

UNIVERSITY OF OSLO

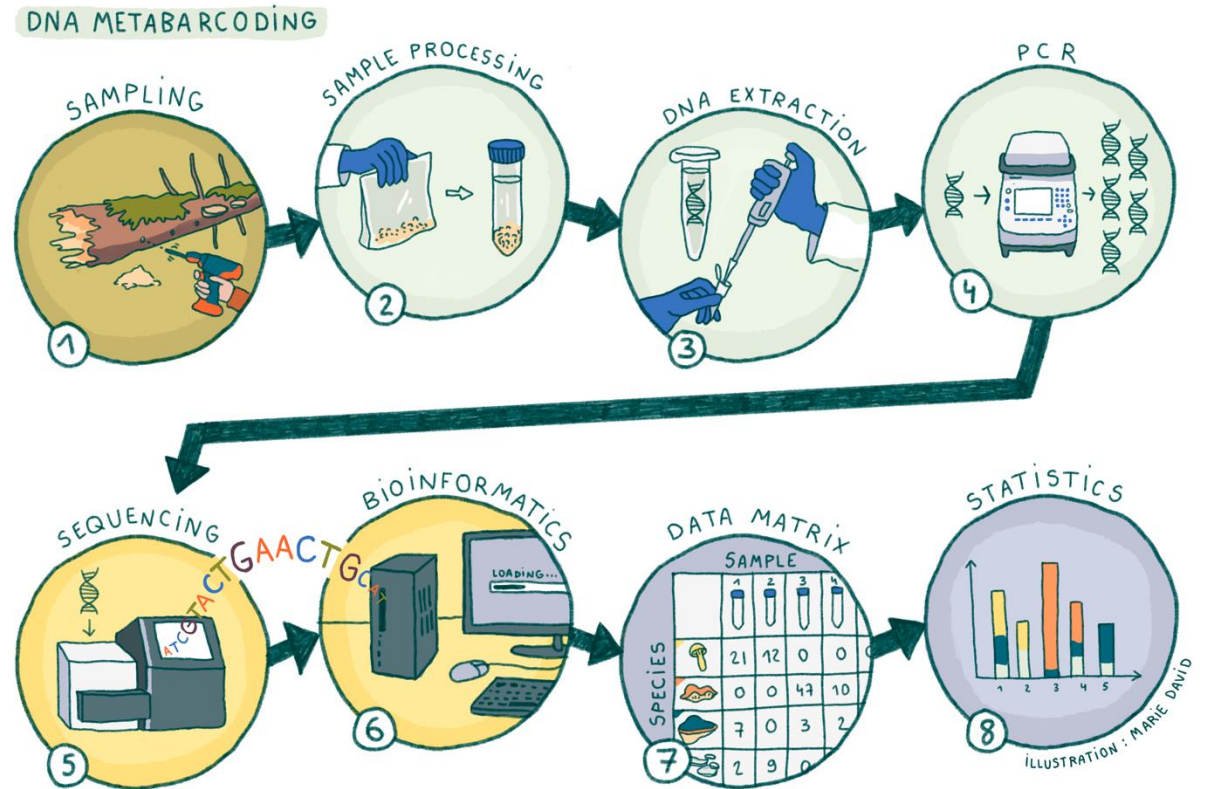
Introduction to long-read metabarcoding

BIO9905MERG1 Spring 2025

Embla Stokke, UiO

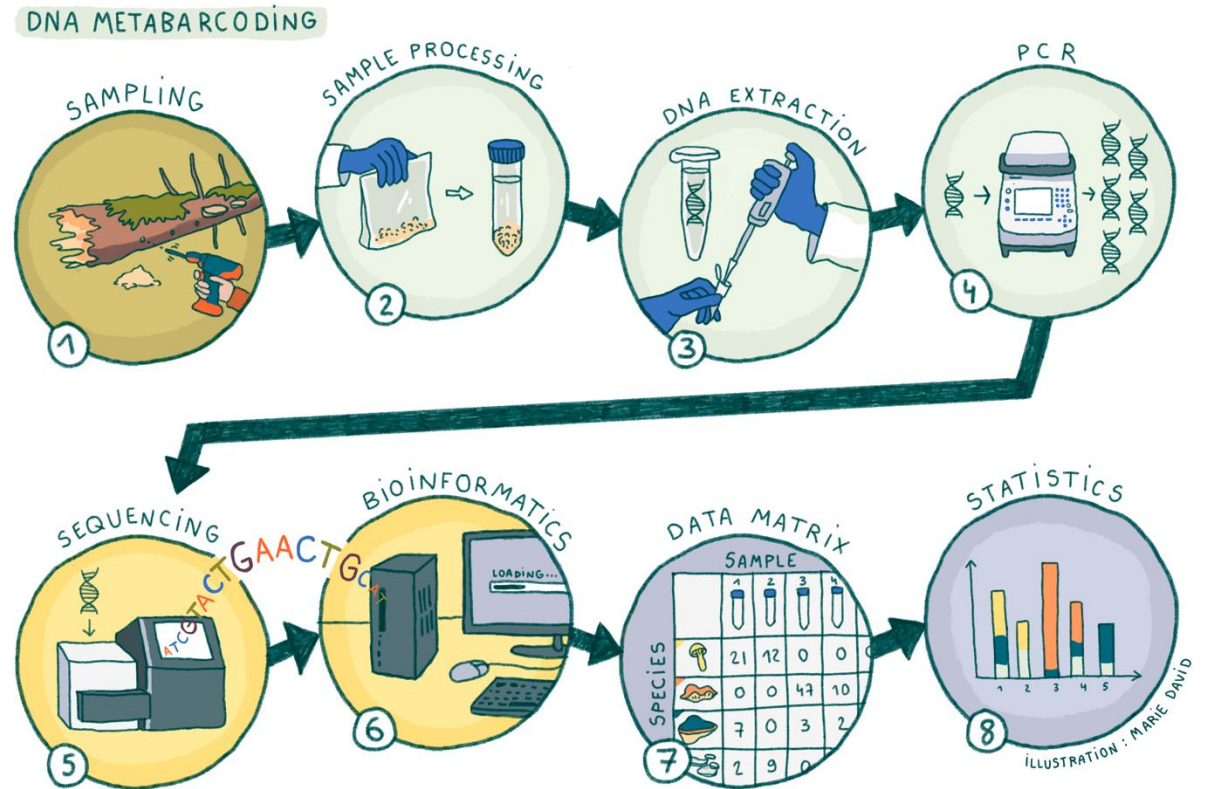
emblaes@student.matnat.uio.no

9/4-2025



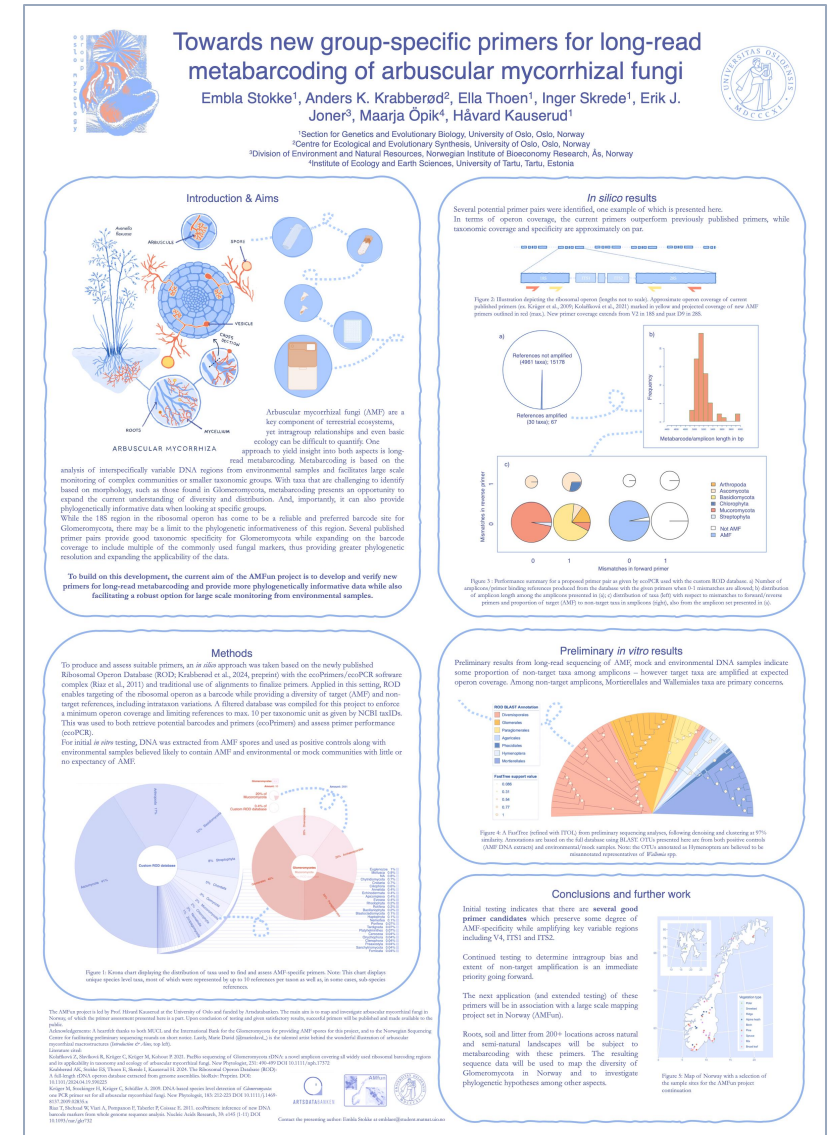
Topics

- Long read metabarcoding versus short read/regular metabarcoding
- Advantages of LRM
- Challenges with LRM
