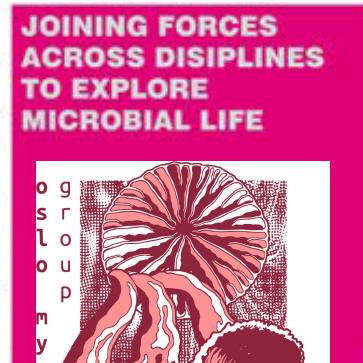


# Case study: metabarcoding microbial communities inhabiting dead wood in boreal forests



Forskningsrådet  
The Research Council of Norway

Sundy Maurice  
Researcher  
Evolution and genetics  
Faculty of Biosciences  
University of Oslo

# Metabarcoding of fungal & bacterial communities from sawdust samples (metasaw)

1. Background
2. Study design & methodological approach
3. Compilation of environmental variables
4. Main results from DNA metabarcoding
5. Perspectives

# Forest practices leading to decline of wood substrates

Natural, protected, old-growth spruce forest with long continuity where dead wood is a characteristic feature



Old-growth forest (Kotinen, Hame)



Managed forest, Southern Finland

- Lack of dead wood resulting in population decline & species extinction

# Fungal & bacterial communities in dead spruce

1. Are there differences in community composition and diversity, for both bacteria and fungi, due to forest management practices and/or landscapes?
2. What are the major variables that correlate with community composition?
3. ~~How are fungi and bacteria distributed along a log?~~

# Study design

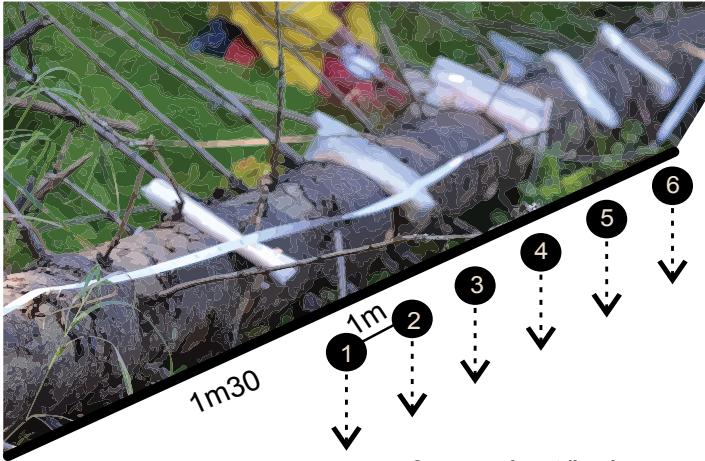
Landscape	Forest type		
	Type I Old-growth occpd	Type II Old-growth unoccpd	Type III Culture
Nordmarka			
Østmarka			
Telemark			

**Type I:** Old-growth forest with large amount of dead wood and trees of large diameters

**Type II:** Old-growth forest with absence of rare polypore species

**Type III:** Cultured forest

# Study design

Forests	Old-growth occupd	Old-growth unoccpd	Culture forest
	20-30 cm DBH Decay stage 2 Absence of 3 focal species fruiting		
5 Logs		 <p>1m30</p> <p>5 cm<sup>3</sup> sawdust/hole</p>	
6 Holes			

Total  $(3 \times 5 \times 6) \times 3$  landscapes = 270 samples,  
PCR replicates, negative PCR and mock

# Sample processing

1. Mixing the sawdust to homogenize
2. Weighing 3g of fresh mass in 50 mL Falcon tubes
3. Freeze drying (30h)
4. Crushing in 50 mL Falcon using 2 \* 6.8mm Zirconium oxide beads (Saveen Werner) using a Fast prep homogenizer (MoBIO). 2\*2 Falcons at a time
5. CTAB & Phenol/Chloroform (40 mL) DNA extraction
6. Cleaning DNA extract with the E.Z.N.A Soil DNA kit [DNA] Qbit average 60 ng/ $\mu$ l
7. Diluting each sample to *ca.* 5 ng/ $\mu$ l



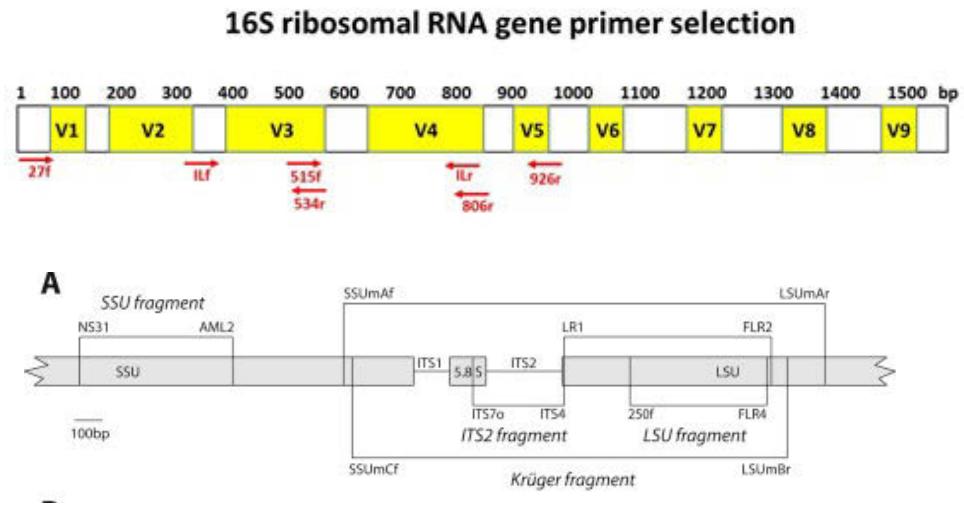
Wood particles of different sizes & not only the sawdust  
Freeze dry 48 hrs, ~1 g lost.

# Sample processing

## Primers used & regions

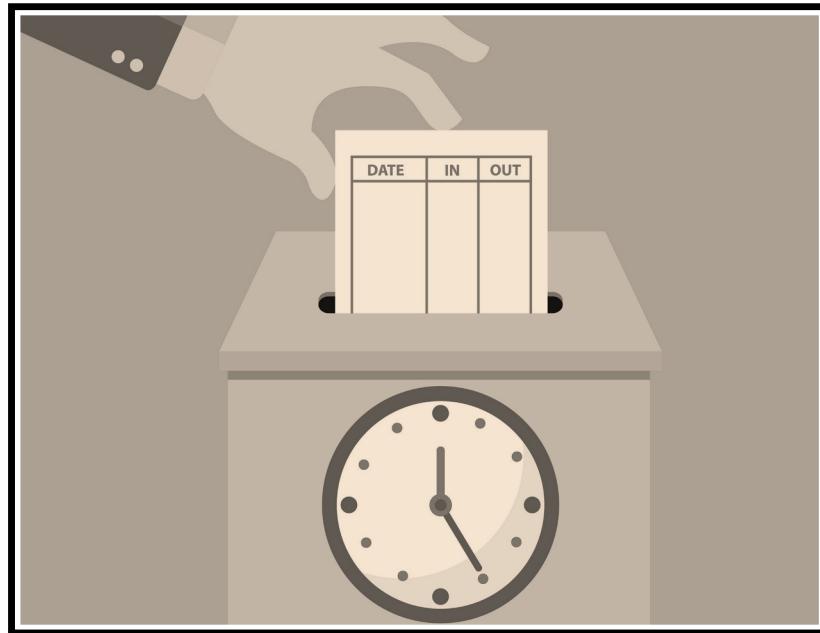
16S: 515F & 806R (240 pb)  
modified by Caporaso et al. 2011

ITS: IT4 and gITS7 (ca. 300pb)  
IT4 White et al 1990  
gITS7 from Ihrmark et al 2012



8. Preparing metabarcoding libraries
9. Normalising PCR product (SequelPrep plates)
10. Pooling \*96 samples & AMPure beads cleaning
11. Sequencing barcoded libraries on 2 Miseq lanes: 16S & ITS

Send for library preparation & sequencing, *ca.* 1 month



What can I do in the meanwhile?

