and ORP) do not. This suggested that algal community changes are associated with a different property of trees, which was hypothesised to be shading (with percentage canopy cover a proxy of this). Increasing tree canopy cover was expected to change the algal community composition, as those algal species which were able to survive in lower light conditions would have a competitive advantage where shading was more prevalent.

However, for the second bioinformatics method, where a much higher proportion of the reads were successfully classified, no tree variables were significant in any of the

dbRDAs. This difference may be because the majority of reads (and therefore likely a high proportion of the reads unidentified by the first method) were allocated to the host *Sphagnum*. If these were previously unidentified, they would not have been able to be filtered out as plants, and so would have been included in the initial algae dataset. While there was no significant correlation between *Sphagnum* percentage cover and tree presence, variables such as soil moisture which would be expected to interact with moss abundance were significantly correlated with tree size, and this may still have been enough to skew the results of the first dbRDA. This may also explain the difference in significance of transect effects between the two bioinformatics methods.

There were also some differences in the algal communities identified by the two bioinformatics strategies. Notably, the Euglenozoa / Euglenaphyceae were missing from the second bioinformatics method. This suggests a loss of sensitivity, but this is outweighed by increased reliability from exclusion of (likely previously unidentified) *Sphagnum* reads. When reference databases are not optimised for the primers, the decision to sacrifice some sensitivity for greater reliability was judged to be the best compromise. However, this demonstrates the need to be aware of the differences in results possible from changes to bioinformatic methods when working with molecular data. In an ideal world, laboratory and field method development to prevent as much host plant DNA entering the PCR may be the best way to avoid this compromise. Regardless, we must accept when using molecular methods to assess biodiversity, especially for organisms not well represented in reference databases, that an absence from molecular data does not necessarily mean the organism is truly absent, and that this may differ even between different bioinformatics pipelines of the same sequence data.

#### *5.4.1.2 dbRDA and Mantel test findings- what drives algal community dissimilarity?*

Across the three successful dbRDA models (models 3 and 5/6), moss percentage moisture and percentage *Sphagnum* cover were consistently identified as statistically significant explanatory variables. This is unsurprising, as *Sphagnum* mosses represent a relatively specific niche for algae, and moss moisture levels will determine whether

algal species which require constant moisture persist, or if the community is made up of species which are more resilient to drought stress.

The percentage of *Trichophorum germanicum* and ORP/pH were also identified as significant variables in the dbRDA models. While there were no significant correlations between *Trichophorum* percentage cover and any other variables in this study (Fig. 5.44), on the biplot (Fig. 5.46) it is plotted as approximately opposite to *Sphagnum* percentage cover, suggesting it is indicative of a vegetation community where *Sphagnum* is less dominant.

Surface vegetation community, and especially the *Sphagnum* species which microbes were extracted from, was also expected to be an important driver of the algal community, as many previous studies have shown microbial community specificity to different *Sphagnum* species (e.g. Hamard *et al.*, 2019; Bragina *et al.*, 2012). However, the dominant *Sphagnum* species was not a significant variable in determining variance of the algal community in any dbRDAs. Dominant *Sphagnum* species was also not a statistically significant predictor of moss percentage moisture, despite the samples for laboratory analysis being taken from the dominant moss species (or a representative sample of the moss community within the quadrat if one species was not clearly dominant), and different species of *Sphagnum* holding different amounts of water both intracellularly and extracellularly (van de Koot *et al.*, 2024, Bengtsson *et al.*, 2020). The reason for this is unclear but may indicate that relationships between bryophyte species and both the hydrological condition and their microbiomes are more complex when there are multiple *Sphagnum* species in close proximity than in discrete single species colonies. It could also indicate that the *Sphagnum* algal microbiome has no clear patterns between different moss species at the family level, and differences in algal communities between different *Sphagnum* species instead occur at the species or genus level and therefore were not detected with this methodology. The significance of *Sphagnum* cover (and *T. germanicum*) to the algal community shows that the vegetation substrate is important for the algae, even if not at the level expected.

While there were limited species-level vegetation variables identified as significantly explaining variance in the algal community by the dbRDAs, Mantel tests did show a

significant correlation between the algal community and the wider vegetation community. However, this was only the case for the vegetation community as detected in the molecular data, rather than that observed and identified in the field. This seems counterintuitive, as the molecular data, being agglomerated at family level, did not contain the species level classifications of *Sphagnum* species, which were hypothesised to be an important determinate of microbial community. The difference between the two vegetation community methodologies may have been a matter of scale; the field observed vegetation included the whole quadrat, whereas molecular data was gathered from the same sample (and same DNA extraction) as the algae, over a smaller area than the quadrat. Therefore, the molecular data could be argued to be more representative of the exact microhabitat experienced by the algae. However, DNA was detected for plant species which were not visible in the (predominantly *Sphagnum*) sample, suggesting DNA inputs not only from the bryophyte substrate. The Mantel test between algal community and molecular vegetation community should also have slightly greater statistical power, as one sample had to be excluded from the observed vegetation dataset (and corresponding algal dataset) due to missing data. This, however, seems unlikely to be responsible for such a large difference in p-value. Regardless of the reason, this demonstrates the importance of clarity when discussing measurements of the vegetation community, and merits further investigation. This second example of *Sphagnum* species not being significant may also suggest that the algal community is less strongly associated with specific *Sphagnum* species than the wider moss surface microbial community.

It is notable that there was also no significant correlation between the observed and molecular vegetation communities. This may be because the Mantel test is a test of correlation between dissimilarity matrices rather than the raw community data. The dissimilarly distances between the communities may be different because *Sphagnum* dominate the molecular samples more than the observed communities. The different taxonomic levels involved may also play a role; much of the difference between different observed vegetation communities was between different *Sphagnum* species, but these would have been agglomerated (order Sphagnales) in the molecular data.

Canopy cover and tree presence likely do explain some algal variance, but only a small proportion; the increase in variance explained between model 5 and 6, where canopy cover was the only addition, was 0.76%. Tree related variables were also repeatedly found not to be statistically significant by dbRDA. Trees therefore have less of a direct impact on variation within the algal community than the other key variables identified (moss moisture, *Sphagnum* percentage etc.). Tree cover can affect vegetation by changes in moisture levels and shading (Kettridge *et al.*, 2013). This also raises the possibility that any impact of trees on algae is itself mediated via shifts in the vegetation community, especially of the bryosphere which forms the immediate habitat of surface-dwelling algae. In peatland restoration following forestry, recovery of vegetation (specifically *Sphagnum*) following disturbance is thought to control the recovery of testate amoeba communities (Creevy *et al.*, 2023). It does not seem unreasonable for the response of algae more generally to tree cover to also be mediated by the bryophyte community. Mixotrophic testate amoebae have been identified as possible bioindicators in forest-bog restoration systems (Creevy *et al.*, 2023), and, like algae, they are sensitive to both light and moisture levels. Whilst the forest-bog ecotone at Abernethy is caused by successional processes rather than restoration processes, similar results could have been expected from this system. The limited response of the algal community to tree variables may suggest that the effect of trees in this natural ecotone is much less extreme than for forest-bog restoration systems. This may be because of lower density of trees than would be found in plantation forestry, dominance of *Pinus sylvestris* rather than standard commercial forestry species, or the lack of anthropogenic drainage in addition to the hydrological impact of the trees.

The effects of tree proximity may not be generalisable to all tree species, or even all sites. Varying canopy density of different species (and different growth forms/ health statuses within species) would be expected to have a direct impact on the degree of shading caused. *Pinus sylvestris* growth form can be variable, with changes in height and diameter of trees associated with different climatic conditions (Mathisen and Hofgaard, 2011).

It was hypothesised that, where climatic (or hydrological) conditions meant that trees were smaller or less healthy, their impact on microbes from either shading or drying soil would be altered, with weakened trees having less impact on the algal community. However, no tree variables were significant with the updated data. Tree health had no significant correlations with any other variables, although this may be because of the relatively small proportion of trees included in this study which had lower health scores. It is possible that if more dead trees (which would be expected to have less hydrological impact on the peatland) had been included, more of a signal would have been possible to detect. However, as tree health had to be excluded from dbRDA models due to NAs for plots without trees, it is not possible to conclusively rule out effects on the algal community. An approach which excluded treeless sampling points and had a more even spread of different tree health statuses would likely be needed to investigate this with any sort of statistical power.

Seasonality should also be considered when interpreting the results of this work. Sample collection took place over a three-day period, providing a snapshot at the end of the growing season. Shading conditions in this *Pinus sylvestris* woodland are likely to show less seasonal variation than deciduous tree cover would. However, seasonal differences in hydrological and nutrient impacts from trees would still be expected, as *P. sylvestris* does have a distinct growing season (Mathisen and Hofgaard, 2011). Therefore, effects of trees may be detectable at different time periods, for example earlier in the growing season. If conditions had been drier for a period before sampling this could also increase the hydrological impact of trees if bryophytes and algae were already drought-stressed, possibly resulting in measurable differences. Adding repeated seasonal measures, linked to climatic conditions, may allow this to be unpicked in future work.

While there was no statistically significant effect of tree variables on the algal community, it is notable that samples A23 and A24, the only two samples where the microbial community was dominated by prokaryotes, were also two of the three plots with the highest numbers of tree stems within 2m of the sampling point, and had the two highest mean percentage canopy covers. This is suggestive of changes in biomass

of the algal (and potentially non-algal) surface microbial community. Methods which allow for biomass to be measured, even by proxy, may therefore be a valuable next step in understanding the surface microbial dynamics of forest – bog ecotones.

#### *5.4.1.3 Transects*

During data analysis, tests were carried out to investigate differences in the algal community between transects. Transects were designed as a way of grouping samples for ease of field sampling, to take samples across the ecotone. Therefore, differences in the algal communities between transects is not in itself ecologically informative, other than in demonstrating the heterogeneity of the system, as the transects have not been designed to differ in any controlled way. However, it does demonstrate the differences between biodiversity indices and PERMANOVA methods which use ASV abundance tables for the whole (algal) community. This suggests that detail is lost in using the simpler biodiversity indices, supporting the use of more complex statistical methods.

#### *5.4.1.4 Ecological indicators*

Initially it was hoped that, if a strong shading effect was detected, there would be potential for microbes to act as an indicator of shading conditions averaged out over the turnover timeframe of the community (or of those key species). This would provide useful data, allowing a single 'snapshot' field sampling session to collect data on light conditions over the medium term which would otherwise require placement (and subsequent collection) of long-term monitoring equipment. However, as no significant shading effect, or indeed effect of canopy cover at all, was detected, there is less potential to use algae as an indicator species potential. Algal heterogeneity even at a small scale remains very high, and it is generally easier to carry out surveys at the tree or vegetation scale. However, algae remain highly diverse in peatlands, and their value as a source of cryptic biodiversity remains important, even if their indicator potential for this type of system is low.

#### *5.4.1.5 Spatial autocorrelation*

The lack of spatial autocorrelation in the algal community is surprising, as this is regularly observed in soil communities (e.g. Malard *et al*., 2021; Bakker *et al.,* 2014; Franklin and Mills, 2003), and also for surface terrestrial algae (Darcy *et al*., 2017). This suggests that variables which do not have clear patterns in spatial heterogeneity over the study site (e.g. *Sphagnum* moss presence, as this was found at all sites) are directly driving the algal community, rather than variables which would be expected to have more spatial structure such as hydrology. The alternative is that spatial structuring of the algal community is occurring, but only at the genus/ species level, which is lost due to the lower sensitivity of the second bioinformatics method. However, the high levels of heterogeneity in algal community composition even at higher taxonomic levels suggests that this is not the case.

### 5.4.2 Vegetation

Canopy cover calculations were an attempt to uncouple shading from other tree presence variables, but, as no significant effects of any tree variables were observed with the improved bioinformatics method, this was of limited success. Canopy cover inherently shows tree presence (and therefore is a proxy for other impacts of trees such as hydrological changes), because the mean canopy cover if no trees are present is 0. An interventionist approach, such as artificial shading, rather than observational studies would likely be needed to properly uncouple the effects of shading alone vs. other tree impacts on the algal community. However, these results suggest that the impact of the amount of shading seen on this site (with the highest canopy cover of 61.7%) on the algal community is low, and that once again, where *Sphagnum* moss is able to grow, the algal community will be relatively constant. In areas where the canopy is denser, and therefore shading is more intense, a larger effect would be expected, and this may be detected in the algal community. However, as most *Sphagnum* species are sensitive to heavy shading (Norby *et al*., 2023), the algal community may in those circumstances be affected more by absence of *Sphagnum* substrate than by the shading itself.

Percentage *Sphagnum* cover had no significant correlations with any other measured variables except for soil temperature. This is likely to be because *Sphagnum* acts as a thermal insulator (Soudzilovskaia, van Bodegom and Cornelissen, 2013) meaning that soil temperature is higher where there is less *Sphagnum*. However, the lack of correlation between *Sphagnum* cover and other variables is surprising, as both hydrological and shading effects of trees would be expected to reduce *Sphagnum* cover (Norby *et al.*, 2023; Moore and Waddington, 2015; Schipperges and Rydin, 1998).

