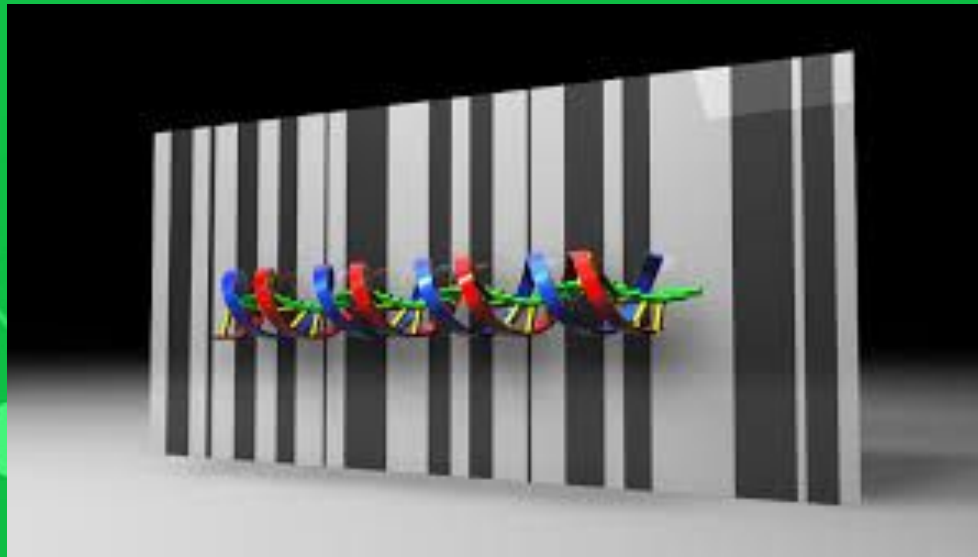


# Metabarcoding for diet analysis and environmental DNA – day 3



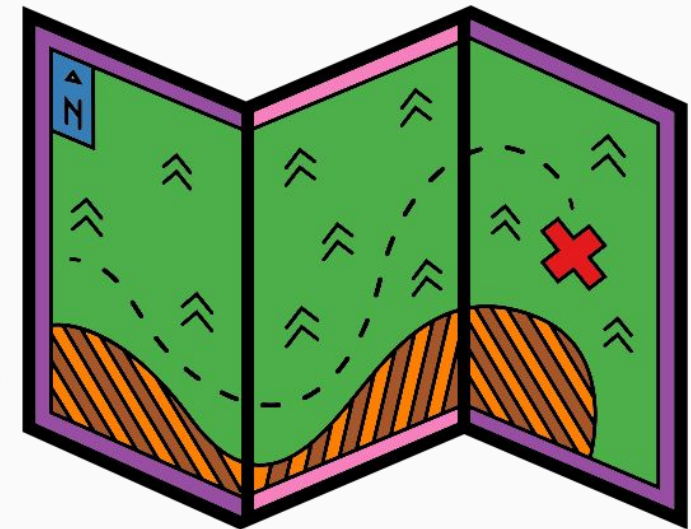
UNIVERSITY OF  
**LIVERPOOL**



The University  
Of Sheffield.