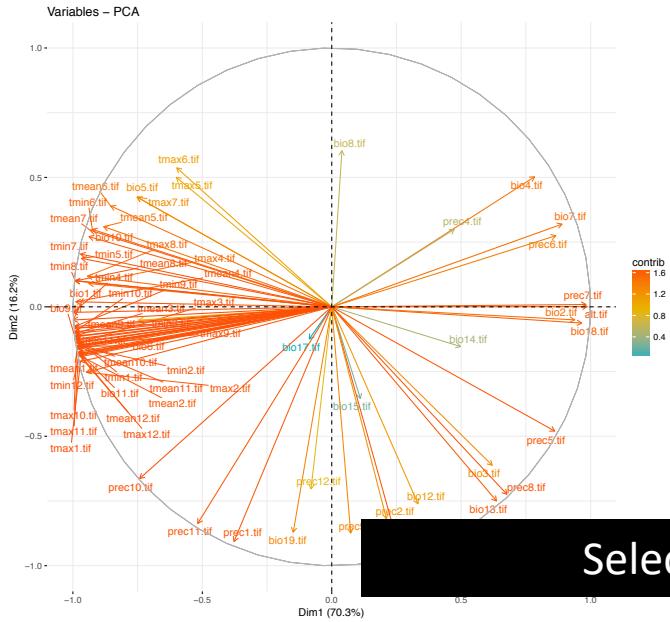
# Take Home tips-1

- ❑ Learn basic Unix command
- ❑ Talk to supervisors/collab. about organization of the analyses
- ❑ Access a cluster (<https://www.metacenter.no>)
- ❑ Take a course in metabarcoding & practice
- ❑ Compile & prepare meta-data

# Selection of environmental variables

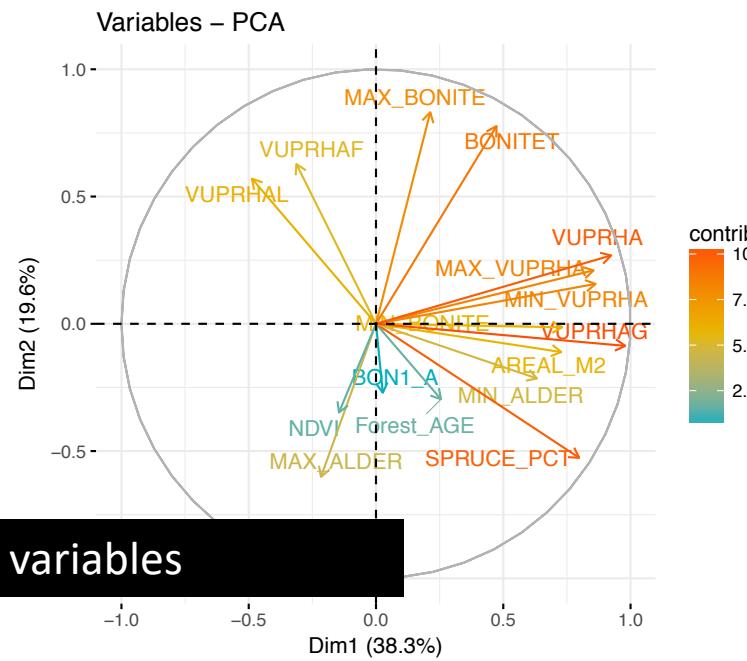
## (1) Climatic variables

68 climatic variables (bioclim)



## (2) Landscaped variables

43 landscape variables (Kilden + Satskog)



Selection of 15 variables

### 3 Selected variables:

Annual mean temperature,  
Annual precipitation,  
Temperature seasonality

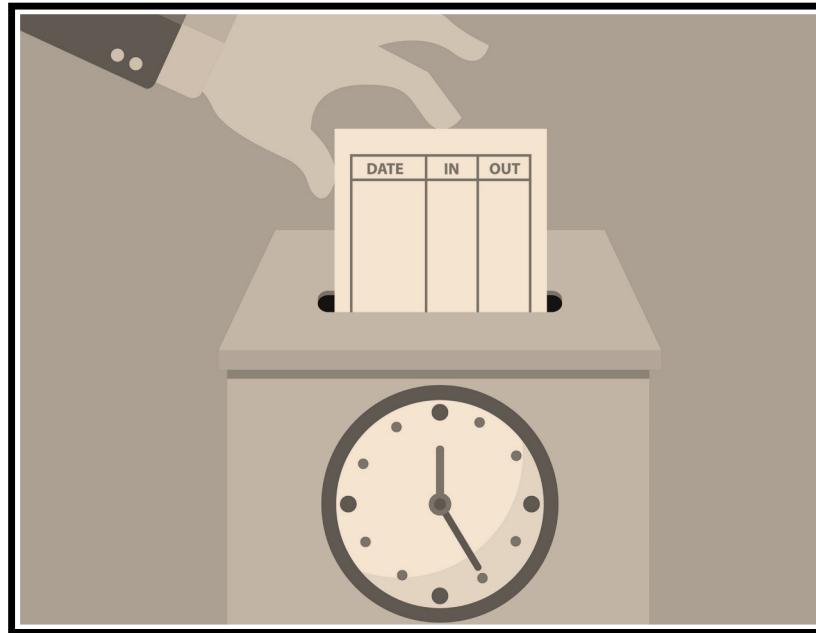
### (3) Field measurements

7 selected variables:  
Decay stage, Ground contact,  
Bark cover, Epiphyte cover, Length, 3-point  
diameter, Moisture content

### 5 Selected variables

Area , Altitude, Forest quality,  
Spruce stand, Dead spruce

*After 1 month .... data arrives*



How do I get started with the data analyses?

# Take Home tips-2

- ❑ Very important! Check quality of sequencing  
e.g. FASTQC
- ❑ Encounter problems with upload/downloading data

```
/projects/clusteringits/rawdata/BaseCallsMaurice_112017  
-bash-4.1$ ls -lah
```

total 14G

drwxrwx--- 3 sandym users 2.0K Jan 29 10:46 .

drwxrwx--- 3 sandym users 2.0K Dec 18 18:37 ..

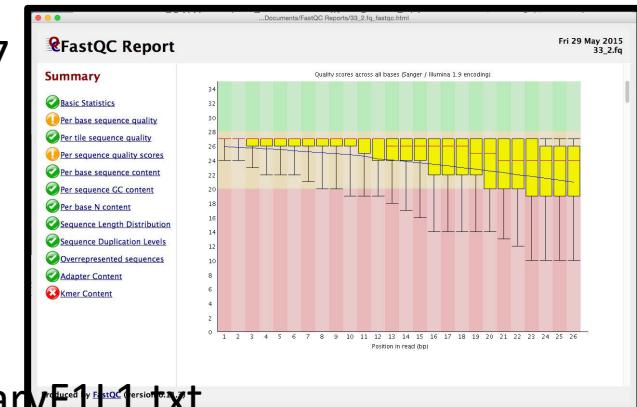
drwxrwx--- 5 sandym users 2.0K Dec 18 18:37 Alignment2

-rwxrwx--- 1 sandym users 2.5K Dec 18 18:37 FastqSummary\_E1-1.txt

-rwxrwx--- 1 sandym users 1.8G Dec 18 18:38 Issa-ITS-lib1\_S1\_L001\_R1\_001.fastq.gz

-rwxrwx--- 1 sandym users 2.0G Dec 18 18:39 Issa-ITS-lib1\_S1\_L001\_R2\_001.fastq.gz

-rwxrwx--- 1 sandym users 4.0G Dec 18 18:40 Issa-ITS-lib2\_S2\_L001\_R1\_001.fastq.gz



- ❑ Insufficient amount of data generated, problem with clustering e.g. 5 M sequences for reverse read (rather than 30-40 M for 16S)

# Bioinformatics workflow (16S & ITS)

## Comparable steps

#	Analysis	Software	Bacteria	Fungi
1	Base correction	BayesHammer ( Spades)	Yes	Yes
3	Merge reads 1 & 2	PEAR	Yes	Yes
4a	Quality filtering	fastx-toolkit	Yes	No
4b	Quality filtering	vsearch	Yes	Yes
5a	Re-orientation of the reads in same direction, 5'->3'	fqgrep	Yes	Yes
5b	Reverse complement the reverse read 3'-5' to 5'-3'	fastx-toolkit	Yes	Yes
6	Demultiplexing	sdm	Yes	Yes
8a	Remove primers both of forward and reverse ends	cutadapt	Yes	Yes
8b	ITS extraction	itsx	No	Yes
9	Dereplication & sort by abundance	vsearch	Yes	Yes
10	Clustering 97 %		Yes	Yes
11	denovo Chimera checking		Yes	Yes
12	Filtering OTU table & reset file with OTUs < 10 reads		Yes	Yes
13a	Generate OTUs table		Yes	Yes
14	Assign taxonomy	Qiime	Yes (NCBI, GG)	Yes (UNITE & Polypore DB)
15	Filter OTU tables to remove singleton otus		Yes	Yes
16	Convert biom to txt		Yes	Yes

Primers:  
gits7 + its4

515F + 806R

# Bioinformatic workflow

Error correction

Bayeshammer (spades)

45.4 M

7.58 M

Merging

Pear

22.6 M

7.18 M

Quality filtering 1

Vsearch -> min 100 bp, min q21

20.3 M

7.12 M

Demultiplexing

sdm (lotus)

20.1 M

5.6 M

ITS extraction + cutadapt

ITSx -> ITS2

19.8 M

-/ 5.4 M

DerePLICATION

Vsearch

6,918,192

1,809,116

Clustering (95 or 97%)

Vsearch

258,388

Quality filtering 2

Vsearch (uchime)

5054  
15 310

Vsearch -> min 10 reads

2,727

OTUs

4,652

Assign taxonomy

qiime

Unite db/GG

2,024

2,387

Post clustering curation

LULU

Blast

FunGuild

1,584

1,802

# Importance of post bioinformatics filtering

Example from a **different** fungal dataset

Filtering steps (post bio-informatics)	# of samples	# of OTUs	# of reads	Remarks
<b>Raw OTU table</b>	192	2188	16,196,863	
<b>LULU</b>	192	1632	16,196,863	rm erroneous OTUs by combining seq. similarity & co-occurrence patterns
<b>Remove replicates and no blast hit OTUs</b>	181	1385	14,871,904	
<b>Tag-switching correction owi_renormalized_10%</b>	181	1385	13,511,207	rm <b>most</b> reads from tag-switching
<b>Remove mock OTUs</b>	181	1369	13,269,587	to rm tag leakage from mock community
<b>Filter samples</b>	176	1367	13,089,836	rm mock, PCR control & poscae12 & rhizaria OTUs
<b>Remove tag-switching coming from host species</b>	176	1367	12,816,168	rm manually the host leakage
<b>Rarefy 10 000</b>	176	1367	1,794,144	sample min 10,194 (Fompin3)