- Long read metabarcoding of soil eukaryotes



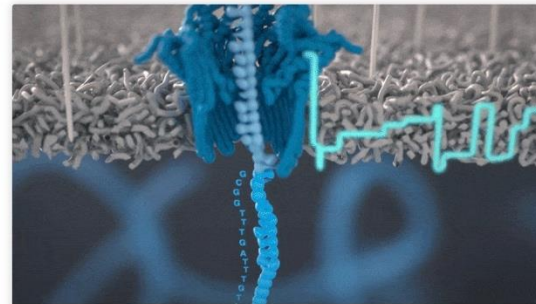
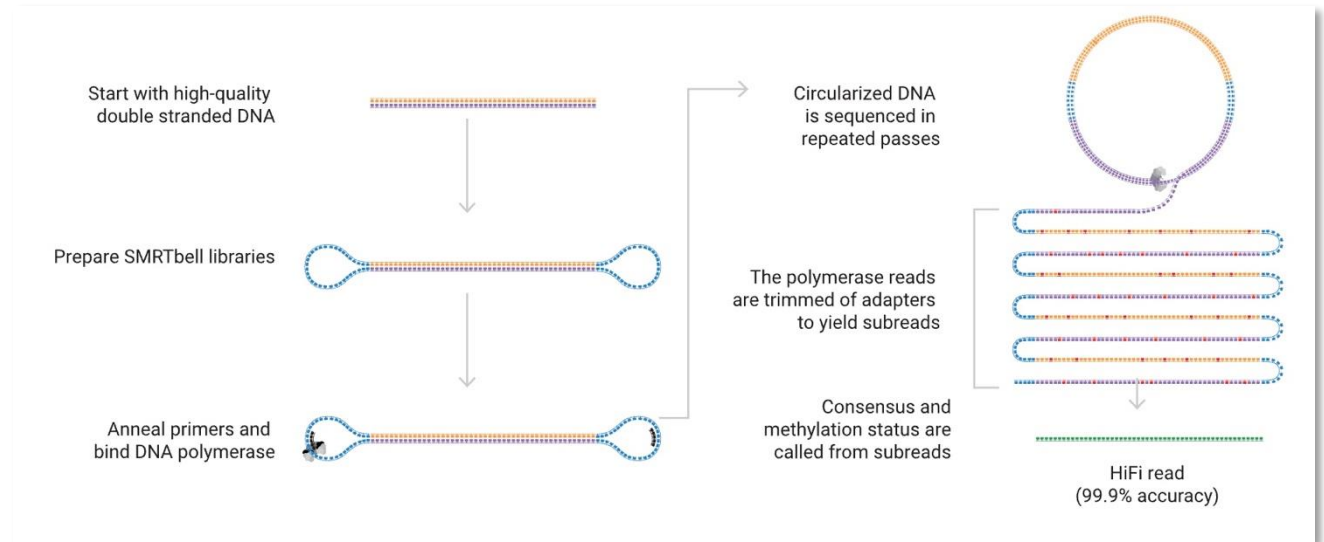
Background

- MSc in Biodiversity & Systematics (NABiS/UiO)
- MSc project with OMG (Håvard Kauserud) – Long read metabarcoding of soil Eukaryotes
- PhD (ongoing) with OMG – Long read metabarcoding of Arbuscular Mycorrhizal Fungi



Long read vs short read

- Generally > 500 bp, 20 000 bp <
- Typically: singular marker vs multimarker
- Pacific Biosciences (PacBio)
- Oxford Nanopore
- Both provide
 - Continuous reads
 - Low sequencing error rate*
 - High yield



You can think of the current as water flowing through a pipe. When an object enters the pipe, the flow of water is disrupted, just as DNA disrupts the current as it passes through the nanopore.