# Recap & summary: dada2 pipeline

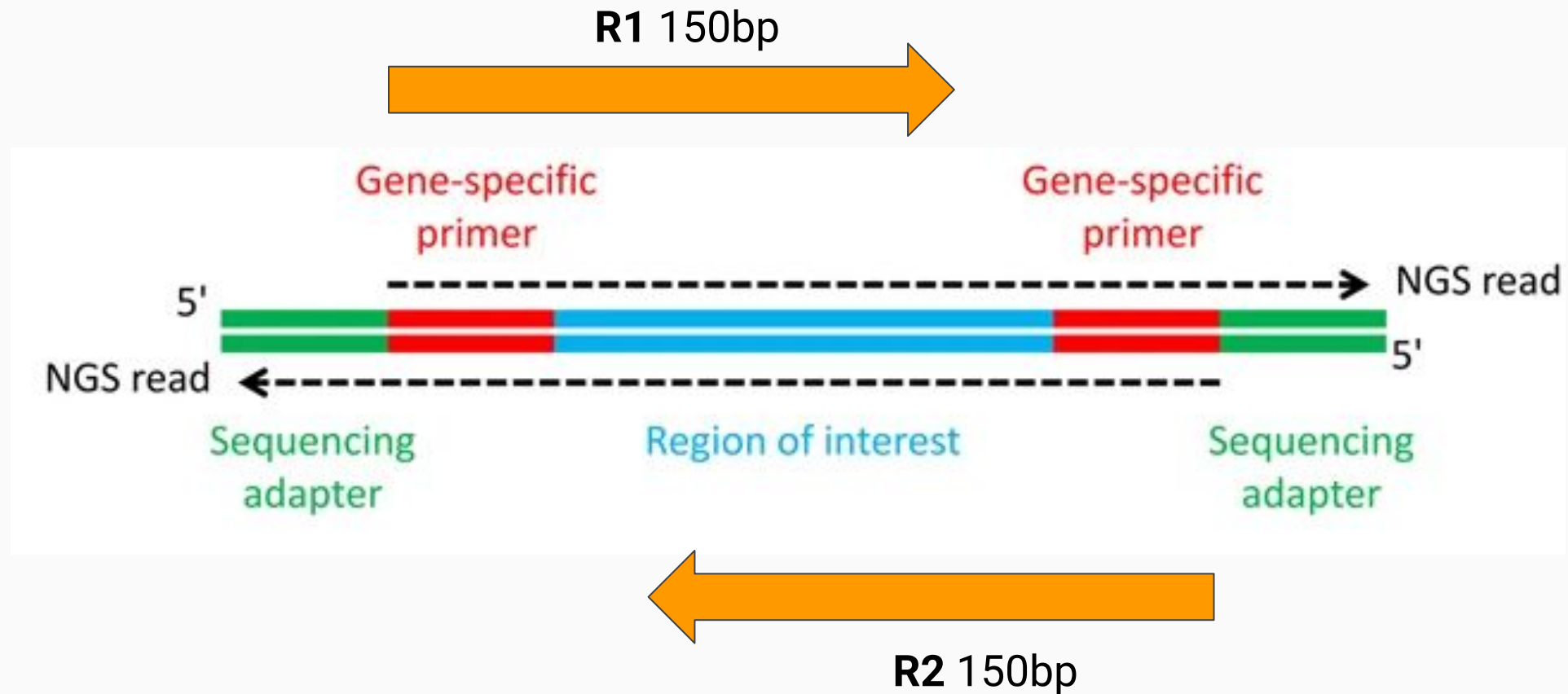
- Primer removal - cutadapt
- Quality assessment & trimming
- Error models
- Merging, chimera removal & sanity check
- Taxon plots
- Diversity analyses