# OTU= Operational Taxonomic Unit Table

OTUs

Sample\_ID

Taxonomic assignment

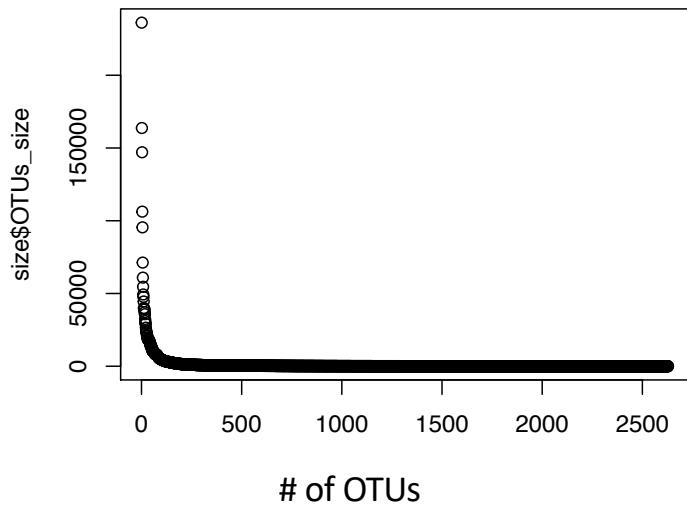
A	taxonomy	S	T
1	OTU_ID k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Boliniiales; f_Boliniaceae; g_ungrouped_Boliniaceae; s_Boliniaceae_sp	TS-No	ITS-Nord
2	OTU_347;size=21 k_Fungi; p_Ascomycota; c_Leotiomycetes; o_Helotiales; f_Hyaloscyphaceae; g_Haplographium; s_Haplographium_sp	4	0
3	OTU_4058;size=3 k_Fungi; p_Ascomycota; c_Lecanoromycetes; o_Lecanorales; f_Parmeliaceae; g_Parmelia; s_Parmelia_serrana	0	0
4	OTU_2547;size=3 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Hypocreales; f_ungrouped_Hypocreales; g_ungrouped_Hypocreales; s_Hypocreales_sp	0	0
5	OTU_870;size=22 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Coniochaetales; f_Coniochaetaceae; g_Lecythophora; s_Lecythophora_sp	0	0
6	OTU_1490;size=7 k_Fungi; p_Basidiomycota; c_Agaricomycetes; o_Polyporales; f_Fomitopsidaceae; g_Antrodia; s_Antrodia_primeva	0	0
7	OTU_11147;size=1 k_Fungi; p_Ascomycota; c_Eurotiomycetes; o_Chaetothyriales; f_ungrouped_Chaetothyriales; g_ungrouped_Chaetothyriales; s_Chaetothyriales_sp	0	0
8	OTU_1700;size=1 k_Fungi; p_Ascomycota; c_Dothideomycetes; o_Capnodiales; f_Mycosphaerellaceae; g_Mycosphaerella; s_Mycosphaerella_sp_Ston1	0	0
9	OTU_2067;size=4 k_Fungi; p_Basidiomycota; c_Tremellomycetes; o_Tremellales; f_Tremellales_family_Incertae_sedis; g_Cryptococcus; s_Cryptococcus_sp	0	0
10	OTU_3177;size=1 k_Fungi; p_Ascomycota; c_Leotiomycetes; o_Helotiales; f_Helotiales_family_Incertae_sedis; g_Hyphodiscus; s_Hyphodiscus_hymenophilus	1	0
11	OTU_1992;size=8 k_Fungi; p_Ascomycota; c_Agaricomycetes; o_Agaricales; f_Mycenaceae; g_ungrouped_Mycenaceae; s_Mycenaceae_sp	0	0
12	OTU_1465;size=8 k_Fungi; p_Basidiomycota; c_Agaricomycetes; o_Agaricales; f_Hyaloscyphaceae; g_Hyaloscypha; s_Hyaloscypha_aureliella	0	0
13	OTU_520;size=78 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Hypocreales; f_ungrouped_Hypocreales; g_ungrouped_Hypocreales; s_Hypocreales_sp	0	0
14	OTU_32;size=947 k_Fungi; p_Ascomycota; c_Leotiomycetes; o_Helotiales; f_ungrouped_Helotiales; g_ungrouped_Helotiales; s_Helotiales_sp	22	0
15	OTU_1500;size=9 k_Fungi; p_Ascomycota; c_Eurotiomycetes; o_Chaetothyriales; f_Chaetothyriaceae; g_Ceramothryium; s_Ceramothryium_carniolicum	0	0
16	OTU_3339;size=1 k_Fungi; p_Basidiomycota; c_Microbotryomycetes; o_Sporidiobolales; f_Sporidiobolales_family_Incertae_sedis; g_ungrouped_Sporidiobolales; s_Sporidiobolales_sp	0	0
17	OTU_4879;size=1 k_Fungi; p_Zygomycota; c_Zygomycota_class_Incertae_sedis; o_Mucorales; f_Mucoraceae; g_Mucor; s_Mucor_silvaticus	0	0
18	OTU_10668;size=1 k_Fungi; p_Ascomycota; c_Leotiomycetes; o_Helotiales; f_Hyaloscyphaceae; g_Hyaloscypha; s_Hyaloscypha_aureliella	0	0
19	OTU_1654;size=6 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Ophiostomatales; f_Ophiostomataceae; g_Grosmannia; s_Grosmannia_cucullata	0	0
20	OTU_3029;size=2 k_Fungi; p_ungrouped_Fungi; c_ungrouped_Fungi; f_ungrouped_Fungi; g_ungrouped_Fungi; s_Fungi_sp	0	0
21	OTU_1381;size=9 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Microascales; f_Ceratostigidaceae; g_Ambrosiella; s_Ambrosiella_sp_2PGP_A1	0	0
22	OTU_2003;size=5 k_Fungi; p_Ascomycota; c_Eurotiomycetes; o_Eurotiales; f_Trichocomaceae; g_Penicillium; s_Penicillium_lividum	0	0
23	OTU_6075;size=5 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Helotiales; f_ungrouped_Pseudeurotiaceae; g_ungrouped_Pseudeurotiaceae; s_Pseudeurotiaceae_sp	0	0
24	OTU_4437;size=1 k_Fungi; p_Ascomycota; c_Dothideomycetes; o_Dothideomycetes_order_Incertae_sedis; f_Pseudeurotiaceae; g_ungrouped_Pseudeurotiaceae; s_Pseudeurotiaceae_sp	0	0
25	OTU_139;size=18 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Microascales; f_Ceratostigidaceae; g_Ambrosiella; s_Ambrosiella_ferruginea	3	0
26	OTU_8680;size=4 k_Fungi; p_Ascomycota; c_Pezizomycetes; o_Pezizales; f_Discomycetidae; g_Gyromitra; s_Gyromitra_esculenta	15	0
27	OTU_16668;size=1 k_Fungi; p_Ascomycota; c_Dothideomycetes; o_Dothideomycetes_order_Incertae_sedis; f_Pseudeurotiaceae; g_ungrouped_Pseudeurotiaceae; s_Pseudeurotiaceae_sp	0	0
28	OTU_4534;size=1 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Lulworthiales; f_Lulworthiaceae; g_Zalerion; s_Zalerion_sp	0	0
29	OTU_1904;size=5 k_Fungi; p_Ascomycota; c_Agaricomycetes; o_Agaricales; f_Typhulaceae; g_ungrouped_Typhulaceae; s_Typhulaceae_sp	0	0
30	OTU_3083;size=2 k_Fungi; p_Ascomycota; c_ungrouped_Ascomycota; o_ungrouped_Ascomycota; f_ungrouped_Ascomycota; g_ungrouped_Ascomycota; s_Ascomycota_sp	0	0
31	OTU_196;size=66 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Diaporthales; f_Gnomoniaceae; g_Gnomonia; s_Gnomonia_cf_ischnostyla_CBS_121908	9	1
32	OTU_2150;size=1 k_Fungi; p_ungrouped_Fungi; c_ungrouped_Fungi; o_ungrouped_Fungi; f_ungrouped_Fungi; g_ungrouped_Fungi; s_Fungi_sp	0	0
33	OTU_1576;size=9 k_Fungi; p_Ascomycota; c_Eurotiomycetes; o_Chaetothyriales; f_ungrouped_Chaetothyriales; g_Chaetothyriales_sp	0	0
34	OTU_6635;size=3 k_Fungi; p_Ascomycota; c_Dothideomycetes; o_Capnodiales; f_Capnodiales_family_Incertae_sedis; g_Capnobotryella; s_Capnobotryella_sp	0	0
35	OTU_2432;size=3 k_Fungi; p_Ascomycota; c_Leotiomycetes; o_Helotiales; f_Hyaloscyphaceae; g_Hyaloscypha; s_Hyaloscypha_albohyalina_var_spiralis	0	0
36	OTU_3797;size=1 k_Fungi; p_Ascomycota; c_Saccharomycetes; o_Saccharomycetales; f_Saccharomycetales_family_Incertae_sedis; g_Candida; s_Candida_ontarioensis	6	0
37	OTU_568;size=74 k_Fungi; p_Ascomycota; c_Agaricomycetes; o_Agaricales; f_ungrouped_Agaricales; g_ungrouped_Agaricales; s_Agaricales_sp	0	0
38	OTU_1747;size=6 k_Fungi; p_Basidiomycota; c_Agaricomycetes; o_Agaricales; f_ungrouped_Agaricales; g_ungrouped_Agaricales; s_Agaricales_sp	0	0
39	OTU_8597;size=1 k_Fungi; p_ungrouped_Fungi; c_ungrouped_Fungi; o_ungrouped_Fungi; f_ungrouped_Fungi; g_ungrouped_Fungi; s_Fungi_sp	0	0
40	OTU_2221;size=3 k_Fungi; p_Basidiomycota; c_Agaricomycetes; o_Cantharellales; f_Botryobasidiaceae; g_ungrouped_Botryobasidiaceae; s_Botryobasidiaceae_sp	0	0
41	OTU_4707;size=1 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Hypocreales; f_Nectriaceae; g_ungrouped_Nectriaceae; s_Nectriaceae_sp	0	0
42	OTU_2064;size=3 k_Fungi; p_Ascomycota; c_Sordariomycetes; o_Hypocreales; f_Nectriaceae; g_ungrouped_Nectriaceae; s_Nectriaceae_sp	0	0