Moss percent moisture was not inherently driven by the soil moisture status, as the correlation between them was not statistically significant. However, moss percentage moisture and soil moisture were significantly negatively correlated with tree height, DBH and canopy cover variables (while not correlated themselves). This suggests that the changes in moss moisture state (which do significantly drive changes in algal community composition) may be influenced by tree variables. Indeed, tree variables had stronger negative correlations with moss moisture than with subsurface moisture, and moss moisture was consistently returned by dbRDAs as the more significant variable for explaining the algal community. This suggests that, while mosses are key and the actual direct effect of trees on algae seems limited, the indirect effects of the trees on the moss may still influence the algal community. Reflectance has been used to detect drought stress in *Sphagnum* for many years (Harris, Bryant and Baird, 2005) and could be used to detect water stress from *Sphagnum* near tree cover, rather than simply relying on percentage moisture by weight as an indirect measure, which might allow the relationship between trees and moss moisture (and therefore algal communities) to be understood in greater detail.

Moss moisture is not a proxy for *Sphagnum* species, although the statistical power of this test may have been reduced by the uneven distribution of different *Sphagnum* species. This is surprising, as different *Sphagnum* species have different levels of resistance to desiccation (Nibau *et al.*, 2022). This has functional implications; desiccation in *Sphagnum* is linked to reductions in photosynthetic capacity, and recovery from this is also variable between *Sphagnum* species (Nibau *et al.,* 2022). Therefore, if the dominant *Sphagnum* species does indirectly drive algal variation in a

way detectable by the level of taxonomic detail in this study, it is not via moisture. The complex chemistry of *Sphagnum* may allow species specific controls to be mediated by pH or ORP, but it seems unlikely that this would be a major explanatory variable when *Sphagnum* species (and vegetation community which included species level identification of *Sphagnum*) had no statistically significant interactions with the algal community.

Soil moisture was relatively consistent between samples, varying between 0.71 and 0.8 $m^3/m^3$. In contrast, moss percentage moisture by weight was more varied, between 93.8% and 99.7%. However, scaling with z-scores should have prevented the difference in variation from preventing correlations being identified. This, therefore, suggests that moss moisture and soil moisture are more uncoupled than would be expected. Moss percentage moisture also had a stronger negative correlation with tree variables than soil moisture, suggesting that trees can affect the bryosphere moisture conditions (which are significant drivers of the algal community) more than the soil moisture conditions. This again highlights the importance of studying the bryosphere conditions when looking at surface microbial communities, rather than extrapolating from more standard peatland 'below-ground' measurements.

Whilst tree cover did not have a strong effect on algal community, it did impact climatic data. Even relatively sparse tree cover in the forest had an insulating effect on surface temperatures (Fig. 5.24). The forested area also had consistently lower subsurface soil moisture levels than the open bog (Fig. 5.23). Moisture and temperature both showed strong seasonality, suggesting that the level that these are stressors on the vegetation varies through the year.

## 5.5 Conclusions

The dbRDA models were successful in explaining a good proportion of variation, especially when ground vegetation variables were added to the model, and it was optimised to reduce autocorrelation between tree variables (model 3). The most significant variables on the algal community were moss percentage water and percentage *Sphagnum* cover. ORP, pH, *Trichophorum germanicum* cover and soil

moisture were statistically significant in some dbRDA models but not all. Tree variables explained some variance in the algal community, but this was limited and they were consistently not statistically significant. This suggests that the moss mediated microenvironment is the most important factor for the peatland microalgae community, and contrary to other studies, the effect of tree cover on the algal community is limited. If trees do impact the algal community, it is likely to be mediated via the vegetation community, and particularly *Sphagnum* mosses.

As algae, which are photosynthetic organisms by definition, contribute to carbon fixation in peatlands where they are present, any changes in algae also have potential implications for carbon budgets of forested peatlands. In open bogs, microalgae contribute measurably to total carbon fixation (Jassey *et al*., 2022). Whilst this study did not seek to quantify microbial carbon fixation, understanding changes in algal community patterns across a forest-bog ecotone may allow future work which incorporates their contribution into treed peatland carbon calculations.

These results suggest that tree cover in mosaic, relatively open systems is unlikely to detrimentally impact the biodiversity of surface algae. However, where tree cover is denser, or trees have greater hydrological impacts, this may be more of a risk. In general, *Sphagnum* cover and moisture appear to be the most important variables for the algal community. Conservation work looking to minimise loss of this cryptic biodiversity should therefore focus on ensuring that tree conditions allow *Sphagnum* to persist without desiccation; if this is the case then the algal communities which depend on it should also be maintained.

# 6 Conclusions

## 6.1 Methodology

Both molecular and microscope methods can have issues assigning taxonomy to species level, and, as seen in this thesis, sometimes only identification at higher taxonomic levels will be possible with any degree of accuracy. In molecular work, when reference databases are limited, taxonomic assignment of sequences can have low success rates, and species level identifications may have reduced accuracy. This in turn can limit our ability to detect if changes in community metrics are indeed for target taxa. Microscopy likewise cannot identify small single celled organisms to species level, and even some of the more distinctive taxa can need electron microscopy for confident species level identification. Therefore, with these methods, there remains cryptic diversity within the cryptic diversity of microalgae.

This thesis demonstrates that a DADA2 pipeline can be adapted to use Kraken2 and Bracken for taxonomic assignment. The tradeoff of reduced accuracy at lower taxonomic levels, and therefore the loss of species-level information when taking a cautious approach, for a higher identification rate of amplicons which do not have an extensive and well-curated database was judged to be worthwhile. Kraken2 and Bracken allow the use of extremely large NCBI databases in a way which, with access to research computing, is relatively less computationally intensive for large datasets. This allows bioinformatics of taxa and amplicons without a dedicated reference database to be carried out more quickly and easily. This principle could be used for other genes and gene regions which do not have established databases, allowing functional genes to be studied as long as they have the level of conservation required for good taxonomic discrimination.

## 6.2 Algal community responses to stressors

The results of Chapters 2, 4 and 5 show that algae are surprisingly resistant to the stressors investigated in this thesis. Neither tree presence and cover nor any nutrient addition treatments corresponded with statistically significant changes in the algal

community. While trees do explain a small amount of the variance in the algal community (Chapter 5), the majority of their impact appears to be indirect. As long as the trees are not affecting the variables which do impact the algal community, such as *Sphagnum* cover and moisture, they are unlikely to have a large impact on the photosynthetic microbes in the kind of near-natural ecotonal system investigated here. The algal community appears to be similarly resistant to long-term nutrient pollution as long as the *Sphagnum* which they were sampled from persisted. Perhaps most surprisingly, the algal community also appeared to be resistant to the impact of wildfire.

We cannot say that the algal community is obligate on *Sphagnum*; it would require development of new methodologies to standardize sampling from *Sphagnum* and non-*Sphagnum* substrates to determine this. However, it does seem ubiquitous- if the *Sphagnum* is able to survive then there are few consistent changes in the algal community with any of the other stressors included in this study.

## 6.3 Management

One of the initial aims for this thesis was to investigate the potential of algae to be useful bioindicators of peatland condition and perhaps recovery. However, a community fingerprinting approach has little value as an ecological indicator when few consistent shifts are detected in the algal community with the environmental stressors investigated in this thesis. Ecological indicator value instead should focus on identifying key taxa which can both be consistently identified and show consistent changes. However, if, as in this study, the algal communities appear to be more impacted by the bryosphere than any of the stressors, it may be better (and perhaps less expensive and time consuming) to monitor vegetation, with a particular focus on mosses.

As the algal community is more driven by the bryophyte community than any of the investigated stressors, it suggests that specific management for this cryptic biodiversity will not be necessary. Instead, management and restoration effort should focus on ensuring the persistence of *Sphagnum* and avoiding desiccation stress. To increase biodiversity of algae, management should focus on diversity of *Sphagnum* species and ecological niches, and in particular ensuring microtopographic variation is protected

and considered when restoring damaged sites. Alongside increasing algal diversity, this may also have a protective action for the peatland surface vegetation (and the photosynthetic algae which dwell on it) from low and medium severity wildfire, and therefore one of the predicted pressures of climate change.

# 7 Supplementary data

**Supplementary Table 3.1.** Prokaryote taxa found by *in silico* PCR

| Sequence | Group and notes |
|---|---|
| *Streptomyces ruber* strain JCM 3131 sequence069, whole genome shotgun sequence | Actinomycetes<br>(Heterotrophic) |
| *Streptomyces lomondensis* strain JCM 4866 sequence39, whole genome shotgun sequence | Actinomycetes |
| *Streptomyces xanthochromogenes* strain JCM 4594 sequence49, whole genome shotgun sequence | Actinomycetes |
| *Streptomyces termitum* strain JCM 4518 sequence35, whole genome shotgun sequence | Actinomycetes |
| *Streptomyces omiyaensis* strain JCM 4806 sequence077, whole genome shotgun sequence | Actinomycetes |
| *Streptomyces olivaceoviridis* strain JCM 4499 sequence100, whole genome shotgun sequence | Actinomycetes |
| *Streptomyces violarus* strain JCM 4534 sequence22, whole genome shotgun sequence | Actinomycetes |
| *Actinokineospora fastidiosa* strain JCM 3276 sequence17, whole genome shotgun sequence | Actinomycetes<br>Aerobic, heterotrophic (Labeda *et al*., 2010) |
| *Halomonas heilongjiangensis* strain 9-2 1252, whole genome shotgun sequence | Pseudomonadota<br>(This phyla -Pseudomonadota- contains purple bacteria - anoxic phototrophs, widespread especially in aquatic environments. Halomonads are heterotrophic.) |

**Supplementary Table 5.1.** Thresholding values used during canopy cover calculation for images with visible trees

| Image | Automatic thresholding level | Adjusted thresholding level |
|---|---|---|
| 5_e | 86 | - |
| 5_n | 82 | - |
| 7_e | 116 | - |
| 7_n | 116 | - |
| 8_e | 112 | - |
| 8_n | 99 | - |
| 9_e | 87 | - |
| 9_n | 80 | - |
| 10_e | 91 | - |
| 10_n | 99 | - |
| 11_e | 113 | - |
| 11_n | 127 | - |
| 12_e | 113 | - |
| 12_n | 112 | - |
| 13_e | 89 | - |
| 13_n | 91 | - |
| 14_e | 97 | - |
| 14_n | 91 | - |
| 15_e | 86 | - |
| 15_n | 106 | 75 |
| 16_e | 71 | - |

| | | |
|---|---|---|
| 16_n | 82 | 70 |
| 17_e | 99 | - |
| 17_n | 103 | - |
| 18_e | 99 | - |
| 18_n | 104 | - |
| 19_e | 87 | - |
| 19_n | 91 | - |
| 20_e | 104 | - |
| 20_n | 96 | - |
| 21_e | 97 | - |
| 21_n | 99 | - |
| 22_e | 101 | - |
| 22_n | 97 | - |
| 23_e | 107 | - |
| 23_n | 122 | - |
| 24_e | 122 | - |
| 24_n | 123 | - |
| 25_e | 60 | 86 |
| 25_n | 51 | 128 |
| 26_e | 82 | 133 |
| 26_n | 79 | 110 |
| 27_e | 74 | 145 |
| 27_n | 70 | 136 |
| 28_e | 84 | - |

| | | |
|---|---|---|
| 28_n | 164 | 90 |
| 29_e | 102 | - |
| 29_n | 98 | - |
| 30_e | 106 | - |
| 30_n | 102 | - |