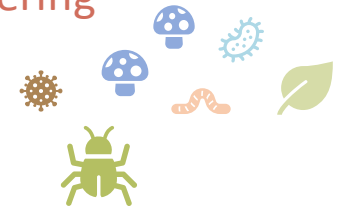
Advantages of LRM

- Longer reads
 - Better lineage resolution
- Consistent quality
- Less noise*
- Cross-database referencing
- Increased multiplex specificity with tagged libraries*

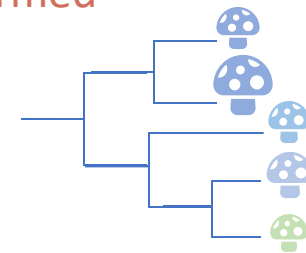
Taxa with tricky barcode gaps



Broad diversity of taxa with differing barcode optima



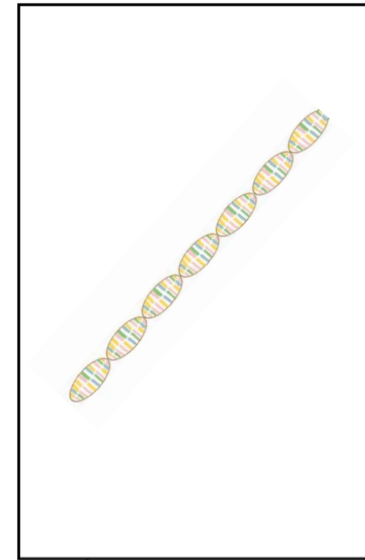
Phylogenetically informed studies



Challenges with LRM

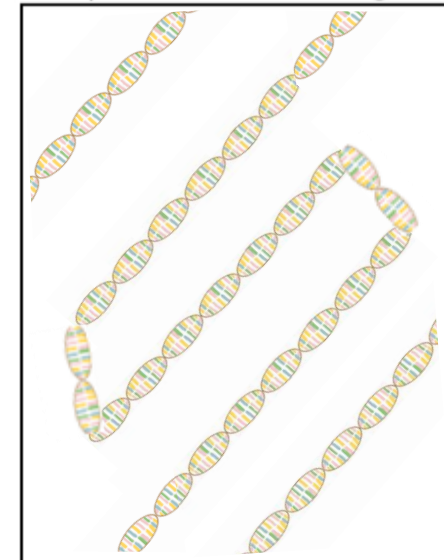
- Not just longer «short» amplicons
- Lab protocols
 - DNA extraction
 - PCR set-up
- Different artefacts/issues to regular amplicons
- Bioinformatics
 - Modified short-read approaches
 - Clustering/OTU synthesis
- Cost

I require a different approach to a number of aspects in the metabarcoding process and data interpretation



who are you?

~~longer~~
im you but ~~stronger~~



Long read metabarcoding of soil eukaryotes

An example of applied LRM*

Long-read sequencing of soil eukaryote eDNA indicates vegetation as driver of soil community structure

MSc thesis by Embla Stokke

Supervised by Ella Thoen

Co-Supervised by Inger Skrede, Anders K. Krabberød & Håvard Kauserud






Methods


Sample collection

- Archaeorhizomycetes Project
- Norway & Svalbard
- Diverse ecological landscapes
- 90 sample sites

Geographic distribution of sample sites



Variable	Description
<i>SITE</i>	Short-name for site
<i>TYPE</i>	Substrate; Litter or Soil
<i>forest_type</i>	Dominant vegetation or forest type at sample site
<i>N</i>	% Nitrogen i sample
<i>C_N</i>	Carbon-to-Nitrogen ratio, calculated from raw percentages
<i>C</i>	% Carbon in sample
<i>P</i>	% Phosphorus in sample
<i>pH</i>	pH measured in sample
<i>lat</i>	Latitude
<i>long</i>	Longitude
<i>plant_richness</i>	Plant richness per site
<i>MAS</i>	Metres above sea level
<i>bio_1</i>	Annual mean temperature
<i>bio_2</i>	Mean diurnal ranges (mean of monthly (max temp - min temp))
<i>bio_3</i>	Isothermality (BIO2/BIO7)(x100)
<i>bio_4</i>	Temperature Seasonality (standard deviation x 100)
<i>bio_8</i>	Mean Temperature of Wettest Quarter
<i>bio_9</i>	Mean Temperature of Driest Quarter
<i>bio_10</i>	Mean Temperature of Warmest Quarter
<i>bio_11</i>	Mean Temperature of Coldest Quarter
<i>bio_12</i>	Annual Precipitation
<i>bio_15</i>	Precipitation Seasonality (coefficient of Variation)