related\_otu\_tax\_no\_singelton

# Number of OTUs obtained

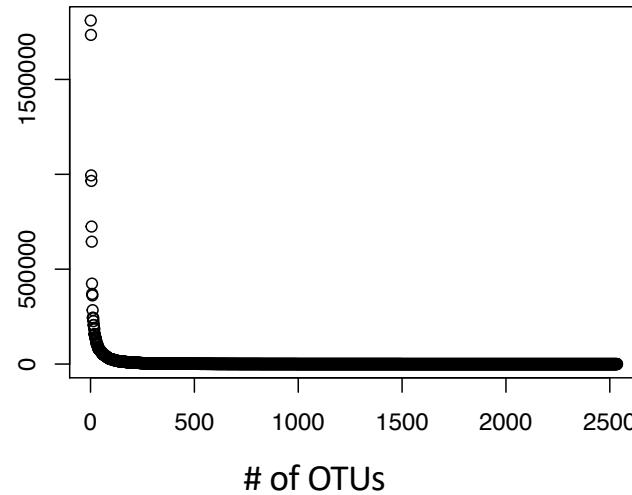
(288 incl. replicates & -ve control)

16S = 2519 OTUs = **2 953 094** reads



16S = **5 607 492** read, post sdm

ITS = 2458 OTUs = **17 990 661** reads

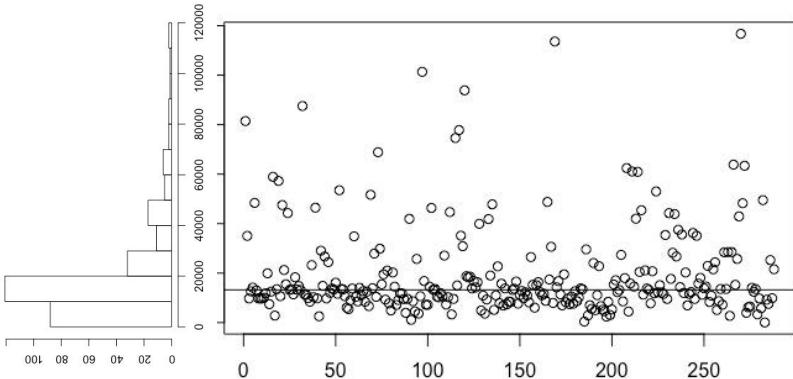


ITS = **19 809 305** reads, post sdm

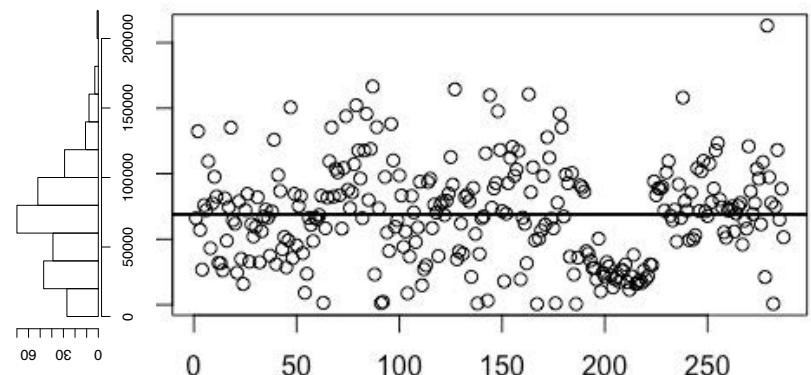
# Number of sequences per sample

(288 incl. replicates & -ve control)