**Supplementary Table 5.2:** VIFs calculated for dbRDA models including various tree variables.

| dbRDA number | dbRDA1 – all variables | dbRDA2 - All variables less max tree height | dbRDA5 - all variables less mean tree height | dbRDA3 - all variables less max tree height & mean tree DBH | dbRDA4 - all variables less max tree height & max tree dbh | dbRDA6 - all variables less mean tree height & mean tree dbh | dbRDA7 - all variables less mean tree DBH & max tree height | dbRDA8 - dbrda7 less tree health | dbRDA9 - DBRDA3 less tree health |
|---|---|---|---|---|---|---|---|---|---|
| tree_stem_count | 12.11023 | 11.63435 | 11.3997 | 10.35155 | 11.63048 | 11.38538 | 10.67771 | 4.364147 | 3.62758 |
| ph | 72.90351 | 65.52622 | 56.71555 | 23.4617 | 57.46289 | 46.17708 | 28.3589 | 6.185794 | 6.362476 |
| orp | 15.0477 | 14.46724 | 10.45802 | 7.58731 | 13.51992 | 7.261839 | 10.17728 | 7.973227 | 5.965408 |
| soil_temperature | 27.51216 | 25.79837 | 23.83168 | 14.55447 | 24.57611 | 19.54081 | 18.02092 | 5.618568 | 5.67617 |
| moisture_.m3.m3. | 16.19117 | 14.6472 | 15.47617 | 12.58948 | 9.769651 | 15.39064 | 8.567184 | 3.04804 | 3.095074 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| moss_percent_water | 17.75267 | 13.24453 | 16.52124 | 5.862164 | 13.09446 | 13.62277 | 10.46653 | 3.960941 | 3.820865 |
| max_tree_height | 135.25727 | 42.46404 | 100.1917 | | | 94.77462 | 30.13838 | 17.74195 | |
| max_tree_dbh | 145.81332 | | 131.8565 | 42.26704 | | 128.9645 | | | 19.67192 |
| mean_tree_height | 111.08931 | 82.28932 | | 18.49708 | 76.86201 | | | | 11.95844 |
| mean_tree_dbh | 111.78434 | 97.08043 | 23.06913 | | 96.63004 | | 22.56316 | 8.632716 | |
| min_tree_health2 | 35.61842 | 30.26393 | 33.9138 | 26.56163 | 16.94775 | 33.70169 | 13.26358 | | |
| min_tree_health3 | 65.0499 | 59.39775 | 48.23752 | 25.51838 | 49.56313 | 44.23134 | 21.00513 | | |
| mean_canopy_cover | 85.47071 | 83.46084 | 80.13506 | 75.80533 | 47.79496 | 74.9458 | 64.98173 | 18.98751 | 24.37056 |
| dominant_sphag_species *Sphagnum_austinii* | | | | | | | | 3.561462 | 3.57827 |
| dominant_sphag_species *Sphagnum_capillifolium* | 12.03201 | 11.52252 | 11.93969 | 11.50804 | 10.61786 | 11.89093 | 10.7887 | 4.696444 | 4.50636 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| dominant_sphag_species *Sphagnum_fallax* | 43.8367 | 37.83484 | 43.74146 | 37.76585 | 28.17639 | 43.08232 | 29.05585 | 11.18843 | 10.93892 |
| dominant_sphag_species *Sphagnum_papillosum* | 35.759 | 35.74689 | 28.9117 | 19.16679 | 32.51358 | 17.95962 | 28.84067 | 3.876643 | 3.670004 |
| dominant_sphag_species *Sphagnum_rubellum* | 59.97921 | 56.08882 | 59.96267 | 56.08247 | 39.13156 | 58.61919 | 43.63082 | 17.08171 | 16.55516 |
| latitude | 46.41057 | 41.37104 | 44.72241 | 39.01776 | 20.48621 | 39.00444 | 27.43198 | 14.7028 | 14.2714 |
| longitude | 58.23716 | 57.96546 | 43.54914 | 28.34627 | 57.81765 | 33.09903 | 41.32851 | 9.415186 | 9.360547 |

# 8 References

A'Bear, A. D. *et al.* (2014). 'Interactive effects of temperature and soil moisture on fungal-mediated wood decomposition and extracellular enzyme activity', *Soil Biology and Biochemistry*, 70, pp. 151–158. Available at: https://doi.org/10.1016/j.soilbio.2013.12.017

Adl, S. M. (2007). 'Motility and migration rate of protozoa in soil columns', *Soil Biology and Biochemistry*, 39(2), pp. 700–703. Available at: https://doi.org/10.1016/j.soilbio.2006.09.008

Adl, S. M. *et al.* (2012). 'The revised classification of Eukaryotes', *The Journal of Eukaryotic Microbiology*, 59(5), pp. 429–493. Available at: https://doi.org/10.1111/j.1550-7408.2012.00644.x

Akhtar, H., Lupascu, M. and Sukri, R. (2022). 'Interactions between microtopography, root exudate analogues and temperature determine $CO_2$ and $CH_4$ production rates in fire-degraded tropical peat', *Soil Biology and Biochemistry*, 169, p. 108646. Available at: https://doi.org/10.1016/j.soilbio.2022.108646

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

An, S. M. *et al.* (2018). 'Next-generation sequencing reveals the diversity of benthic diatoms in tidal flats', *Algae*, 33(2), pp. 167–180. Available at: https://doi.org/10.4490/ALGAE.2018.33.4.3

An, S. M., Choi D. H. and Noh, J. H. (2020). 'High-throughput sequencing analysis reveals dynamic seasonal succession of diatom assemblages in a temperate tidal flat', *Estuarine, Coastal and Shelf Science,* 237, p. 106686. Available at: https://doi.org/10.1016/j.ecss.2020.106686

Andersen, R. *et al.* (2017) 'An overview of the progress and challenges of peatland restoration in Western Europe', *Restoration Ecology*, 25(2), pp. 271–282. Available at: https://doi.org/10.1111/REC.12415

Andersen, R. *et al.* (2024). 'Blanket bog vegetation response to wildfire and drainage suggests resilience to low severity, infrequent burning', *Fire Ecology*, 20(1), pp. 1–16. Available at: https://doi.org/10.1186/s42408-024-00256-0

Andersen, R., Chapman, S. J. and Artz, R. R. E. (2013). 'Microbial communities in natural and disturbed peatlands: A review', *Soil Biology and Biochemistry*, 57, pp. 979–994. Available at: https://doi.org/10.1016/j.soilbio.2012.10.003

Anderson, K. and Downing, J. (2006). 'Dry and wet atmospheric deposition of nitrogen, phosphorus and silicon in an agricultural region', *Water, Air, and Soil Pollution,* 176, pp. 351–374. Available at: https://doi.org/10.1007/s11270-006-9172-4

Anderson, M., Walsh, D. (2013). 'PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing?', *Ecological Monographs*, 83(4), pp. 557-574. Available at: https://doi.org/10.1890/12-2010.1

Artz, R.*et al.* (2019). 'The State of UK Peatlands: an update'. IUCN UK Peatland Programme. [Online] Available at: https://www.iucn-uk-peatlandprogramme.org/sites/default/files/2019-11/COI%20State_of_UK_Peatlands.pdf

Atherton, I., Bosanquet, S. and Lawley, M. (Eds.). (2010). *Mosses and Liverworts of Britain and Ireland – a field guide*. Plymouth: British Bryological Society. Latimer Trend & Co. Ltd.

Azam, F. *et al.* (1983). 'The Ecological Role of Water-Column Microbes in the Sea', *Marine Ecology - Progress Series*, 10, pp. 257–263. Available at: https://doi.org/10.3354/meps010257

Baek, C., Shin, S.K. and Yi, H. (2019). '*Phenylobacterium parvum* sp. Nov., isolated from lake water', *International Journal of Systematic and Evolutionary Microbiology*, 69(4), pp. 1169–1172. Available at: https://doi.org/10.1099/IJSEM.0.003284

Baker W. J. *et al*. (2022). 'A comprehensive phylogenomic platform for exploring the angiosperm tree of life', *Systematic Biology*, 71, pp. 301–319. Data release 3.0 (April 2023). Available at: https://doi.org/10.1093/sysbio/syab035

Bakker, M.G. *et al.* (2014). 'Diffuse symbioses: roles of plant–plant, plant–microbe and microbe–microbe interactions in structuring the soil microbiome', *Molecular Ecology*, 23(6), pp. 1571–1583. Available at: https://doi.org/10.1111/mec.12571

Ballesteros, I. *et al.* (2021). 'DNA barcoding approach to characterize microalgae isolated from freshwater systems in Ecuador', *Neotropical Biodiversity*, 7(1), pp. 170-183. Available at: https://doi.org/10.1080/23766808.2021.1920296

Barreiro, A. and Díaz-Raviña, M. (2021). 'Fire impacts on soil microorganisms: Mass, activity, and diversity', *Current Opinion in Environmental Science & Health,* 22, article number 100264. Available at: https://doi.org/10.1016/j.coesh.2021.100264

Basilier, K. (1979). ‘Moss-associated nitrogen fixation in some mire and coniferous forest environments around Uppsala, Sweden’, *Lindbergia*, 5(2), pp. 84–88. Available at: https://www.jstor.org/stable/pdf/20149332.pdf

Basilier, K. (1980). ‘Fixation and uptake of nitrogen in *Sphagnum* blue-green algal associations’, *Oikos*, 34(2), pp. 239–242. Available at: https://www.jstor.org/stable/pdf/3544188.pdf

Bates, S. T. *et al.* (2013). ‘Global biogeography of highly diverse protistan communities in soil’, *The ISME Journal*, 7(3), pp. 652–659. Available at: https://doi.org/10.1038/ismej.2012.147

Bavestrello, G. *et al.* (2000). ‘Parasitic diatoms inside Antarctic sponges’, *The Biological Bulletin*, 198(1), pp. 29–33. Available at: https://doi.org/10.2307/1542801

Beaulne, J. *et al.* (2021). ‘Peat deposits store more carbon than trees in forested peatlands of the boreal biome’. *Scientific Reports,* 11, article no. 2657 [no pagination] . Available at: https://doi.org/10.1038/s41598-021-82004-x

Belcher *et al*. (2021). ‘UK wildfires and their climate challenges’. Expert Led Report prepared for the third UK Climate Change Risk Assessment. [Online] Available at: https://www.ukclimaterisk.org/wp-content/uploads/2021/06/UK-Wildfires-and-their-Climate-Challenges.pdf

Bengtsson, F. *et al.* (2020). ‘Mechanisms behind species-specific water economy responses to water level drawdown in peat mosses’, *Annals of Botany*, 126(2), pp. 219-30. Available at: https://doi.org/10.1093/aob/mcaa033

Berendse, F. *et al.* (2001). ‘Raised atmospheric $CO_2$ levels and increased N deposition cause shifts in plant species composition and production in *Sphagnum* bogs’, *Global Change Biology*, 7(5), pp. 591-598. Available at: https://doi.org/10.1046/j.1365-2486.2001.00433.x

Berg, A., Danielsson, Å. and Svensson, B. H. (2013). ‘Transfer of fixed-N from $N_2$-fixing cyanobacteria associated with the moss *Sphagnum riparium* results in enhanced growth of the moss’, *Plant and Soil*, 362(1–2), pp. 271–278. Available at: https://doi.org/10.1007/S11104-012-1278-4.

Berney, C. *et al.* (2017). ‘UniEuk: Time to speak a common language in protistology!’, *The Journal of Eukaryotic Microbiology,* 64(1), pp. 407–411. Available at: https://doi.org/10.1111/jeu.12414

Bianchi, S. *et al*., *(2017)* 'Rapid assessment of forest canopy and light regime using smartphone hemispherical photography', *Ecology and Evolution*, 7(24), pp. 10556–10566. Available at: https://doi.org/10.1002/ece3.3567

Bishop, R. R. *et al*. (2018). 'Deforestation and human agency in the North Atlantic region: archaeological and palaeoenvironmental evidence from the Western Isles of Scotland', *Proceedings of the Prehistoric Society,* 84, pp. 145-184. Available at: https://doi.org/10.1017/ppr.2018.8

Bobbink, R., Hornung, M. and Roelofs, J. G. M. (1998). 'The effects of air-borne nitrogen pollutants on species diversity in natural and semi-natural European vegetation', *Journal of Ecology*, 86(5), pp. 717–738. Available at: https://doi.org/10.1046/j.1365-2745.1998.8650717.x

Boenigk, J. *et al.* (2002). 'Food concentration-dependent regulation of food selectivity of interception-feeding bacterivorous nanoflagellates', *Aquatic Microbial Ecology*, 27, pp. 195–202. Available at: https://doi.org/10.3354/ame027195

Booth, R. K. (2002). 'Testate amoebae as paleoindicators of surface-moisture changes on Michigan peatlands: modern ecology and hydrological calibration', *Journal of Paleolimnology*, 28(3), pp. 329–348. Available at: https://doi.org/10.1023/A:1021675225099

Borcard, D., Gillet, F., and Legendre, P. (2018). *Numerical Ecology with R*. 2nd edition. Cham: Springer International Publishing.

Bothe, H. *et al*. (2010). 'Nitrogen fixation and hydrogen metabolism in cyanobacteria' *Microbiology and Molecular Biology Review,* 74(4), pp. 529-51. Available at: https://doi.org/10.1128/mmbr.00033-10

Boucher, O. (2012). 'Comparison of physically- and economically-based $CO_2$-equivalences for methane', *Earth System Dynamics*, 3(1), pp. 49–61. Available at: https://doi.org/10.5194/ESD-3-49-2012

Bragazza, L. *et al.* (2004). 'Nutritional constraints in ombrotrophic *Sphagnum* plants under increasing atmospheric nitrogen deposition in Europe', *The New Phytologist,* 163(3), pp. 609–16. Available at: https://doi.org/10.1111/j.1469-8137.2004.01154.x

Bragazza, L. *et al.*(2006). 'Atmospheric nitrogen deposition promotes carbon loss from peat bogs', *PNAS,* 103(51), pp. 19386-19389. Available at: https://doi.org/10.1073/pnas.0606629104

Bragina, A. *et al*. (2012). '*Sphagnum* mosses harbour highly specific bacterial diversity during their whole lifecycle', *The ISME Journal*, 6(4), pp. 802-13. Available at: https://doi.org/10.1038/ismej.2011.151

Breitwieser, F. P. and Salzberg, S. L. (2020). 'Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification', *Bioinformatics*, 36(4), pp. 1303–1304. Available at: https://doi.org/10.1093/bioinformatics/btz715

Brown, S. P. *et al.* (2019). 'Context dependent fungal and bacterial soil community shifts in response to recent wildfires in the Southern Appalachian Mountains', *Forest Ecology and Management*, 451, p. 117520. Available at: https://doi.org/10.1016/J.FORECO.2019.117520

BSBI (2025a). *Taxon Lists* [Online]. The Botanical Society of Britain and Ireland. Available at: https://bsbi.org/taxon-lists [Accessed 6 August 2025].