Coring

Separation and pooling

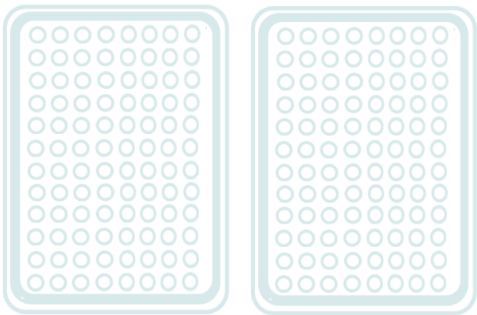
Homogenizing

DNA extraction





Library assembly

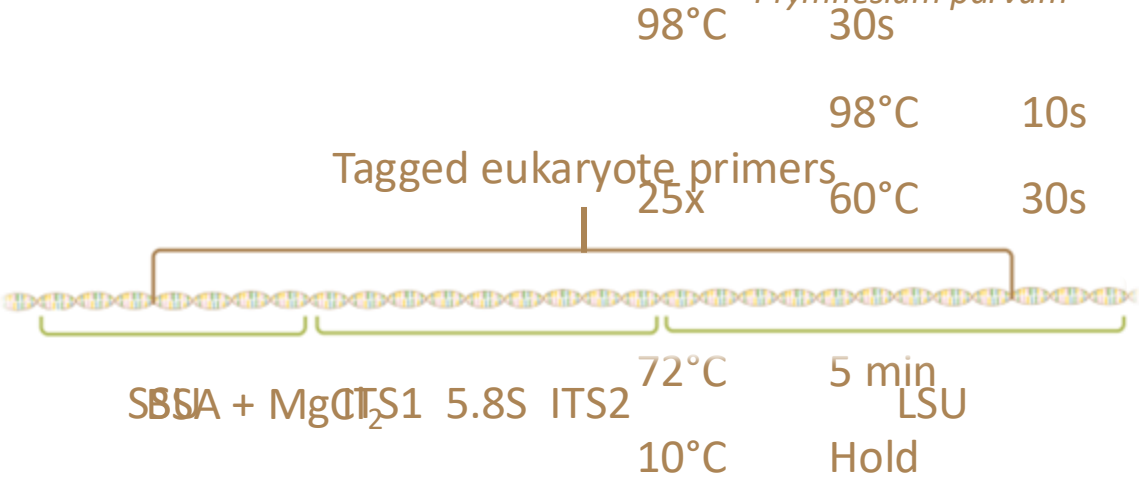


Sequencing

- 180 samples
- 8 replicates
- 2 mock community samples

Serpula similis
Pleurotus ostreatus
Prymnesium parvum

Amplification



*cutadapt – linked
adapter approach*

Demultiplexing

*VSEARCH – denovo
chimera checking with
conservative approach*

Chimera checking

PR²

UNITE

NCBI