16S = 5 607 492 reads, post sdm



ITS = 19 809 305 reads, post sdm

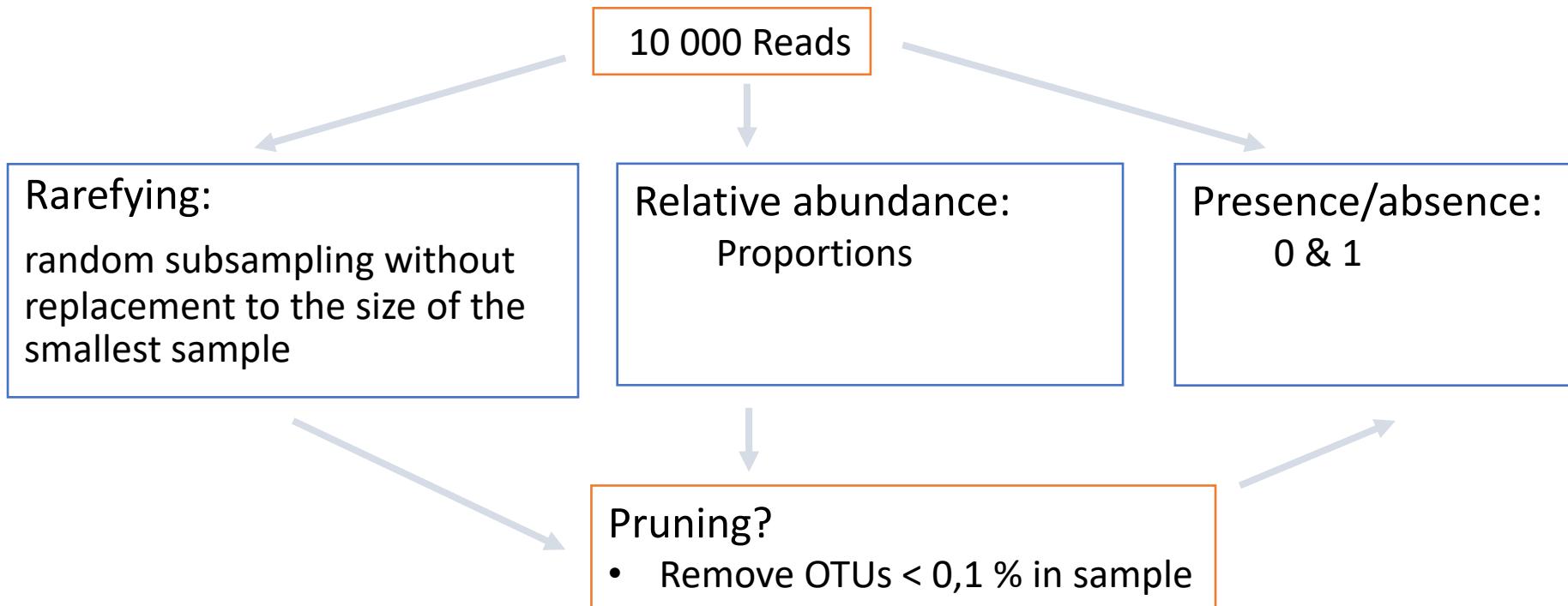


Min.	1 <sup>st</sup> Quar.	Median	Mean	Max.
51	9464	13238	19470	116727

Min.	1 <sup>st</sup> Quar.	Median	Mean	Max.
548	39006	69509	69022	213033

Conclusion: Large amount and high quality bacterial & fungal sequences

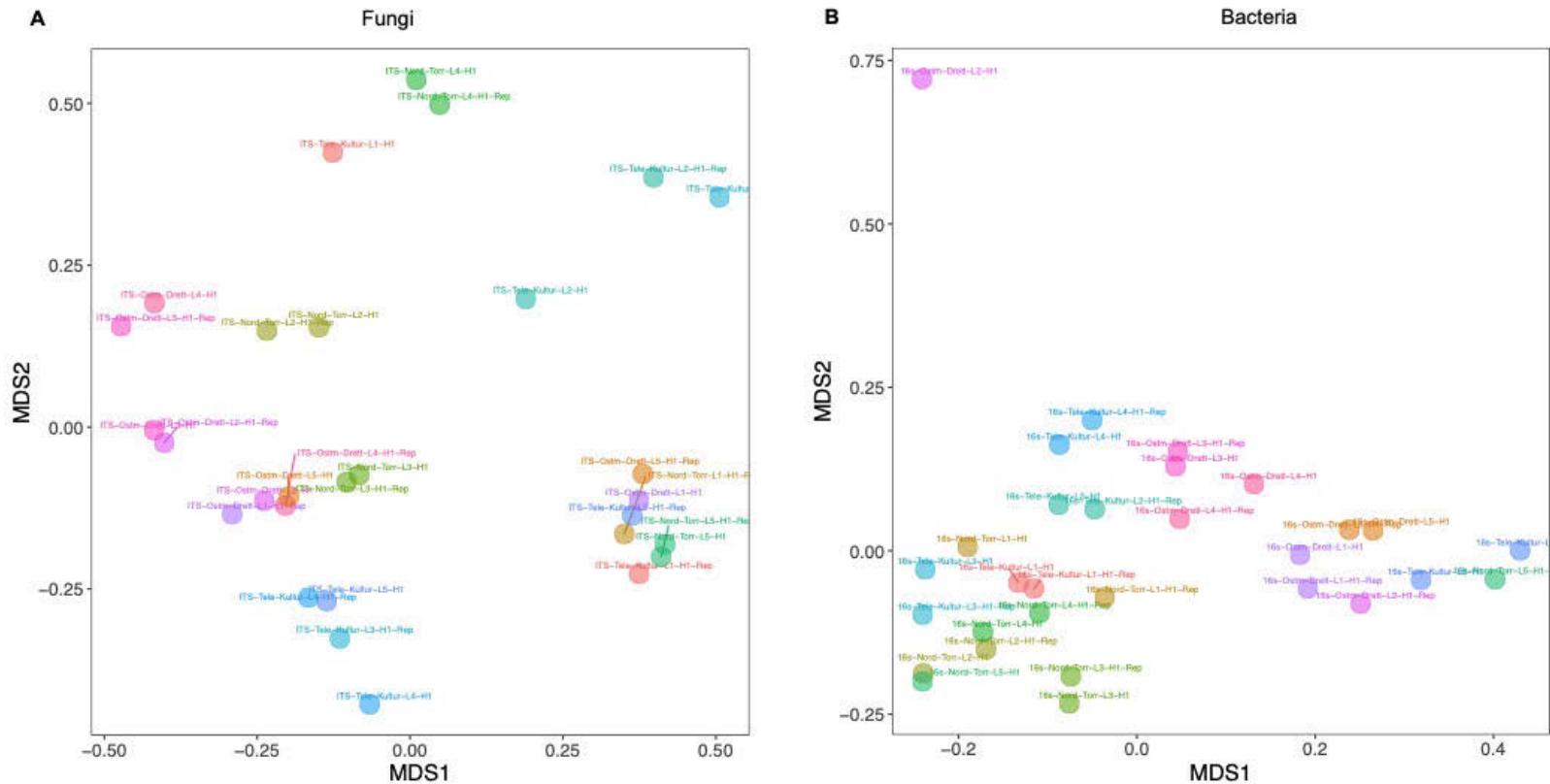
# Normalisation methods



Rarified + Pruned  
Protest correlation in procrustes's rotation

# Are PCR replicates more similar to each other

-> 15 replicates (Culture forests, hole 1 for each of the 5 logs)



\* Excluding one outlier 16S-Ostm-L2-H1 (29 points)

## Take Home tips-3

- ❑ Play & learn from your data, optimize threshold/filtering parameters ... but learn to say STOP and move on with analyses
- ❑ Try to understand what you are doing
- ❑ Do the analyses in One-Go
- ❑ Account & correct tag hopping (0-10 % revealed!)

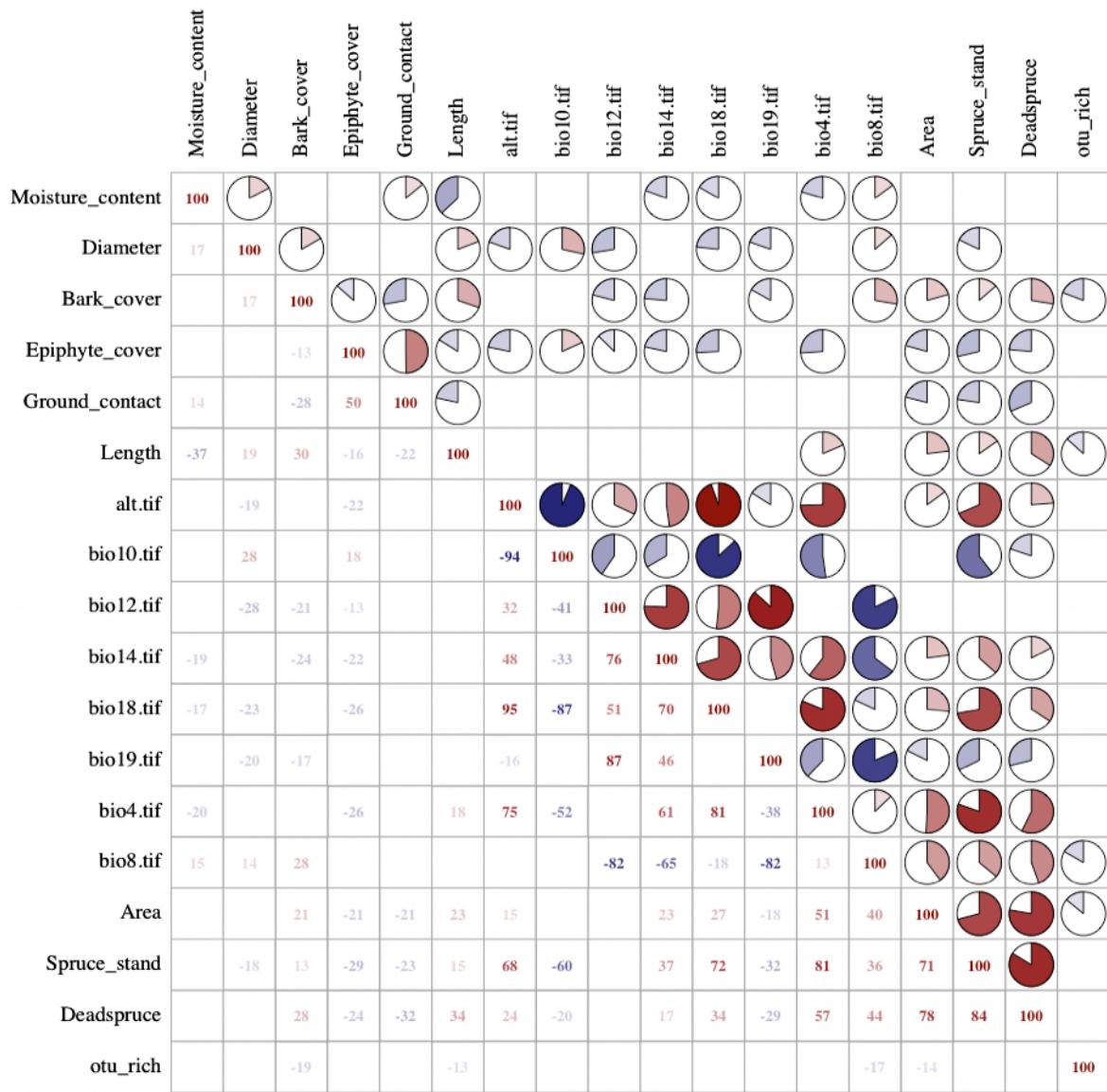
Tag switching (syn. hitch-hacking, jumping, swapping, hopping) can be a big issue on Illumina, evidence with known mock community (see Pauvert al., 2019)