BSBI (2025b). *Taxon search: Zingiber mioga (Thunb.) Roscoe* [Online]. The Botanical Society of Britain and Ireland. Available at: https://bsbi.org/taxon?taxonid=2cd4p9h.gwdp5f&class=TaxonInstance [Accessed 6 August 2025].

Bu, Z.J. *et al.* (2013). 'Facilitation vs. competition: Does interspecific interaction affect drought responses in *Sphagnum*?', *Basic and Applied Ecology*, 14(7), pp. 574–584. Available at: https://doi.org/10.1016/J.BAAE.2013.08.002

Burke, D. J., Smemo, K. A. and Hewins, C. R. (2014). 'Ectomycorrhizal fungi isolated from old-growth northern hardwood forest display variability in extracellular enzyme activity in the presence of plant litter', *Soil Biology and Biochemistry*, 68, pp. 219–222. Available at: https://doi.org/10.1016/J.SOILBIO.2013.10.013

Burn, W. (2021). *The hidden half: Blanket bog microbial communities across a spectrum of site and management conditions and impacts on peatland carbon cycling.* University of York. PhD.

Busch, A. and Hess, S. (2017). 'The cytoskeleton architecture of algivorous protoplast feeders (Viridiraptoridae, Rhizaria) indicates actin-guided perforation of prey cell walls', *Protist*, 168(1), pp. 12–31. Available at: https://doi.org/10.1016/J.PROTIS.2016.10.004

Buttler, A. *et al.* (1996). 'Vertical patterns of testate amoebae (Protozoa: Rhizopoda) and peat-forming vegetation on cutover bogs in the Jura, Switzerland', *New Phytologist,* 134(2), pp. 371–382. Available at: https://doi.org/10.1111/j.1469-8137.1996.tb04642.x

Caisová, L. (2016). '*Dicranochaete* – an enigmatic green alga with surprising adaptive capabilities', *Phycologia*, 55(2), pp. 219-229. Available at: https://doi.org/10.2216/15-135.1

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

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

Cardon, Z., Gray, D. and Lewis, L. (2008). 'The green algal underground: Evolutionary secrets of desert cells', *BioScience,* 58(2), pp. 114-121. Available at: https://doi.org/10.1641/B580206

Carrell, A.A. *et al*. (2022). 'Novel metabolic interactions and environmental conditions mediate the boreal peatmoss-cyanobacteria mutualism', *The ISME Journal*,16(4), pp. 1074–1085. Available at: https://doi.org/10.1038/s41396-021-01136-0

Chen, X. *et al*. (2016). 'Diatoms are better indicators of urban stream conditions: A case study in Beijing, China', *Ecological Indicators*, 60, pp. 265–274. Available at: https://doi.org/10.1016/J.ECOLIND.2015.06.039

Clymo, R. S. (1963). 'Ion exchange in *Sphagnum* and its relation to bog ecology', *Annals of Botany*, 27(106), pp. 309–324. Available at: https://doi.org/10.2307/42907700

Clymo, R. S. and Bryant, C. L. (2008). 'Diffusion and mass flow of dissolved carbon dioxide, methane, and dissolved organic carbon in a 7-m deep raised peat bog', *Geochimica et Cosmochimica Acta*, 72(8), pp. 2048–2066. Available at: https://doi.org/10.1016/j.gca.2008.01.032

Clymo, R. S., Turunen, J. and Tolonen, K. (1998). 'Carbon Accumulation in Peatland', *Oikos*, 81(2), pp. 368–388. Available at: https://doi.org/10.2307/3547057

Clymo, R. S. and Williams, M. M. R. (2012). 'Diffusion of gases dissolved in peat pore water', *Mires and Peat,* 10, article no. 6 [no pagination]. Available at: http://www.mires-and-peat.net/pages/volumes/map10/map1006.php

Coesel, P. F. M. (1982). 'Structural characteristics and adaptations of Desmid communities', *The Journal of Ecology*, 70(1), p. 163. Available at: https://doi.org/10.2307/2259871

Coppins, S. and Coppins, B. (2008). 'Findhorn Dunes Lichen Survey'. The Findhorn Dunes Trust. [Online] Available at: https://www.findhornhinterland.org/ecology/lichen-survey

Creevy, A. L. *et al.* (2016). 'Protist diversity on a nature reserve in NW England—With particular reference to their role in soil biogenic silicon pools', *Pedobiologia*, 59(1–2), pp. 51–59. Available at: https://doi.org/10.1016/J.PEDOBI.2016.02.001

Creevy, A. *et al*. (2018). 'Testate amoebae as functionally significant bioindicators in forest-to-bog restoration', *Ecological Indicators,* 84, pp. 274-282. Available at: https://doi.org/10.1016/j.ecolind.2017.08.062

Creevy, A. *et al*. (2023). 'Testate amoebae response and vegetation composition after plantation removal on a former raised bog', *European Journal of Protistology*, 89, p. 125977. Available at: https://doi.org/10.1016/j.ejop.2023.125977

Cremen, M.C.M. *et al.* (2019). 'Reassessment of the classification of Bryopsidales (Chlorophyta) based on chloroplast phylogenomic analyses', *Molecular Phylogenetics and Evolution*, 130, pp. 397–405. Available at: https://doi.org/10.1016/J.YMPEV.2018.09.009

Czech, L. *et al.* (2022). 'Metagenomic analysis using phylogenetic placement—a review of the first decade', *Frontiers in Bioinformatics*, 2, p. 871393. Available at: https://doi.org/10.3389/FBINF.2022.871393

Darcy, J. L. *et al.* (2017). 'Spatial autocorrelation of microbial communities atop a debris-covered glacier is evidence of a supraglacial chronosequence', *FEMS Microbiology Ecology*, 93(8), p. 95. Available at: https://doi.org/10.1093/FEMSEC/FIX095

Davies, G. M. *et al.* (2016). 'Vegetation structure and fire weather influence variation in burn severity and fuel consumption during peatland wildfires', *Biogeosciences*, 13(2), pp. 389–398. Available at: https://doi.org/10.5194/bg-13-389-2016

De Boer, S. H. *et al*. (1995). 'Attenuation of PCR inhibition in the presence of plant compounds by addition of BLOTTO', *Nucleic Acids Research*, 23(13), pp. 2567–2568. Available at: https://doi.org/10.1093/nar/23.13.2567

De Cáceres, M. and Legendre, P. (2009). 'Associations between species and groups of sites: indices and statistical inference', *Ecology*, 90(12), pp. 3566-3574. Available at: https://doi.org/10.1890/08-1823.1

Defra (2020). *Defra UK air information resource.* [Online]. Available at: https://uk-air.defra.gov.uk/data/ [Accessed 9 October 2020]. © Crown 2025 copyright Defra via uk-air.defra.gov.uk, licenced under the Open Government Licence (OGL).

Del Campo, J. *et al*. (2018). 'EukRef: Phylogenetic curation of ribosomal RNA to enhance understanding of eukaryotic diversity and distribution' *PLoS Biology,* 16(9), pp. e2005849–14. Available at: https://doi.org/10.1371/journal.pbio.2005849

Diaz, G. M. (2017). Gastón Mauro Díaz's Quick Files lens_calibration_sheet.pdf [Online] OSF. Last updated: 30 January 2025. Available at: https://osf.io/tudzc [Accessed 21 January 2024]

Díaz, G. M. (2023). rcaiman: CAnopy IMage Analysis. R package (version 1.2.2). [Software] Available at: https://CRAN.R-project.org/package=rcaiman

Díaz, G. M., Lang, M. and Kaha, M. (2024). 'Simple calibration of fisheye lenses for hemispherical photography of the forest canopy', *Agricultural and Forest Meteorology*, 352, p. 110020. Available at: https://doi.org/10.1016/J.AGRFORMET.2024.110020

Dooley, S. and Treseder, K. (2012). 'The effect of fire on microbial biomass: A meta-analysis of field studies', *Biogeochemistry*, 109, pp. 49-61. Available at: https://doi.org/10.1007/s10533-011-9633-8

Dragosits *et al*. (2020). 'Nitrogen Futures. JNCC Report No. 665.' JNCC, Peterborough, ISSN 0963-8091. Available at: https://data.jncc.gov.uk/data/04f4896c-7391-47c3-ba02-8278925a99c5/JNCC-Report-665-FINAL-WEB.pdf

Dufrêne, M., and Legendre, P. (1997). 'Species assemblages and indicator species: the need for a flexible asymmetrical approach', *Ecological Monographs,* 67(3), pp. 345–66. Available at: https://doi.org/10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2

Dunthorn, M. *et al.* (2014). 'Placing environmental next-generation sequencing amplicons from microbial eukaryotes into a phylogenetic context', *Molecular Biology and Evolution*, 31(4), pp. 993–1009. Available at: https://doi.org/10.1093/MOLBEV/MSU055

Dyer, A. F., and Lindsay, S. (1992). 'Soil spore banks of temperate ferns', *American Fern Journal*, 82(3), pp. 89–123. Available at: https://doi.org/10.2307/1547792

Eddy, A. (1977). 'Sphagnales of tropical Asia', *Bulletin of the British Museum (Natural History) Botany,* 5, pp. 359 – 445. Available at: https://www.biodiversitylibrary.org/page/2297912

Ekelund, F. and Rønn, R. (1994). 'Notes on protozoa in agricultural soil with emphasis on heterotrophic flagellates and naked amoebae and their ecology', *FEMS Microbiology Reviews*, 15(4), pp. 321–353. Available at: https://doi.org/10.1111/j.1574-6976.1994.tb00144.x

Elbert, W. *et al.* (2012). 'Contribution of cryptogamic covers to the global cycles of carbon and nitrogen', *Nature Geoscience*, 5(7), pp. 459–462. Available at: https://doi.org/10.1038/ngeo1486

Elliott, D. R. *et al.* (2015). 'Bacterial and fungal communities in a degraded ombrotrophic peatland undergoing natural and managed re-vegetation.', *PloS One*, 10(5), p. e0124726. Available at: https://doi.org/10.1371/journal.pone.0124726

European Environment Agency (2016). 'European Union emission inventory report 1990–2014 under the UNECE Convention on Long-range Transboundary Air Pollution (LRTAP).' EEA report 16/2016 [Online]. Available at: https://doi.org/10.2800/628267

Evans, C. *et al*. (2017). 'Implementation of an Emissions Inventory for UK Peatlands.' A report to the Department for Business, Energy and Industrial Strategy. [Online] Available at: https://uk-air.defra.gov.uk/reports/cat07/1904111135_UK_peatland_GHG_emissions.pdf

Fazekas, A. J. *et al*. (2012) DNA barcoding methods for land plants. In Kress, W., Erickson, D. (Eds). *DNA Barcodes*. Methods in Molecular Biology, vol 858. Totowa, NJ: Humana Press. pp. 223–252. Available at: https://doi.org/10.1007/978-1-61779-591-6_11

Fehling, J., Stoecker, D. and Baldauf, S. (2007). Chapter 6 - Photosynthesis and the Eukaryote tree of life. In Falkowski, P. G. and Knoll, A. H. (Eds). *Evolution of Primary Producers in the Sea*. Cambridge, Massachusetts: Academic Press. pp. 75-107. Available at: https://doi.org/10.1016/B978-012370518-1/50007-2

Ferrenberg, S. *et al.* (2013). 'Changes in assembly processes in soil bacterial communities following a wildfire disturbance', *The ISME Journal*, 7(6), pp. 1102–1111. Available at: https://doi.org/10.1038/ISMEJ.2013.11

Flannigan, M. *et al.* (2013). 'Global wildland fire season severity in the 21st century', *Forest Ecology and Management*, 294, pp. 54–61. Available at: https://doi.org/10.1016/J.FORECO.2012.10.022

Forestry Commission Scotland (2015). 'Deciding future management options for afforested deep peatland.' Forestry Commission Scotland, Edinburgh.

Forestry Commission and Natural England (2023). ‘Decision support framework for peatland protection, the establishment of new woodland and re-establishment of existing woodland on peatland in England.’