Taxonomic annotation

Sample inference
and denoising

*dada2 – filtering,
denoising and sample
inference*

Clustering

*VSEARCH – clustering at
97% similarity*



*Coherence between database
annotations & phylogenetic
groupings*

Phylogenetic
annotation check



Phylogenetic inference

*SILVA concatenated reference
sequences*

MAFFT – sequence alignment

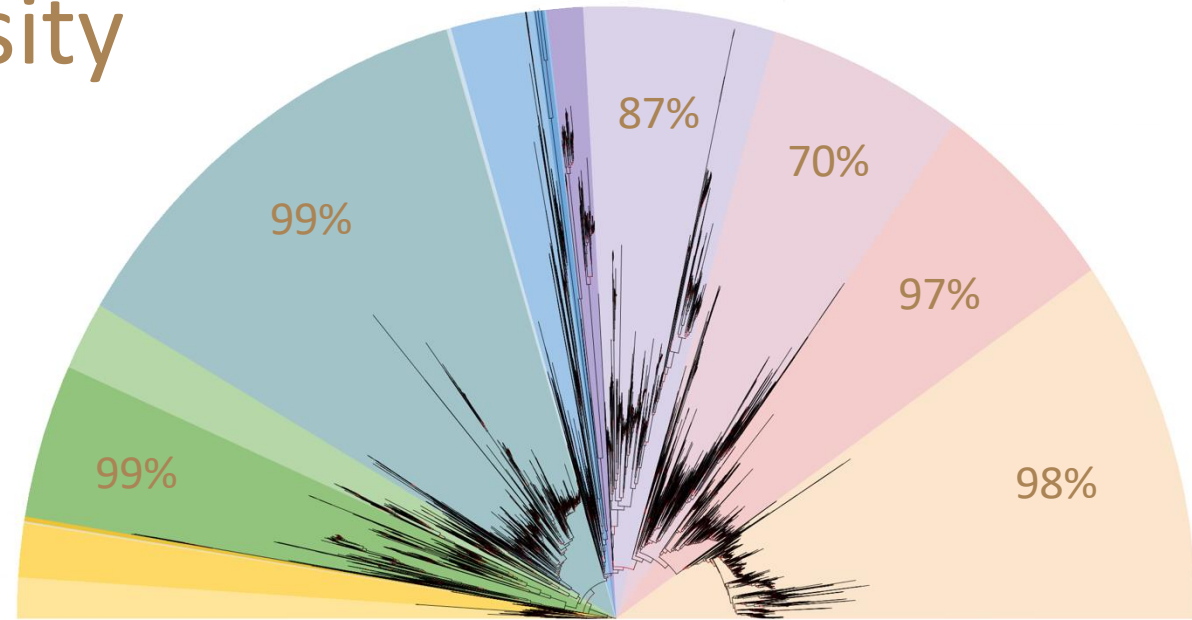
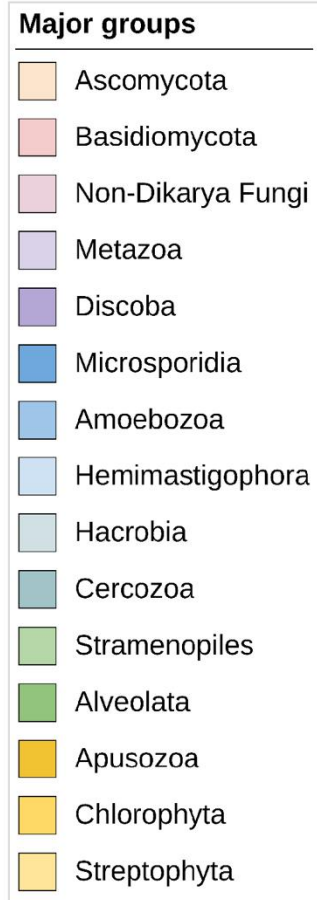
*Geneious - 50% gaps & ITS
removal*

