Bacterial diversity of Northern peatlands

# Abstract

# Introduction

## Why peatlands?

Peatlands represent 3% of global land surfaces but are estimated to store 30% of terrestrial carbon (REF). Usually being carbon sinks, their future role in atmospheric CO2 dynamics and thus climate changes are uncertain (REF).

## What are peatlands?

Most peatlands are found in the northern hemisphere (REF) and are dominated by mosses of the sphagnum genus (REF) which represents most of the photosynthetic activity (REF). However, peatlands are also homes of diverse microbial communities that can fix (e.g. methanotrophs and cyanobacteria) or release (e.g. methanogens) atmospheric carbon (REF).

## Why bacteria?

These microbial communities are both functionally and taxonomically diverse (REF) but remains largely unknown at global scale (naze).

## Why global dataset?

Climate change being a global phenomenon, it is essential to assess peatlands functioning at a transcontinental scale.

## What are the objectives?

We want to identify most abundant and ubiquitous OTUs as well as characterize rare OTUs.

We want to identify the drivers of these (most abundants?) OTUs.

Map this taxo-ecological diversity

Cluster – RVB clusters

# Materials and methods

199 Peat samples were collected from XX peatlands across the Northern Hemisphere (MAP). Samples were shipped to Toulouse for DNA extractions and measurements.

## DNA extraction and sequencing

To characterize bacterial communities, 16s nuclear DNA gene was amplified using the V4-V5 region with the 515F (GTGYCAGCMGCCGCGGT) and 909R (CCCCGYCAATTCMTTTRAGT) primers (Tuan et al. 2014). Forward and reverse primers were tagged for post-sequencing discrimination.

PCR protocol

## Bioinformatics

Sequencing outputs were processed using the FROGs (V4.1) pipeline implemented on a Galaxy server (Galaxy Toulouse INRAE). Parameters details for each step are available in Supplementary (XXX) and only the main steps are briefly described here. Sequences were dereplicated and then clustered using SWARM (Mahé et al. XX) with an aggregation distance of 1 resulting in the equivalent of ASVs. Chimera, aberrant sequences resulting from PCR errors, were removed from the dataset before taxonomic assignation using the 16s SILVA database (release 16S\_SILVA\_138.1) and RDP classifier (REF). The resulting files were exported to .TSV for further processing within R (Version) and RStudio (Version).

Further processing was performed using the MetabaR R package (REF). Chloroplastic and mitochondrial sequences were removed from the dataset as well as sequences poorly assigned at the Kingdom level (less than 70% of RDP bootstraps).

We also removed singletons sequences (i.e. sequences present only once in the dataset) that represented an important diversity but only XX% of the observations (i.e., reads).

## Statistics

# Results

## Bacterial diversity

We found high diversity (~XX ASVs) but also important variation across samples with alphadiversity ranging from XX to XX ASVs per sample. The dominant bacterial orders were *Acetobacterales* (33.9%), *Acidobacteriales* (12.3%), *WD260* (8.7%), *Cyanobacteriales* (4.8%), *Caulobacterales* (4.4%), *Burkholderiales* (3.1%), *Rhizobiales* and *Chitinophagales* (2.5% each). The remaining 24% belonged to XX different orders for a total of XX bacterial orders (DETAILS SUPP XX) while 3.8% of our sequences were not assigned to any known orders. We found XX genera with two particularly dominant across Northern *sphagnum* peatlands with 14.1% and 8.6% of our sequences respectively belonging to *Acidocella* and *Granulicella* while the 3rd genus (*Acidisoma*) represented only 1.6% of our sequences. 53.2% had no assignation at this taxonomic rank suggesting an important unknown diversity.

Weird sites

# Discussion

# References

# Supplementary