Franklin, R. B. and Mills, A. L. (2003). ‘Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field’, *FEMS Microbiology Ecology*, 44(3), pp. 335–346. Available at: https://doi.org/10.1016/S0168-6496(03)00074-6

Freeman, K. R. *et al.* (2009). ‘Soil CO2 flux and photoautotrophic community composition in high-elevation, “barren” soil’, *Environmental Microbiology,* 11(3), pp. 674–686. Available at: https://doi.org/10.1111/j.1462-2920.2008.01844.x

Frey, B. *et al.* (2013). ‘Molecular characterization of phototrophic microorganisms in the forefield of a receding glacier in the Swiss Alps’, *Environmental Research Letters,* 8(1), p. 015033. https://doi.org/10.1088/1748-9326/8/1/015033

Frisk, C. *et al*. (2022). ‘Atmospheric transport reveals grass pollen dispersion distances’, *Science of The Total Environment*, 814, p. 152806.Available at: https://doi.org/10.1016/j.scitotenv.2021.152806

GBIF (2025). *Zingiber mioga (Thunb.)* Roscoe [Online]. The Global Biodiversity Information Facility. Available at: https://www.gbif.org/species/2757086 [Accessed 6 August 2025].

Geisen, S. *et al.* (2018). ‘Soil protists: a fertile frontier in soil biology research’, *FEMS Microbiology Reviews*, 42(3), pp. 293–323. Available at: https://doi.org/10.1093/femsre/fuy006

Geisen, S. and Bonkowski, M. (2018). ‘Methodological advances to study the diversity of soil protists and their functioning in soil food webs’, *Applied Soil Ecology*, 123, pp. 328–333. Available at: https://doi.org/10.1016/J.APSOIL.2017.05.021

Geneious Prime (version 2023.2.1) [Software] Available at: https://www.geneious.com/

Gibbs, S.P. (1981). ‘The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae’, *Annals of the New York Academy of Sciences*, 361(1), pp. 193–208. Available at: https://doi.org/10.1111/J.1749-6632.1981.TB54365.X

Gifford, D. J. (1991). ‘The Protozoan-Metazoan trophic link in pelagic ecosystems’, *The Journal of Protozoology,* 38(1), pp. 81–86. Available at: https://doi.org/10.1111/j.1550-7408.1991.tb04806.x

Gilbert, D. *et al*. (1998a). ‘Short-term effect of nitrogen enrichment on the microbial communities of a peatland’, *Hydrobiologia*, 373/374(0), pp. 111–119. Available at: https://doi.org/10.1023/A:1017091926454

Gilbert, D. *et al*. (1998b). ‘The microbial loop at the surface of a peatland: structure, function, and impact of nutrient input’, *Microbial Ecology*, 35(1), pp. 83–93. Available at: https://doi.org/10.1007/s002489900062

Gilbert D. and Mitchell E. A. D. (2006). ‘Microbial diversity in *Sphagnum* peatlands. In Martini, I. P., Matinez Cortizas, A. and Chesworth, W. (Eds). *Peatlands: basin evolution and depository of records on global environmental and climatic changes*. Developments in Earth Surface Processes series. Amsterdam: Elsevier, pp. 287–318.

Godfrey, M. and Rogers, K. (2021). *Sphagnum* mosses*: field key to the mosses of Britain and Ireland*. Telford: FSC Publications AIDGAP.

Gomaa, F. *et al*. (2014). ‘One alga to rule them all: unrelated mixotrophic testate amoebae (Amoebozoa, Rhizaria and Stramenopiles) share the same symbiont (Trebouxiophyceae)’, *Protist*, 165(2), pp. 161–176. Available at: https://doi.org/10.1016/J.PROTIS.2014.01.002

Gomes da Silva, Y. *et al*. (2022). ‘Testate amoebae (Amorphea, Amoebozoa, Cercozoa) as bioindicators: a scientometric review’, *Acta Limnologica Brasiliensia,* 34, p. e20. Available at: https://doi.org/10.1590/S2179-975X0722

González Garraza, G., Burdman, L. and Mataloni, G. (2019). ‘Desmids (Zygnematophyceae, Streptophyta) community drivers and potential as a monitoring tool in South American peat bogs’, *Hydrobiologia*, 833, pp. 125-141. Available at: https://doi.org/10.1007/s10750-019-3895-x

Goodyer, E. (2014). *Quantifying the Desmid diversity of Scottish blanket mires*. University of Aberdeen. PhD.

Gorham, E. (1991). ‘Northern peatlands: role in the carbon cycle and probable responses to climatic warming’, *Ecological Applications*, 1(2), pp. 182–195. Available at: https://doi.org/10.2307/1941811

Granhall, U. and Selander, H. (1973), ‘Nitrogen Fixation in a Subarctic Mire’, *Oikos*, 24(1), pp. 8–15. Available at: https://www.jstor.org/stable/pdf/3543247.pdf

Guiry, M. D. and Guiry, G. M. (2025). *AlgaeBase :: Listing the World’s Algae*. World-wide electronic publication, National University of Ireland, Galway. [Online] Available at: https://www.algaebase.org/ [Accessed 4 May 2025].

Guiry, M.D. in Guiry, M. D. and Guiry, G. M. (2018). *AlgaeBase*. World-wide electronic publication, National University of Ireland, Galway. [Online]. Available at: https://www.algaebase.org [Accessed 24 April 2025].

Haas, B.J. *et al.* (2011). 'Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons', *Genome research*, 21(3), pp. 494–504. Available at: https://doi.org/10.1101/GR.112730.110

Hamard, S. *et al.* (2021). 'Contribution of microbial photosynthesis to peatland carbon uptake along a latitudinal gradient', *Journal of Ecology*, 109, pp. 3424–3441. Available at: https://doi.org/10.1111/1365-2745.13732

Hambley, G. (2016). *The effect of forest-to-bog restoration on net ecosystem exchange in Flow Country peatlands*. University of St Andrews. PhD.

Hamilton, T. L., Bryant, D. A. and Macalady, J. L. (2016). 'The role of biology in planetary evolution: cyanobacterial primary production in low-oxygen Proterozoic oceans', *Environmental microbiology*, 18(2), pp. 325–40. Available at: https://doi.org/10.1111/1462-2920.13118

Harris, A., Bryant, R. G. and Baird, A. J. (2005). 'Detecting near-surface moisture stress in *Sphagnum* spp.', *Remote Sensing of Environment*, 97(3), pp. 371–381. Available at: https://doi.org/10.1016/J.RSE.2005.05.001

Hartmann, M. *et al.* (2012). 'Mixotrophic basis of Atlantic oligotrophic ecosystems', *PNAS*, 109(15), pp. 5756–5760. Available at: https://doi.org/10.1073/pnas.1118179109

Heal, O. W. (1962). 'The abundance and micro-distribution of testate amoebae (Rhizopoda:Testacea) in *Sphagnum*', *Oikos*, 13(1), pp. 35–47. Available at: https://doi.org/10.2307/3565062

Heger, T. J. *et al*. (2018). 'High-throughput environmental sequencing reveals high diversity of litter and moss associated protist communities along a gradient of drainage and tree productivity', *Environmental Microbiology*, *20*(3), pp. 1185–1203. Available at: https://doi.org/10.1111/1462-2920.14061

Hemond, H. F. (1983). 'The nitrogen budget of Thoreau's Bog', *Ecology*, 64(1), pp. 99–109. Available at: https://doi.org/10.2307/1937333

Henniges, M. C., *et al*. (2021). *A taxonomic, genetic and ecological data resource for the vascular plants of Britain and Ireland*, NERC Environmental Information Data Centre. [Online] Available at: https://doi.org/10.5285/9f097d82-7560-4ed2-af13-604a9110cf6d [Accessed 16 December 2023].

Herbert, R. P. *et al.* (2019). 'Light variability and mixotrophy: Responses of testate amoeba communities and shell δ13C values to a peatland shading experiment', *European Journal of Protistology*, 67, pp. 15–26. Available at: https://doi.org/10.1016/J.EJOP.2018.10.005

Hijmans, R. (2024). *geosphere: Spherical Trigonometry*. R package (version 1.5-20). [Software] Available at: https://github.com/rspatial/geosphere

Hinlo, R. *et al*. (2017). 'Methods to maximise recovery of environmental DNA from water samples', *PLoS One* 12(6), p. e0179251. Available at: https://doi.org/10.1371/journal.pone.0179251

Hleap, J. S. *et al.* (2021). 'Assessment of current taxonomic assignment strategies for metabarcoding eukaryotes', *Molecular Ecology Resources*, 21(7), pp. 2190–2203. Available at: https://doi.org/10.1111/1755-0998.13407

Hodson, M.E. *et al.* (2023). 'Enhanced plant growth in the presence of earthworms correlates with changes in soil microbiota but not nutrient availability', *Geoderma*, 433, p. 116426. Available at: https://doi.org/10.1016/j.geoderma.2023.116426

Holden, S. *et al*. (2016). 'Fire severity influences the response of soil microbes to a boreal forest fire', *Environmental Research Letters*, 11, p. 035004. Available at: https://doi.org/10.1088/1748-9326/11/3/035004

Holden, S. R., Berhe, A. A. and Treseder, K. K. (2015). 'Decreases in soil moisture and organic matter quality suppress microbial decomposition following a boreal forest fire', *Soil Biology and Biochemistry*, 87, pp. 1–9. Available at: https://doi.org/10.1016/J.SOILBIO.2015.04.005

Holden, S., Gutierrez, A. and Treseder, K. (2013). 'Changes in soil fungal communities, extracellular enzyme activities, and litter decomposition across a fire chronosequence in Alaskan boreal forests', *Ecosystems*, 16, pp. 34–46. Available at: https://doi.org/10.1007/s10021-012-9594-3

Holman, I. P. (2009)., 'An estimate of peat reserves and loss in the East Anglian Fens'. Commissioned by the RSPB. [Online] Available at: http://ww2.rspb.org.uk/Images/Fenlandpeatassessment_tcm9-236041.pdf

House, J. I. *et al*. (2010). 'Climate change and the British Uplands: evidence for decision-making', *Climate Research*, *45*, pp. 3–12. Available at: http://doi.org/10.3354/cr00982

Hsieh, T., Ma, K.and Chao, A. (2020). iNEXT: Interpolation and Extrapolation for Species Diversity. R package (version 2.0.20). [Software] Available at: http://chao.stat.nthu.edu.tw/wordpress/software_download/.

Hugin (2009). (version 2023.0.0). [Software] Available at: http://hugin.sourceforge.net/

iNaturalist (2025). *Myoga Ginger (Zingiber mioga)* [Online]. iNaturalist Available at: https://www.inaturalist.org/taxa/151801-Zingiber-mioga [Accessed 6 August 2025].

Ingram, H. A. P. (1978). 'Soil layers in mires: Function and terminology', *Journal of Soil Science*, 29(2), pp. 224–227. Available at: http://doi.org/10.1111/j.1365-2389.1978.tb02053.x

IUCN (2018). 'UK Peatland Strategy 2018–2040', IUCN UK Peatland Programme, Newark. [Online] Available at: https://www.iucn-uk-peatlandprogramme.org/sites/default/files/2022-04/UK%20Peatland%20Strategy%202018_2040.pdf

Jassey, V. E. J. *et al.* (2011). 'Effect of a temperature gradient on *Sphagnum fallax* and its associated living microbial communities: a study under controlled conditions', *Canadian Journal of Microbiology,* 57(3), pp. 226-235. Available at: http://doi.org/10.1139/W10-116

Jassey, V. E. J. *et al.* (2015). 'An unexpected role for mixotrophs in the response of peatland carbon cycling to climate warming', *Scientific Reports*, 5(1), p. 16931. Available at: http://doi.org/10.1038/srep16931

Jassey, V. E. J. *et al*. (2022). 'Photosynthetic microorganisms effectively contribute to bryophyte $CO_2$ fixation in boreal and tropical regions', *ISME Communications*, 2(1), p. 64. Available at: https://doi.org/10.1038/s43705-022-00149-w

Johnson, K. *et al.* (2008). 'Ecological observations and new locations of a rare moss *Ambuchanania leucobryoides* (Ambuchananiaceae)', *Papers and Proceedings of the Royal Society of Tasmania*, 142(2), pp. 79–84. Available at: https://doi.org/10.26749/RSTPP.142.2.79

Joint Nature Conservation Committee (2011). 'Towards an assessment of the state of UK peatlands.' JNCC Report No. 445. [Online] Available at: http://jncc.defra.gov.uk/pdf/jncc445_web.pdf [Accessed 19 February 2019].

Kapustin, D. and Kulikovskiy, M. (2023) '*Sphagnum*-associated chrysophytes from the 'Dermansko-Ostrozsky' National Nature Park (Ukraine)', *Mires and Peat*, 29, article no. 22 [no pagination]. Available at: https://doi.org/10.19189/MaP.2023.OMB.Sc.2111526

Kettridge, N. *et al*. (2013). 'The ecohydrology of forested peatlands: Simulating the effects of tree shading on moss evaporation and species composition', *Journal of Geophysical Research: Biogeosciences*, 118(2), pp. 422–435, Available at: https://doi.org/10.1002/jgrg.20043

Khan, T. F. (2022). *Interactions of microplastic, soil and plant with a particular emphasis on phosphorus*. University of York. PhD.