FastTree v.2 - tree building

Statistical analysis

*R v. 4.2.2 – various
packages*

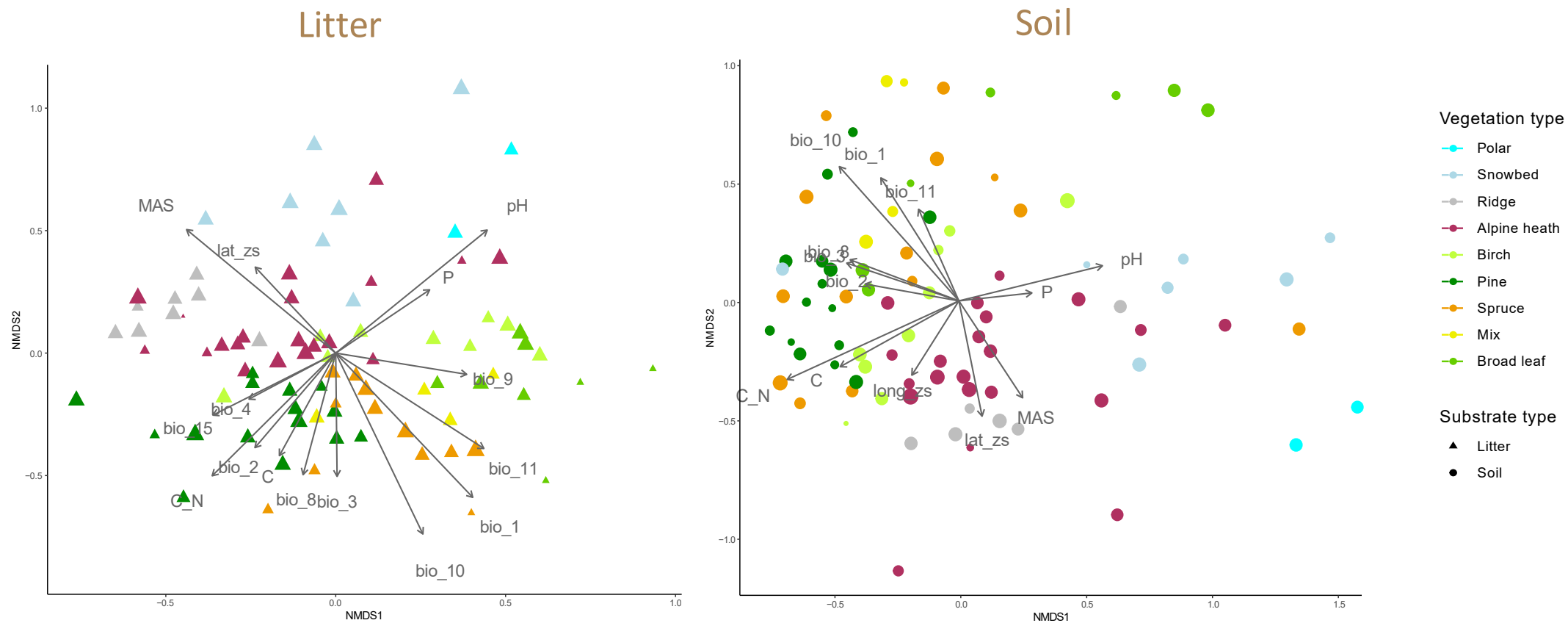
Phylogeny & diversity

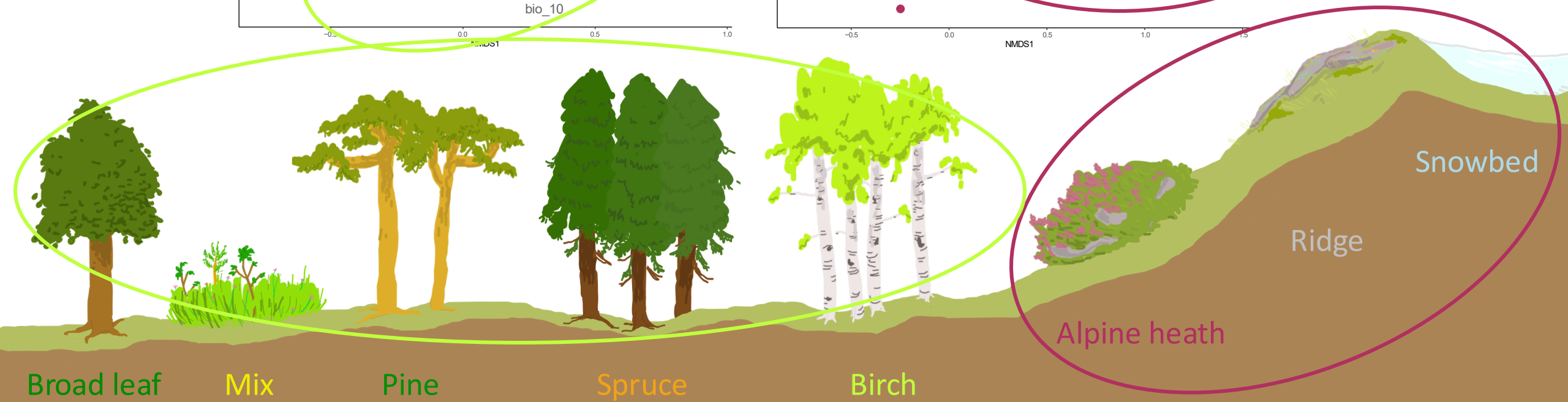