Try to make **statistical** sense of your data

# Fungal & bacterial communities in dead spruce

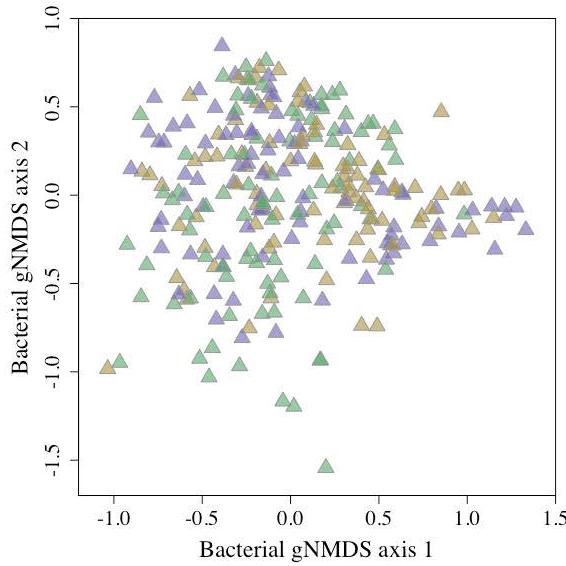
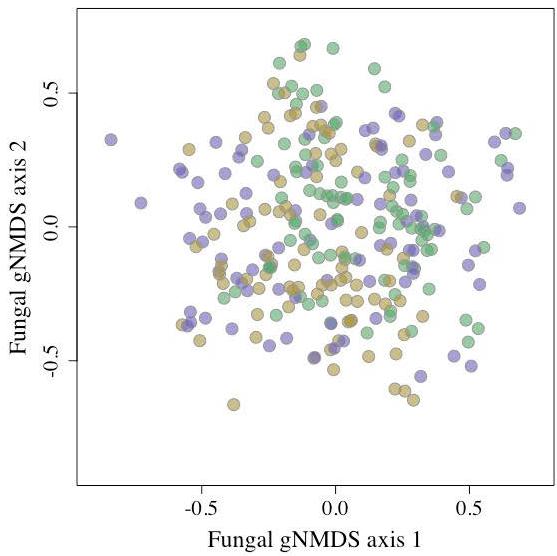
1. Are there differences in community composition and diversity, for both bacteria and fungi, due to forest management practices and/or landscapes?
2. What are the major variables that correlate with community composition?
3. ~~How are fungi and bacteria distributed along a log?~~

# Environmental variables- bacterial communities

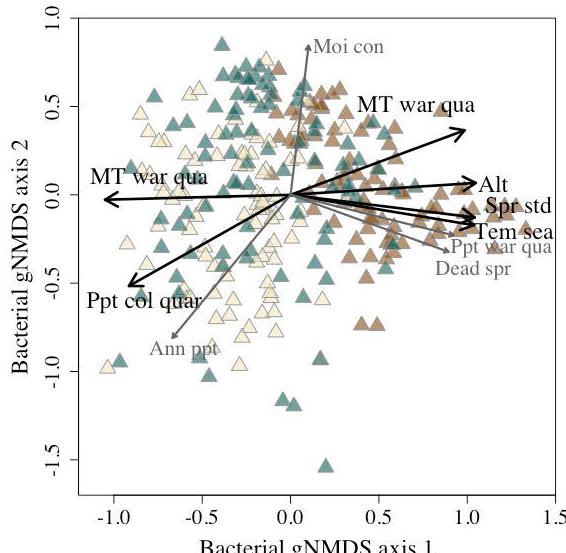
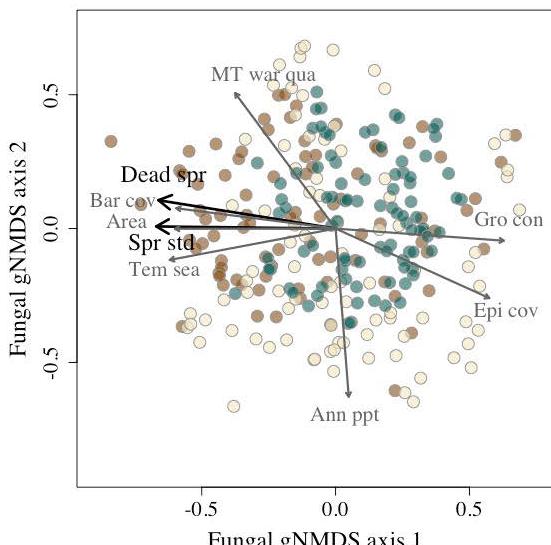


- High collinearities still exist Bio10 (Mean Temp. of Warm. quar), alt, Bio 18= Ppt. of Wat. Quart.
- Select the ones that make sense biologically/ecologically

# Main drivers of community structure



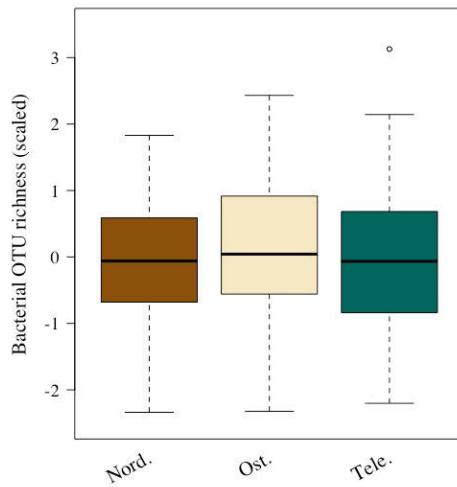
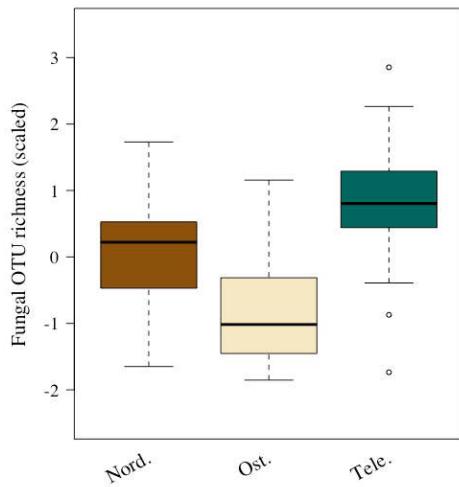
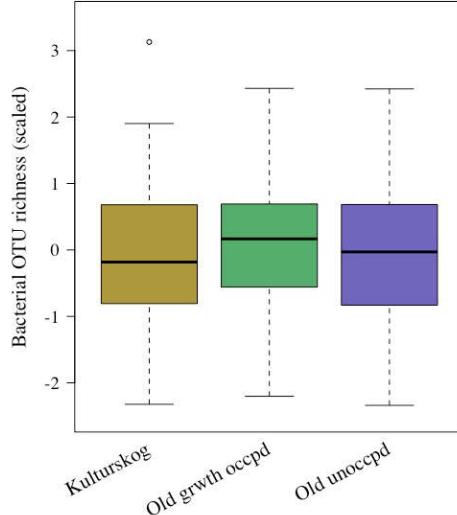
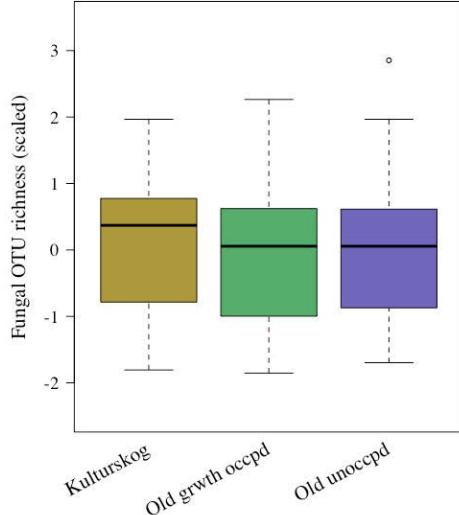
Differences in the environmental and landscape correlates of fungal communities on logs



Selection of shaping variables based on significance:

- marginally significant  
 $0.2 \leq |\text{Kendall } \tau| < 0.3$
- significant  
 $|\text{Kendall } \tau| \geq 0.3$