Killham, K. (1994). *Soil ecology*. 1st Edition. Cambridge: Cambridge University Press.

Kögel-Knabner, I. (2002) 'The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter', *Soil Biology and Biochemistry*, 34(2), pp. 139–162. Available at: https://doi.org/10.1016/S0038-0717(01)00158-4

Kohlenberg, A. J. *et al*. (2018). 'Controls on boreal peat combustion and resulting emissions of carbon and mercury', *Environmental Research Letters*, 13(3), p. 035005. Available at: https://doi.org/10.1088/1748-9326/AA9EA8

Kopczynski, E. D., Bateson, M. M. and Ward, D. M. (1994). 'Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms', *Applied and Environmental Microbiology*, 60(2), pp. 746–748. Available at: https://doi.org/10.1128/AEM.60.2.746-748.1994

Koppelle, S. *et al*. (2022). 'Mixotrophy in the bloom-forming genus *Phaeocystis* and other haptophytes', *Harmful Algae,* 117, 102292. Available at: https://doi.org/10.1016/j.hal.2022.102292

Kosecka, M. *et al*. (2022). 'Phylogeny and ecology of *Trebouxia* photobionts from Bolivian lichens', *Frontiers in Microbiology*, 13, p. 779784. Available at: https://doi.org/10.3389/FMICB.2022.779784

Kress, J. and Erickson, D. (2007). 'A two-locus global DNA barcode for land plants: The coding rbcL gene complements the non-coding trnHpsbA spacer region', *PLoS One* 2(6), p. e508. Available at: https://doi.org/10.1371/journal.pone.0000508

Kress, W. J., Prince, L. M. and Williams, K. J. (2002). 'The phylogeny and a new classification of the gingers (Zingiberaceae): evidence from molecular data', *American Journal of Botany*, 89(10), p. 1682-1696. Available at: https://doi.org/10.3732/ajb.89.10.1682

Kreutz, M. and Foissner, W. (2006). 'The *Sphagnum* ponds of Simmelried in Germany : a biodiversity hot-spot for microscopic organisms', *Protozoological Monographs*. Edited by W. Foissner, 3, pp. 1–267.

Kunin, V. *et al.* (2010). 'Wrinkles in the rare biosphere: Pyrosequencing errors can lead to artificial inflation of diversity estimates', *Environmental microbiology*, 12(1), pp. 118–123. Available at: https://doi.org/10.1111/J.1462-2920.2009.02051.X

Küttim, M. *et al.* (2020). 'Controls of *Sphagnum* growth and the role of winter', *Ecological Research*, 35(1), pp. 219–234. Available at: https://doi.org/10.1111/1440-1703.12074

Labeda, D. P. *et al.* (2010). 'Emended description of the genus *Actinokineospora* Hasegawa 1988 and transfer of *Amycolatopsis fastidiosa* Henssen *et al.* 1987 as *Actinokineospora fastidiosa* comb. nov.', *International Journal of Systematic and Evolutionary Microbiology*, 60(6), pp. 1444–1449. Available at: https://doi.org/10.1099/IJS.0.016568-0

Lahr, D. J. G., Lara, E. and Mitchell, E. A. D. (2012). 'Time to regulate microbial eukaryote nomenclature', *Biological Journal of the Linnean Society*, 107(3), pp. 469–476. Available at: https://doi.org/10.1111/j.1095-8312.2012.01962.x

Lahti, L. and Shetty, S. (2017). Microbiome: Tools for microbiome analysis in R. R package (version 1.28.0). [Software] Available at: http://microbiome.github.com/microbiome

Lauber, C. L. *et al.* (2009). 'Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale', *Applied and Environmental Microbiology*, 75(15), pp. 5111–5120. Available at: https://doi.org/10.1128/AEM.00335-09

Leff, J. W. (2022). mctoolsr: Microbial Community Data Analysis Tools. R package (version 0.1.1.9.). [Software] Available at: https://github.com/leffj/mctoolsr

Lejzerowicz, F. *et al.* (2010). 'Molecular evidence for widespread occurrence of Foraminifera in soils', *Environmental Microbiology*, 12(9), pp. 2518–2526. Available at: https://doi.org/10.1111/j.1462-2920.2010.02225.x

Levy, P. *et al.* (2019). 'Response of a peat bog vegetation community to long-term experimental addition of nitrogen', *Journal of Ecology*, 107(3), pp. 1167–1186. Available at: https://doi.org/10.1111/1365-2745.13107

Liang, Z. and Keeley, A. (2013). 'Filtration Recovery of Extracellular DNA from Environmental Water Samples', *Environmental Science & Technology*. 47(16), pp. 9324−9331. Available at: https://doi.org/10.1021/es401342b

Lindsay, R. et al. (1988). 'The Flow Country: The peatlands of Caithness and Sutherland.' Peterborough: Joint Nature Conservation Committee. [Online] Available at: http://roar.uel.ac.uk/1155/ [Accessed: 19 February 2019].

Lindsay, R., Birnie, R. and Clough, J. (2014). 'IUCN UK Committee Peatland Programme Briefing note no 3: impacts of artificial drainage on peatlands.' [Online] Available at: https://www.researchgate.net/publication/268220701_IUCN_UK_Peatland_Programme_Briefing_Note_No_3_Impacts_of_Artificial_Drainage_on_Peatlands#:~:text=Impacts%20of%20Drainage%20Main%20impact,in%20milk%20than%20in%20peat%22

Lindsay, R., Birnie, R. and Clough, J. (2016). 'IUCN UK Committee Peatland Programme Briefing Note No. 11: Peatland Restoration.' [Online] Available at: https://repository.uel.ac.uk/download/517ff1198b0e95d77107a07e3a648f374413cf516a10d3475d5d86c2af5b193f/6200353/11%20Peatland%20Restoration_FINAL.pdf

Lindsay, R. and Immirzi, P. (1996). 'An inventory of lowland raised bogs in Great Britain.' Scottish Natural Heritage. Perth. Scottish Natural Heritage Research, Survey and Monitoring Report No 78. [Online] Available at: https://repository.uel.ac.uk/item/86q77

Lu, F. *et al.* (2023). 'Long-term phosphorus addition strongly weakens the carbon sink function of a temperate peatland', *Ecosystems*, 26(1), pp. 201–216. Available at: https://doi.org/10.1007/S10021-022-00754-9

Lu, J. *et al.* (2017). 'Bracken: Estimating species abundance in metagenomics data', *PeerJ Computer Science*, 2017(1), p. e104. Available at: https://doi.org/10.7717/peerj-cs.104/

Lukenbach, M. *et al.* (2015). 'Hydrogeological controls on post-fire moss recovery in peatlands', *Journal of Hydrology,* 530, pp. 405-418. Available at: https://doi.org/10.1016/j.jhydrol.2015.09.075

Lund, M. *et al*. (2009). 'Effects of N and P fertilization on the greenhouse gas exchange in two northern peatlands with contrasting N deposition rates', Biogeosciences, 6(10), pp. 2135–2144. Available at: https://doi.org/10.5194/bg-6-2135-2009

Ma, W. *et al*. (2023). 'Editorial: Algal photosynthesis', *Frontiers in Microbiology*, 13, p. 1112301. Available at: https://doi.org/10.3389/fmicb.2022.1112301

Ma, M., Gong, Y. and Hu, Q. (2018). "Identification and feeding characteristics of the mixotrophic flagellate *Poterioochromonas malhamensis*, a microalgal predator isolated from outdoor massive *Chlorella* culture', *Algal Research*, 29, pp. 142–153. Available at: https://doi.org/10.1016/J.ALGAL.2017.11.024

Magnan, G., Lavoie, M. and Payette, S. (2012). 'Impact of fire on long-term vegetation dynamics of ombrotrophic peatlands in northwestern Québec, Canada', *Quaternary Research*, 77(1), pp. 110-121. Available at: https://doi.org/10.1016/j.yqres.2011.10.006

Magurran, A. (2004). *Measuring Biological Diversity*. Oxford: Wiley-Blackwell. ISBN: 978-0-632-05633-0

Majaneva, M. *et al.* (2018). 'Environmental DNA filtration techniques affect recovered biodiversity', *Scientific Reports,*8, article no. 4682 [no pagination]. Available at: https://doi.org/10.1038/s41598-018-23052-8

Malard, L.A. *et al.* (2021). 'Influence of spatial scale on structure of soil bacterial communities across an Arctic landscape', *Applied and Environmental Microbiology*, 87(5), pp. e02220-20. Available at: https://doi.org/10.1128/AEM.02220-20

Malmer, N. *et al*. (2003). 'Interferences between *Sphagnum* and vascular plants: effects on plant community structure and peat formation', Oikos, 100(3), pp. 469-482. Available at: https://doi.org/10.1034/j.1600-0706.2003.12170.x

Marcisz, K. *et al*. (2014). 'Seasonal changes in *Sphagnum* peatland testate amoeba communities along a hydrological gradient', *European Journal of Protistology,* 50(5), pp. 445-455. Available at: https://doi.org/10.1016/j.ejop.2014.07.001

Marcisz, K. *et al.* (2015). 'Long-term hydrological dynamics and fire history over the last 2000 years in CE Europe reconstructed from a high-resolution peat archive', *Quaternary Science Reviews*, 112, pp. 138–152. Available at: https://doi.org/10.1016/J.QUASCIREV.2015.01.019

Martin, M. (2011). 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet.journal*, 17(1), pp. 10-12. Available at: https://doi.org/10.14806/ej.17.1.200

Martinez Arbizu, P. (2020). pairwiseAdonis: Pairwise multilevel comparison using adonis. R package (version 0.4). [Software] Available at: https://github.com/pmartinezarbizu/pairwiseAdonis

Matheson, C. *et al.* (2010). 'Assessing PCR inhibition from humic substances', *The Open Enzyme Inhibition Journal*, 3(1), pp. 38–45. Available at: https://benthamopen.com/contents/pdf/TOEIJ/TOEIJ-3-38.pdf

Mathisen, I. E. and Hofgaard, A. (2011). 'Recent height and diameter growth variation in Scots pine (*Pinus sylvestris* L.) along the Arctic margin: the importance of growing season versus non-growing season climate factors', *Plant Ecology and Diversity*, 4(1), pp. 1-11. Available at: https://doi.org/10.1080/17550874.2011.570797

Mazei, Y. A., Tsyganov, A. N. and Bubnova, O. A. (2007). 'Structure of a community of testate amoebae in a *Sphagnum* dominated bog in upper sura flow (Middle Volga Territory)', *Biology Bulletin*, 34(4), pp. 382–394. Available at: https://doi.org/10.1134/S1062359007040115

Mchaffie, H. *et al.* (2009). 'Scots pine growing on forested mires in Abernethy Forest, Strathspey, Scotland', *Botanical Journal of Scotland*, 54(2), pp. 209–219. Available at: https://doi.org/10.1080/03746600208685038

McMurdie, P. J. and Holmes, S. (2013). 'phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data', PLoS One, 8(4), p. e61217. Available at: https://doi.org/10.1371/journal.pone.0061217

Mieczan, T., Bronowicka-Mielniczuk, U. and Rudyk-Leuska, N. (2022). 'Effects of fires on microbial and metazoan communities in peatlands', *Water,* 14(21), p. 3402. Available at: https://doi.org/10.3390/W14213402

Mitchell, E. A. D. and Gilbert, D. (2004). 'Vertical micro-distribution and response to nitrogen deposition of testate amoebae in *Sphagnum*', *The Journal of Eukaryotic Microbiology*, 51(4), pp. 480–490. Available at: https://doi.org/10.1111/j.1550-7408.2004.tb00400.x

Mitchell, E. A. D. *et al.* (2000). 'Horizontal distribution patterns of testate amoebae (Protozoa) in a *Sphagnum magellanicum* carpet', *Microbial Ecology*, 39(4), pp. 290–300. Available at: https://doi.org/10.1007/s002489900187

Mlaga, K. D. *et al*. (2021).' HCK and ABAA: A newly designed pipeline to improve fungi metabarcoding analysis', *Frontiers in Microbiology*, *12*, p. 640693. https://doi.org/10.3389/FMICB.2021.640693

Moore, P.A. and Waddington, J.M. (2015). 'Modelling *Sphagnum* moisture stress in response to projected 21st-century climate change,' *Hydrological Processes*, 29(18), pp. 3966–3982. Available at: https://doi.org/10.1002/HYP.10484

Moret, B.M.E., Roshan, U. and Warnow, T. (2002). 'Sequence-length requirements for phylogenetic methods,' In: Guigó, R., Gusfield, D. (eds) *Algorithms in Bioinformatics*. WABI 2002. Lecture Notes in Computer Science, vol 2452, pp. 343–356. Available at: https://doi.org/10.1007/3-540-45784-4_26