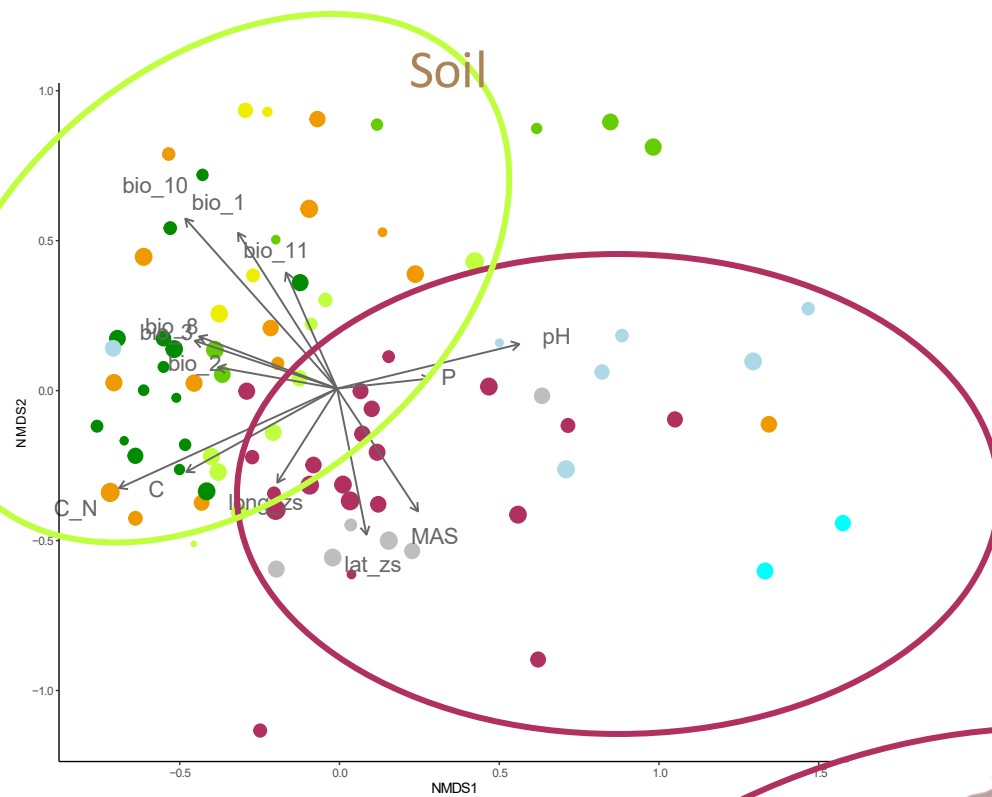
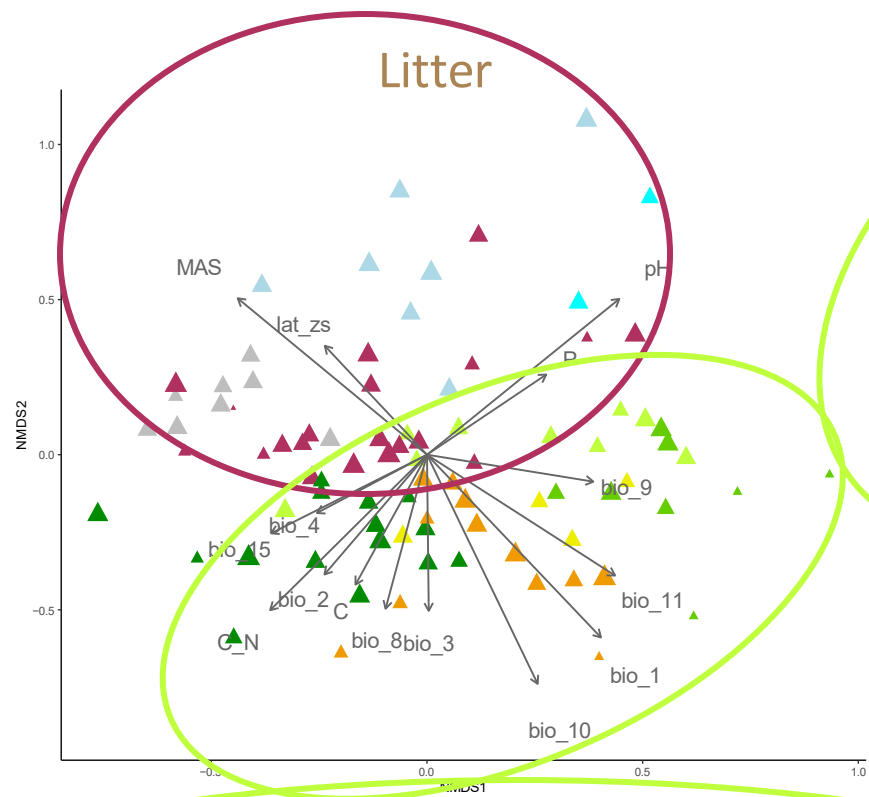


Distribution of relative read abundance



Beta diversity





Thank you for listening!

Questions?