# Forest management vs. Landscape

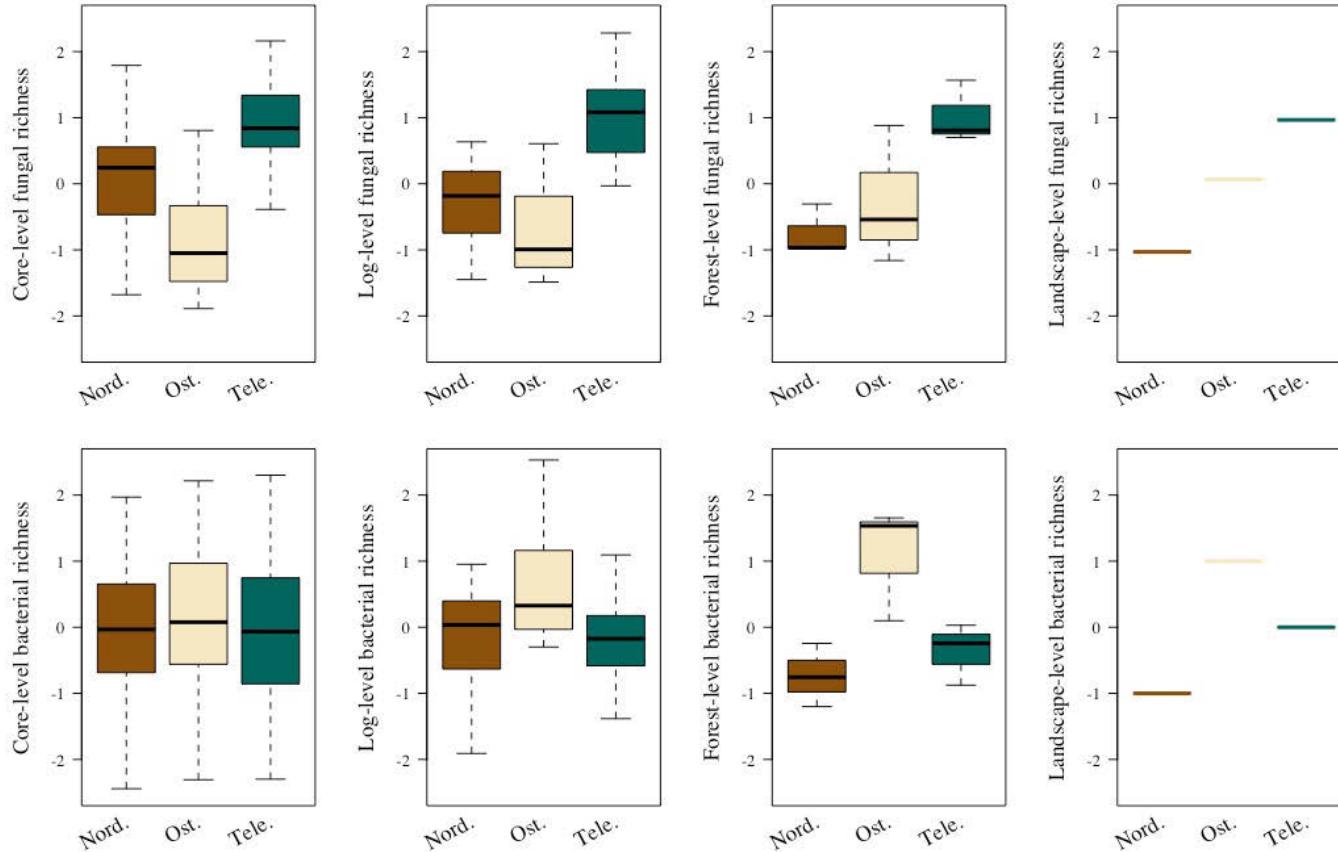


Forest management patterns are not different in terms of richness

Differences in richness were more marked across landscapes for fungi

# Diversity across spatial scales: from hole to landscape

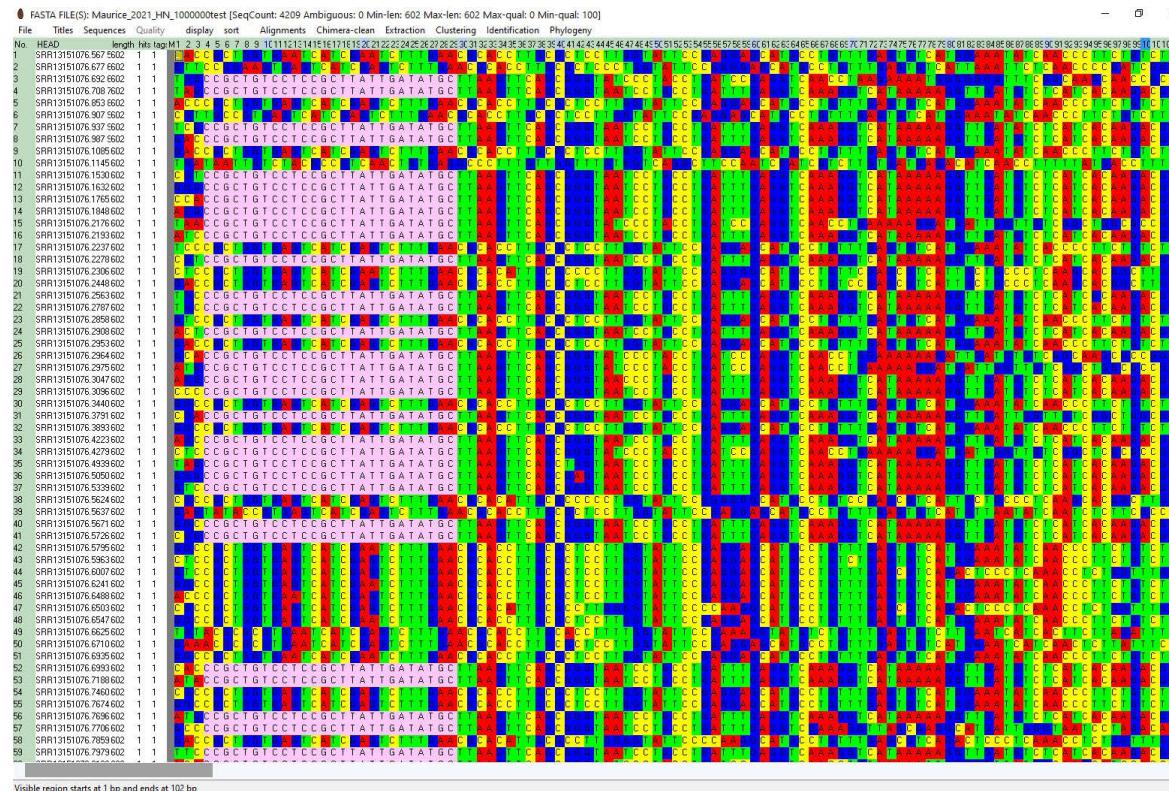
## Fungi



## Bacteria

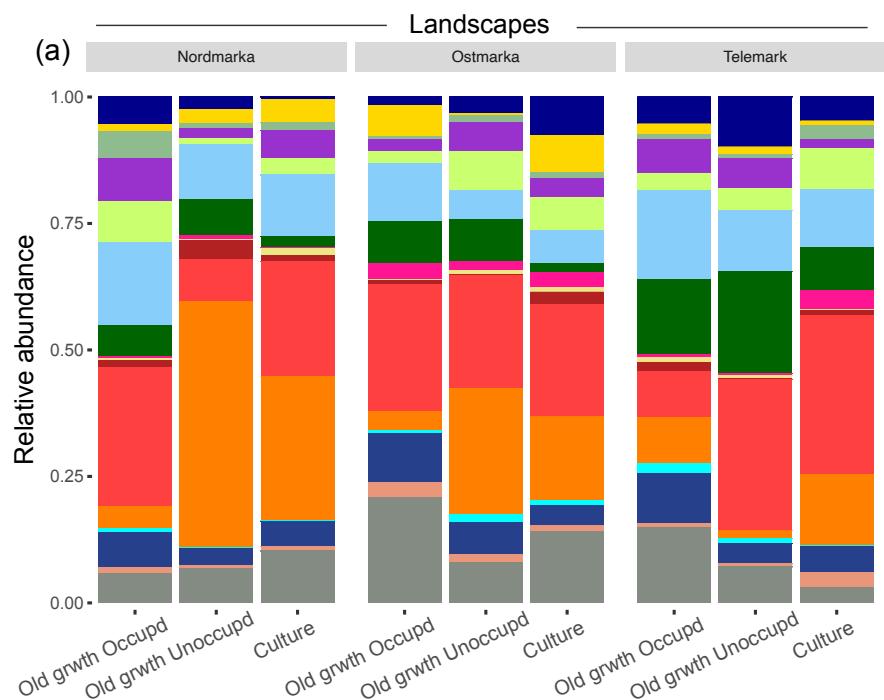
Total richness	Hole	Log	Forest	Landscape
Fungi (1602 OTUs)	57 to 656	486 to 882	946 to 1226	1200 to 1354
Bacteria (1690 OTUs)	84 to 766	474 to 1161	1267 to 1457	1507 to 1597

Try to make **biological** sense of your data

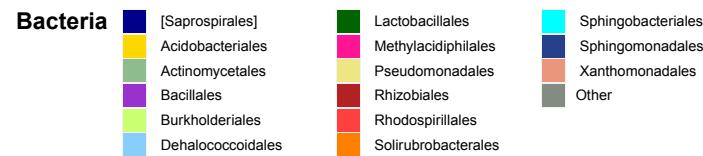
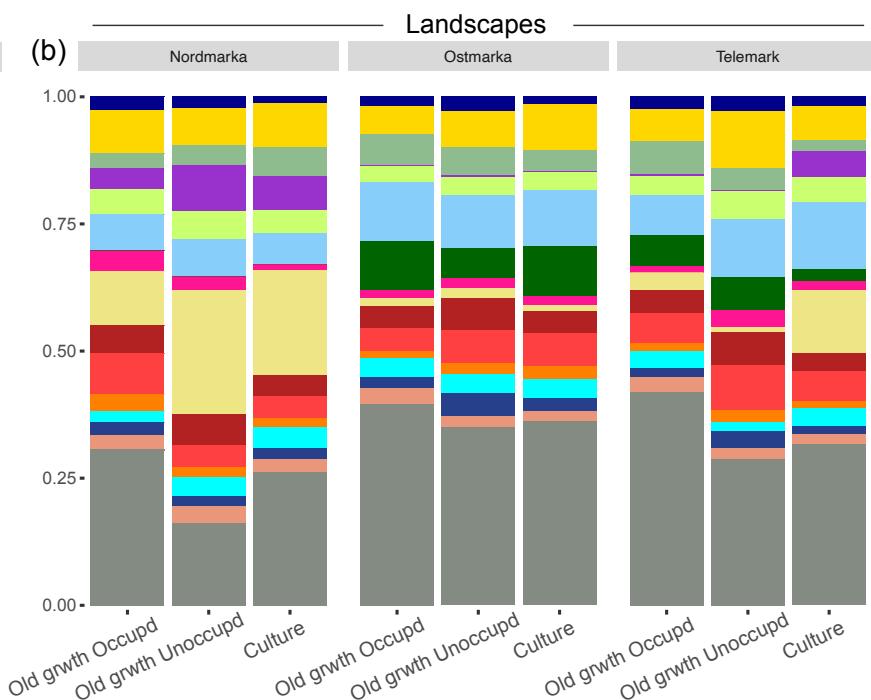


# Community Composition

## Fungi (1602 OTUS @97%)



## Bacteria (1690 OTUS @97%)



# Fungi

- ❑ Wood decay fungi
- ❑ Lichen-associated fungi present in wood
- ❑ Knowledge about patterns of fungal growth in the logs

# Bacteria

- ❑ Make wood more water permeable
- ❑ Bacteria with direct decomposition ability
- ❑ Bacteria that stimulate fungal decomposition
- ❑ Bacteria that inhibit fungal decomposition

# Satisfaction versus Frustration

1-0

or 1-2?