National Parks England (2019). 'Written evidence submitted by National Parks England (PLD0027). State of Peatland in England Inquiry.' [Online] Available at: https://committees.parliament.uk/writtenevidence/105553/pdf/

Nelson, K. *et al*. (2021). 'Peatland-fire interactions: A review of wildland fire feedbacks and interactions in Canadian boreal peatlands', *Science of The Total Environment*, 769, p. 145212. Available at: https://doi.org/10.1016/j.scitotenv.2021.145212

Nibau, C. *et al*. (2022). 'Molecular and physiological responses to desiccation indicate the abscisic acid pathway is conserved in the peat moss, *Sphagnum*', *Journal of Experimental Botany*, 73(13), p. 4576. Available at: https://doi.org/10.1093/JXB/ERAC133

Norby, R.J. *et al*. (2023). 'Shading contributes to *Sphagnum* decline in response to warming,' *Ecology and Evolution*, 13(9), p. e10542. Available at: https://doi.org/10.1002/ECE3.10542

Odom, A. R. *et al*. (2023). 'Metagenomic profiling pipelines improve taxonomic classification for 16S amplicon sequencing data', *Scientific reports*, 13(1), article no. 13957 [no pagination]. Available at: https://doi.org/10.1038/S41598-023-40799-X

Ofiti, N. O. E. *et al*. (2023). 'Climate warming and elevated $CO_2$ alter peatland soil carbon sources and stability', *Nature Communications,* 14, article no. 7533 [no pagination]. Available at: https://doi.org/10.1038/s41467-023-43410-z

Oksanen,J. *et al*. (2022). vegan: Community Ecology Package. R package (version 2.6-4). [Software] Available at: https://CRAN.R-project.org/package=vegan

Olrik, K. (2003). 'The occurrence of planktonic desmid species in Danish lakes and ponds related to total nitrogen and phosphorus contents of the water body', *Biologia,* 58(4), pp. 709–716. Available at: https://www.researchgate.net/publication/286989804_The_occurrence_of_planktonic_desmid_species_in_Danish_lakes_and_ponds_related_to_total_nitrogen_and_phosphorus_contents_of_the_water_body

Opelt, K. *et al*. (2007). 'High specificity but contrasting biodiversity of *Sphagnum*-associated bacterial and plant communities in bog ecosystems independent of the geographical region', *The ISME Journal*, 1(6), pp. 502-16. Available at: https://doi.org/10.1038/ismej.2007.58

Ordnance Survey (2025). OSTN15 Coordinate Transformation Tool - Coordinate transformations [Online] Available at: https://www.ordnancesurvey.co.uk/geodesy-positioning/coordinate-transformations/tool [Accessed 29 April 2025].

Oudot-Le Secq, M.P. *et al*. (2007). 'Chloroplast genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*: Comparison with other plastid genomes of the red lineage,' *Molecular Genetics and Genomics*, 277(4), pp. 427–439. Available at: https://doi.org/10.1007/S00438-006-0199-4

Paulson, J. N. *et al*. (2013). *metagenomeSeq: Statistical analysis for sparse high-throughput sequncing*. Bioconductor package. [Software] Available at: http://www.cbcb.umd.edu/software/metagenomeSeq

Payne, R. (2013). 'Seven reasons why protists make good indicators', *Acta Protozoologica*, 52(3), pp. 105-113. Available at: https://doi.org/10.4467/16890027AP.13.0011.1108

Payne, R. (2014). 'The exposure of British peatlands to nitrogen deposition, 1900–2030', *Mires and Peat*, 14, article no. 04, pp. 1–9. Available at: http://mires-and-peat.net/pages/volumes/map14/map1404.php

Payne, R. J. *et al*. (2006). 'Ecology of testate amoebae (Protista) in south-central Alaska peatlands: building transfer-function models for palaeoenvironmental studies', *The Holocene*, 16(3), pp. 403–414. Available at: https://doi.org/10.1191/0959683606hl936rp

Payne, R. J. *et al*. (2012). 'Impact of simulated nitrogen pollution on heathland microfauna, mesofauna and plants', *European Journal of Soil Biology*, 49, pp. 73–79. Available at: https://doi.org/10.1016/J.EJSOBI.2011.08.003

Payne, R. J. *et al*. (2013). 'Ammonia exposure promotes algal biomass in an ombrotrophic peatland', *Soil Biology and Biochemistry*, 57, pp. 936–938. Available at: https://doi.org/10.1016/J.SOILBIO.2012.09.012

Payne, R. J. *et al*. (2016). 'Tree encroachment may lead to functionally-significant changes in peatland testate amoeba communities', *Soil Biology and Biochemistry*, 98, pp. 18-21. Available at: https://doi.org/10.1016/j.soilbio.2016.04.002

Pérez-Burillo, J. *et al*. (2020). 'Evaluation and sensitivity analysis of diatom DNA metabarcoding for WFD bioassessment of Mediterranean rivers', *Science of The Total Environment,* 727, p. 138445. Available at: https://doi.org/10.1016/j.scitotenv.2020.138445

Pérez-Valera, E. *et al*. (2018). 'Resilience to fire of phylogenetic diversity across biological domains', *Molecular Ecology*, 27(13), pp. 2896-2908. Available at: https://doi.org/10.1111/mec.14729

Pérez-Valera, E., Goberna, M. and Verdú, M. (2019). 'Fire modulates ecosystem functioning through the phylogenetic structure of soil bacterial communities', *Soil Biology and Biochemistry*, 129, pp. 80-89. Available at: https://doi.org/10.1016/j.soilbio.2018.11.007

Pipes, L. and Nielsen, R. (2024). 'A rapid phylogeny-based method for accurate community profiling of large-scale metabarcoding datasets', *eLife*, 13, p. e85794. Available at: https://doi.org/10.7554/ELIFE.85794

Platt, T., Subba Rao, D. v. and Irwin, B. (1983). 'Photosynthesis of picoplankton in the oligotrophic ocean', *Nature*, 301, pp. 702–704. Available at: https://doi.org/10.1038/301702a0

Porter, K. G. *et al*. (1985). 'Protozoa in planktonic food webs', *The Journal of Protozoology*, 32(3), pp. 409–415. Available at: https://doi.org/10.1111/j.1550-7408.1985.tb04036.x

Posa, M. R. C., Wijedasa, L. S. and Corlett, R. T. (2011). 'Biodiversity and conservation of tropical peat swamp forests', *BioScience*, 61(1), pp. 49–57. Available at: https://doi.org/10.1525/bio.2011.61.1.10

Potila, H. and Sarjala, T. (2004). 'Seasonal fluctuation in microbial biomass and activity along a natural nitrogen gradient in a drained peatland', *Soil Biology and Biochemistry*, 36(7), pp. 1047–1055. Available at: https://doi.org/10.1016/J.SOILBIO.2004.02.014

Pouliot,R. *et al*. (2011). 'Initiation of *Sphagnum* moss hummocks in bogs and the presence of vascular plants: Is there a link?', *Acta Oecologica*, 37(4), pp. 346-354. Available at: https://doi.org/10.1016/j.actao.2011.04.001

Prescott, G.W. (1948). "Desmids," *Botanical Review*, 14(10), pp. 644–676. Available at: https://doi.org/10.1007/BF02861551

Pressler, Y., Moore, J. and Francesca Cotrufo, M. (2019). 'Belowground community responses to fire: meta-analysis reveals contrasting responses of soil microorganisms and mesofauna', *Oikos*, 128(3), pp. 309-327. Available at: https://doi.org/10.1111/oik.05738

Rajter Ľ. *et al*. (2021). 'Colpodean ciliate phylogeny and reference alignments for phylogenetic placements', *European Journal of Protistology*, 77, p. 125747. Available at: https://doi.org/10.1016/j.ejop.2020.125747

Rajter, Ľ. and Dunthorn, M. (2021). 'Ciliate SSU-rDNA reference alignments and trees for phylogenetic placements of metabarcoding data', *Metabarcoding and Metagenomics,* 5, p. e69602. Available at: https://doi.org/10.3897/mbmg.5.69602

Richardson, S. J. and Smith, J. (1977). 'Peat wastage in the East Anglian Fens', *Journal of Soil Science*, 28(3), pp. 485–489. Available at: https://doi.org/10.1111/j.1365-2389.1977.tb02256.x

Rimet, F. *et al*. (2018). *Diat.barcode, an open-access barcode library for diatoms*, V11.1 Recherche Data Gouv. [Online] Available at: https://doi.org/10.15454/TOMBYZ [Accessed 2 December 2023].

Ritz, C. *et al*. (2015). 'Dose-Response Analysis Using R', *PLOS ONE*, 10, p. e0146021. Available at: https://doi.org/10.1371/journal.pone.0146021

Robledo-Arnuncio, J. J. (2011). 'Wind pollination over mesoscale distances: an investigation with Scots pine', *New Phytologist*, 190(1), pp. 222-233. Available at: https://doi.org/10.1111/j.1469-8137.2010.03588.x

Rognes, T. *et al*. (2016). 'VSEARCH: a versatile open source tool for metagenomics,' *PeerJ*, 18, p. e2584. Available at: https://doi.org/10.7717/peerj.2584

Rønn, R., Vestergård, M. and Ekelund, F. (2012). 'Interactions between bacteria, protozoa and nematodes in soil', *Acta Protozoologica*, 51(3), pp. 223–235. Available at: https://doi.org/10.4467/16890027AP.12.018.0764

Rose, J. P., and Dassler, C. L. (2017). 'Spore production and dispersal in two temperate fern species, with an overview of the evolution of spore production in ferns', *American Fern Journal*, *107*(3), pp. 136–155. Available at: https://doi.org/10.1640/0002-8444-107.3.136

Roulet, N. T. *et al.* (2007). 'Contemporary carbon balance and late Holocene carbon accumulation in a northern peatland', *Global Change Biology*, 13(2), pp. 397–411. Available at: https://doi.org/10.1111/j.1365-2486.2006.01292.x

Rueden, C. T. *et al*. (2017). 'ImageJ2: ImageJ for the next generation of scientific image data', *BMC Bioinformatics*, *18*(1), article no. 529 [no pagination]. Available at: https://doi.org/10.1186/s12859-017-1934-z

Rydin, H. and Jeglum, J. K. (2006). *The biology of peatlands*. 1st edn. Oxford: Oxford University Press.

Rydin, H., Gunnarsson, U. and Sundberg, S. (2006). 'The Role of *Sphagnum* in Peatland Development and Persistence'. In Wieder, R.K. and Vitt, D.H. (Eds). *Boreal Peatland Ecosystems*. Ecological Studies, vol 188. Berlin, Heidelberg: Springer, pp. 47-65. Available at: https://doi.org/10.1007/978-3-540-31913-9_4

San Millán, R. M. *et al*. (2013**)**. 'Online exercise for the design and simulation of PCR and PCR-RFLP experiments', *BMC Research Notes,* 6, article no. 513 [no pagination]. Available at: https://doi.org/10.1186/1756-0500-6-513

Schindelin, J. *et al*. (2012). 'Fiji: an open-source platform for biological-image analysis', *Nature Methods*, *9*(7), pp. 676–682. Available at: https://doi.org/10.1038/nmeth.2019

Schipperges, B. and Rydin, H. (1998). 'Response of photosynthesis of *Sphagnum* species from contrasting microhabitats to tissue water content and repeated desiccation,' *New Phytologist*, 140(4), pp. 677–684. Available at: https://doi.org/10.1046/j.1469-8137.1998.00311.x

Schmidt, O., Dyckmans, J. and Schrader, S. (2016). 'Photoautotrophic microorganisms as a carbon source for temperate soil invertebrates', *Biology Letters*, 12(1), p. 20150646. Available at: https://doi.org/10.1098/rsbl.2015.0646

Schönborn, W. (1965). Untersuchungen über die Zoochlorellen-Symbiose der Hochmoor-Testaceen. Limnologica (Berlin), 3, pp. 173-176.