## Frustration ....

On-site inventory of these 45 logs

- ❑ 107 polypores fruiting bodies were recorded
  
- ❑ Dried/dead specimens: 49 genus for a total of 268 fruit bodies including polypores & corticioids

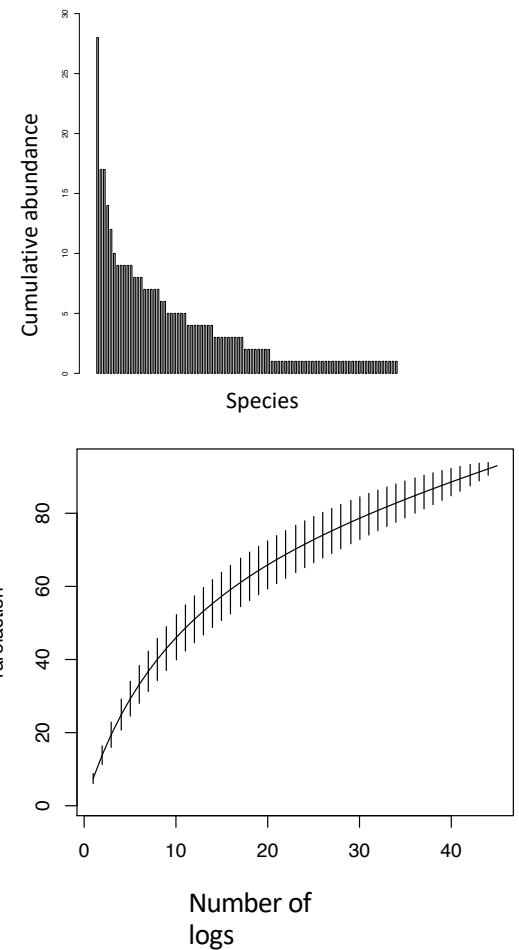
Not always very easy to use all compiled metadata

# Perspectives

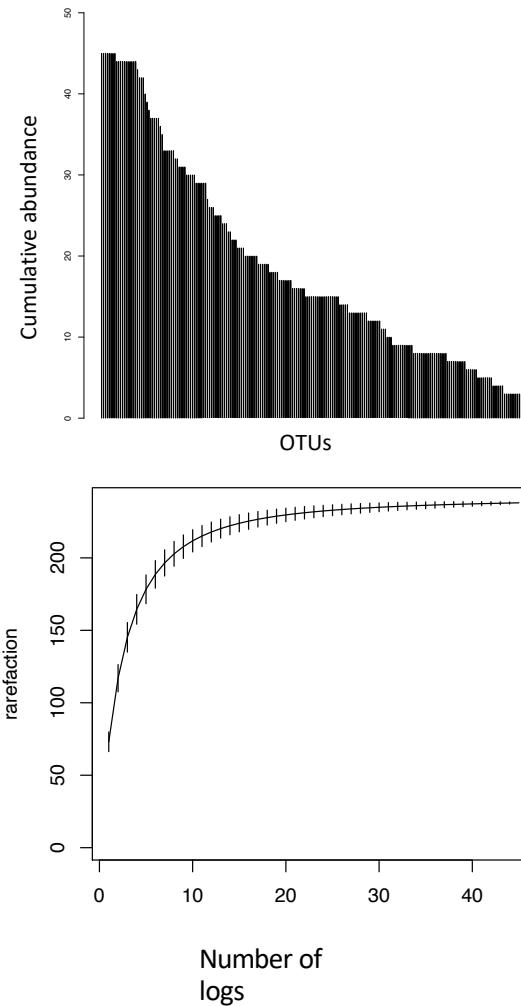
Generate some hypotheses, but...

# Perspective-1

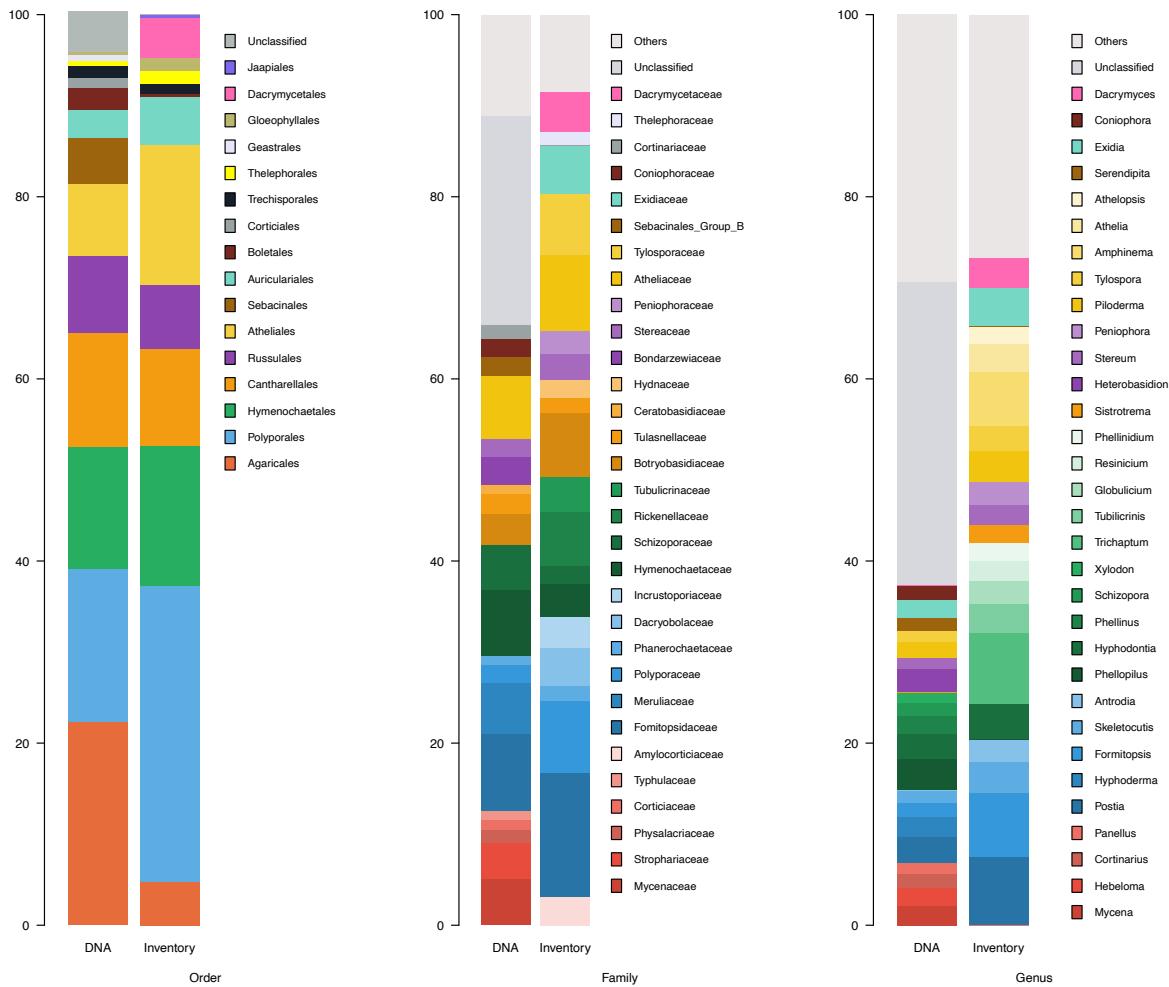
Sporocarp-based inventory



DNA metabarcoding



# Perspective-1



## Perspective-2

Interaction between the bacterial & fungal species?

2.1 Spatial ecology to understand microbial communities disperse and interact

2.2 Are there any co-occurrence patterns observed?  
Characterisation of the relationship

2.2 Correlation of main microbial taxonomic group with potential functions in ecosystem functioning, both for bacteria & fungi in wood

## Take Home tips-4

- ❑ Tag switching is a big issue on Illumina
- ❑ Be cautious with removal of outliers
- ❑ Generate some hypotheses
- ❑ Downscale data output
- ❑ Do not let the knowledge to rest & rust

Be critical and let frustration speaks at times

A photograph of a forest floor covered in green moss and fallen tree trunks. Several tall, thin trees stand in the foreground and middle ground. The background shows more trees and some green foliage.

Thank you for your attention