Schrader, C. *et al*. (2012). 'PCR inhibitors - occurrence, properties and removal', *Journal of Applied Microbiology*, 113(5), pp. 1014–1026. Available at: https://doi.org/10.1111/J.1365-2672.2012.05384.X

Schwalm, M. and Zeitz, J. (2015). 'Concentrations of dissolved organic carbon in peat soils as influenced by land use and site characteristics — A lysimeter study', Catena, 127, pp. 72-79. Available at: https://doi.org/10.1016/j.catena.2014.12.007

Seppey, C. V. W. *et al*. (2017). 'Distribution patterns of soil microbial eukaryotes suggests widespread algivory by phagotrophic protists as an alternative pathway for nutrient cycling', *Soil Biology and Biochemistry*, 112, pp. 68–76. Available at: https://doi.org/10.1016/j.soilbio.2017.05.002

Servín-Garcidueñas, L. E. and Martínez-Romero, E. (2012). 'Complete mitochondrial and plastid genomes of the green microalga Trebouxiophyceae sp. strain MX-AZ01 isolated from a highly acidic geothermal lake', *Eukaryotic cell*, 11(11), pp. 1417–1418. Available at: https://doi.org/10.1128/EC.00244-12

Seymour, M. *et al*. (2020). 'Executing multi-taxa eDNA ecological assessment via traditional metrics and interactive networks', *Science of The Total Environment*, 729, p. 138801. Available at: https://doi.org/10.1016/j.scitotenv.2020.138801

Sheath, R. G. and Wehr, J. D. (2003). 1 -Introduction to freshwater algae. In Wehr, J. D. and Sheath, R. G. (Eds). *Aquatic Ecology, Freshwater Algae of North America*. Cambridge, Massachusetts: Academic Press, pp. 1–9. Available at: https://doi.org/10.1016/B978-012741550-5/50002-7

Sheppard, L. J. *et al*. (2004). 'An automated wet deposition system to compare the effects of reduced and oxidised N on ombrotrophic bog species: Practical

considerations', *Water, Air, and Soil Pollution: Focus*, 4(6), pp. 197–205. Available at: https://doi.org/10.1007/S11267-004-3030-4

Sheppard, L. J. *et al*. (2014). 'Inertia in an ombrotrophic bog ecosystem in response to 9 years' realistic perturbation by wet deposition of nitrogen, separated by form', *Global Change Biology*, 20(2), pp. 566–580, Available at: https://doi.org /10.1111/gcb.12357

Sherr, B. F., Sherr, E. B. and Berman, T. (1983). 'Grazing, growth, and ammonium excretion rates of a heterotrophic microflagellate fed with four species of bacteria', *Applied and Environmental Microbiology*, 45(4), pp. 1196-201. Available at: https://doi.org/10.1128/aem.45.4.1196-1201.1983

Shotter, L. and Wentworth, J. (2019). 'Climate Change and UK Wildfire.' POSTnote. [Online] Available at: https://doi.org/10.58248/PN603

Simons, J. *et al.* (1994). 'Restoration of Botshol (The Netherlands) by reduction of external nutrient load: recovery of a characean community, dominated by *Chara connivens*', *Hydrobiologia,* 275, pp. 243–253. Available at: https://doi-org.libproxy.york.ac.uk/10.1007/BF00026715

Smith, H. G., Bobrov, A. and Lara, E. (2007). Diversity and biogeography of testate amoebae. In Foissner, W. and Hawksworth, D. L. (Eds). *Protist Diversity and Geographical Distribution. Topics in Biodiversity and Conservation vol. 8*. Dordrecht: Springer, pp. 95-109. Available at: https://doi.org/10.1007/978-90-481-2801-3_8

Söderström, L., and Jonsson, B. G. (1989). 'Spatial pattern and dispersal in the leafy hepatic *Ptilidium pulcherrimum*', *Journal of Bryology*, 15(4), pp. 793–802. Available at: https://doi.org/10.1179/jbr.1989.15.4.793

Sotiaux, A., Sotiaux, O., and Vanderpoorten, A. (2006). 'Discovery of the rare liverwort *Haplomitrium hookeri* in Belgium: relictualism or long-distance dispersal?', *Cryptogamie, Bryologie,* 27(3), pp. 367-373. Available at: https://sciencepress.mnhn.fr/en/periodiques/bryologie/27/3/discovery-rare-liverwort-haplomitrium-hookeri-belgium-relictualism-or-long-distance-dispersal

Soudzilovskaia, N. A., van Bodegom, P. M. and Cornelissen, J. H. C. (2013). 'Dominant bryophyte control over high-latitude soil temperature fluctuations predicted by heat transfer traits, field moisture regime and laws of thermal insulation', *Functional Ecology*, 27(6), pp. 1442–1454. Available at: https://doi.org/10.1111/1365-2435.12127

Stevens, P. F. (2017). *Angiosperm phylogeny website* Version 14 [Online] Last updated: 19 June 2025. Available at: http://www.mobot.org/MOBOT/research/APweb/ [Accessed 25 August 2025].

Stoeck, T. *et al*. (2010). 'Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water', *Molecular Ecology*, 19(s1), pp. 21-31. Available at: https://doi.org/10.1111/j.1365-294X.2009.04480.x

Stoecker, D. K. *et al.* (2017). 'Mixotrophy in the Marine Plankton', *Annual Review of Marine Science*, 9, pp. 311–335. Available at: https://doi.org/10.1146/annurev-marine-010816-060617

Stoneburner, A., Lane, D. M., and Anderson, L. E. (1992). 'Spore Dispersal Distances in *Atrichum angustatum* (Polytrichaceae)', *The Bryologist*, *95*(3), pp. 324–328. Available at: https://doi.org/10.2307/3243491

Stovall, A.E.L., Shugart, H. and Yang, X. (2019). 'Tree height explains mortality risk during an intense drought', *Nature Communications*, 10, p. 4385. Available at: https://doi.org/10.1038/s41467-019-12380-6

Swindles, G. T. *et al.* (2016). 'Palaeoecology of testate amoebae in a tropical peatland', *European Journal of Protistology*, 55(B), pp. 181–189. Available at: https://doi.org/10.1016/J.EJOP.2015.10.002

Thompson, D. K., Simpson, B. N. and Beaudoin, A. (2016). 'Using forest structure to predict the distribution of treed boreal peatlands in Canada', *Forest Ecology and Management*, 372, pp. 19–27. Available at: https://doi.org/10.1016/J.FORECO.2016.03.056

Thompson, D.K. and Waddington, J.M. (2013). 'Wildfire effects on vadose zone hydrology in forested boreal peatland microforms', *Journal of Hydrology*, 486, pp. 48–56. Available at: https://doi.org/10.1016/J.JHYDROL.2013.01.014

Turetsky, M. R. *et al.* (2010). 'Recent acceleration of biomass burning and carbon losses in Alaskan forests and peatlands', *Nature Geoscience*, 4(1), pp. 27–31. Available at: https://doi.org/10.1038/ngeo1027

Turetsky, M. R. *et al.* (2015). 'Global vulnerability of peatlands to fire and carbon loss', *Nature Geoscience*, 8(1), pp. 11–14. Available at: https://doi.org/10.1038/ngeo2325

Valenzuela-González, F. *et al.* (2016). 'Studying long 16S rDNA sequences with ultrafast-metagenomic sequence classification using exact alignments (Kraken)', *Journal of Microbiological Methods*, 122, pp. 38–42. Available at: https://doi.org/10.1016/J.MIMET.2016.01.011

van Bellen, S. *et al.* (2014). 'Testate amoebae as a proxy for reconstructing Holocene water table dynamics in southern Patagonian peat bogs', *Journal of Quarternary Science*, 29(5), pp. 463–474. Available at: https://doi.org/10.1002/jqs.2719

van de Koot, W. Q. M. *et al.* (2024). 'Variation in water-holding capacity in *Sphagnum* species depends on both plant and colony structure,' *Plants*, 13(8), p. 1061. Available at: https://doi.org/10.3390/PLANTS13081061

van der Heijden, E., Verbeek, S. K., and Kuiper, P. J. C. (2000). 'Elevated atmospheric $CO_2$ and increased nitrogen deposition: Effects on C and N metabolism and growth of the peat moss *Sphagnum recurvum* P. Beauv. var. *mucronatum* (Russ.) Warnst', *Global Change Biology*, 6(2), pp. 201-212. Available at: https://doi.org/10.1046/j.1365-2486.2000.00303.x

Vasselon, V. *et al*. (2017). 'Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France)', *Ecological Indicators*, 82, pp. 1-12. Available at: https://doi.org/10.1016/j.ecolind.2017.06.024

Wanner, M. and Elmer, M. (2009). '"Hot spots" on a new soil surface - How do testate amoebae settle down?', *Acta Protozoologica*, 48, pp. 281–289. Available at: https://www.researchgate.net/publication/228485045

Weber, C. *et al*. (2014). 'Bacterial composition of soils in ponderosa pine and mixed conifer forests exposed to different wildfire burn severity', *Soil Biology and Biochemistry*, 69, pp. 242-250. Available at: https://doi.org/10.1016/j.soilbio.2013.11.010

Wei, T. and Simko, V. (2021). R package 'corrplot': Visualization of a Correlation Matrix (Version 0.92). [Software]. Available at:https://github.com/taiyun/corrplot

Wget (version 1.19.5) [Software]. Available at: https://www.gnu.org/software/wget/

Whitman, T. *et al*. (2019). 'Soil bacterial and fungal response to wildfires in the Canadian boreal forest across a burn severity gradient', *Soil Biology and Biochemistry*, 138, p. 107571. Available at: https://doi.org/10.1016/j.soilbio.2019.107571

Whitton, B. A. *et al.* (2004). Ecological aspects of phosphatase activity in cyanobacteria, eukaryotic algae and bryophytes. In Turner, B., Frossard, E. and Baldwin, D. (Eds). *Organic Phosphorus in the Environment*. Wallingford: CABI publishing, pp. 205–242. Available at: https://doi.org/10.1079/9780851998220.0205

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. Available at: https://ggplot2.tidyverse.org.

Wilkinson, S.L. *et al.* (2023). 'Wildfire and degradation accelerate northern peatland carbon release', *Nature Climate Change*, 13(5), pp. 456–461. Available at: https://doi.org/10.1038/s41558-023-01657-w

Wilkinson, D. M., Creevy, A. and Valentine, J. (2012). 'The past, present and future of soil protist ecology - Introduction to special issue', *Acta Protozoologica*, 51(3), pp. 189–199. Available at: https://doi.org/10.4467/16890027AP.12.015.0761

Williams, P. J. (1981). 'Incorpooration of microheterotrophic processes into the classical paradigm of the planktonic food web', *Kieler Meeresforschungen - Sonderheft*, 5, pp. 1-28. Available at: https://oceanrep.geomar.de/id/eprint/56172

Wilson, L. *et al*. (2011). 'Ditch blocking, water chemistry and organic carbon flux: Evidence that blanket bog restoration reduces erosion and fluvial carbon loss', *Science of The Total Environment*, 409(11), pp. 2010-2018. Available at: https://doi.org/10.1016/j.scitotenv.2011.02.036

Wood, D. E., Lu, J. and Langmead, B. (2019.) 'Improved metagenomic analysis with Kraken 2', *Genome Biology*, 20(1), pp. 1–13. Available at: https://doi.org/10.1186/S13059-019-1891-0

Wood, D. E. and Salzberg, S. L. (2014). 'Kraken: Ultrafast metagenomic sequence classification using exact alignments', *Genome Biology*, 15(3), pp. 1–12. Available at: https://doi.org/10.1186/GB-2014-15-3-R46

Worden, A. Z. *et al.* (2015). 'Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes', *Science*, 347(6223), p. 1257594. Available at: https://doi.org/10.1126/science.1257594

Worrall, F. *et al*. (2011). 'Carbon fluxes from eroding peatlands – the carbon benefit of revegetation following wildfire', *Earth Surface Processes and Landforms,* 36(11), pp. 1487-1498. Available at: https://doi.org/10.1002/esp.2174

Wyatt, K. H. and Turetsky, M. R. (2015). 'Algae alleviate carbon limitation of heterotrophic bacteria in a boreal peatland', *Journal of Ecology*, 103(5), pp. 1165–1171. Available at: https://doi.org/10.1111/1365-2745.12455

Xu, J. (2018). *The role of peatlands in global and regional drinking water resources*. University of Leeds. PhD.

Yang, Z. *et al.* (2023). 'Phylotranscriptomics unveil a Paleoproterozoic-Mesoproterozoic origin and deep relationships of the Viridiplantae', *Nature Communications,* 14, article no. 5542 [no pagination]. Available at: https://doi.org/10.1038/s41467-023-41137-5

Ye, J. *et al*. (2012). 'Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction', *BMC Bioinformatics*, 13, article no. 134 [no pagination]. Available at: https://doi.org/10.1186/1471-2105-13-134

Yin, T. *et al* (2022). 'Biological Nitrogen Fixation and Nitrogen Accumulation in Peatlands', *Frontiers in Earth Science*, 10, article no. 670867 [no pagination]. Available at: https://doi.org/10.3389/feart.2022.670867

Yool, A. and Tyrrell, T. (2003). 'Role of diatoms in regulating the ocean's silicon cycle', *Global Biogeochemical Cycles*, 17(4), p. 1103. Available at: https://doi.org/10.1029/2002GB002018

Yoon, H. S. *et al*. (2004). 'A molecular timeline for the origin of photosynthetic eukaryotes', *Molecular Biology and Evolution*, 21(5), pp. 809–818. Available at: https://doi.org/10.1093/molbev/msh075

Yu, Z. *et al*. (2010). 'Global peatland dynamics since the Last Glacial Maximum', *Geophysical Research Letters*, 37(13), p. L13402. Available at: https://doi.org/10.1029/2010GL043584

Maps throughout this thesis were created using ArcGIS® software by Esri. ArcGIS® and ArcMap™ are the intellectual property of Esri and are used herein under license. Copyright © Esri. All rights reserved. For more information about Esri® software, please visit www.esri.com.