# Primer removal - cutadapt



MiFish-U amplicon length is ~170bp



# Primer removal - cutadapt



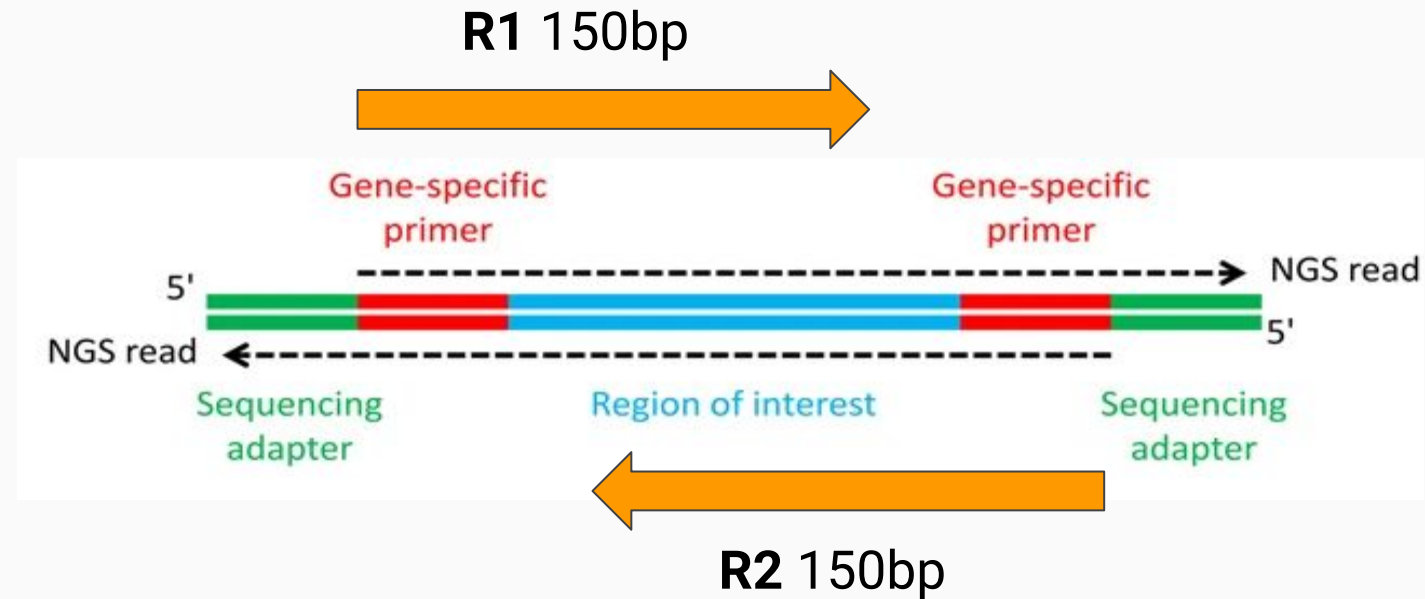
MiFish-U amplicon length is ~170bp

Cutadapt options:

-g: primer 5' end

-a: primer 3' end

-G and -A: for  
paired-end data



Forward primer FWD

5' -GTCGGTAAACTCGTGCCAGC-3'

Reverse primer REV

5' -CATAGTGGGGTATCTAATCCCAGTTTG-3'





# Primer removal - cutadapt



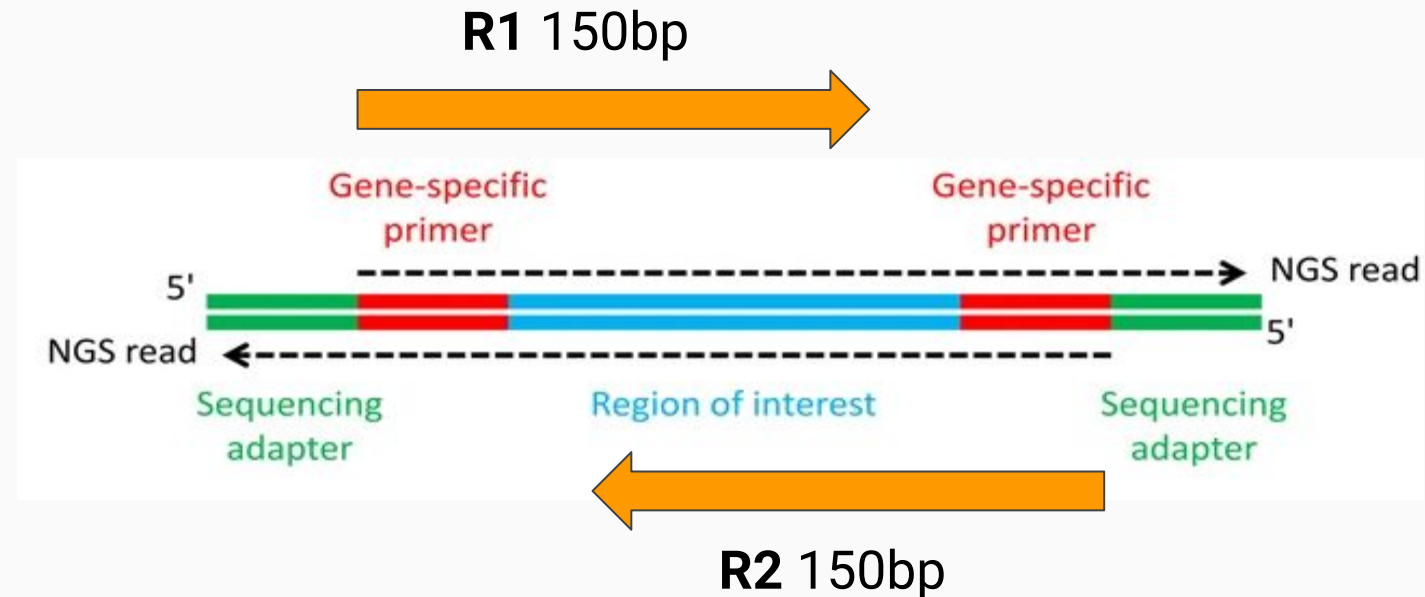
MiFish-U amplicon length is ~170bp

Cutadapt options:

-g: primer 5' end

-a: primer 3' end

-G and -A: for  
paired-end data

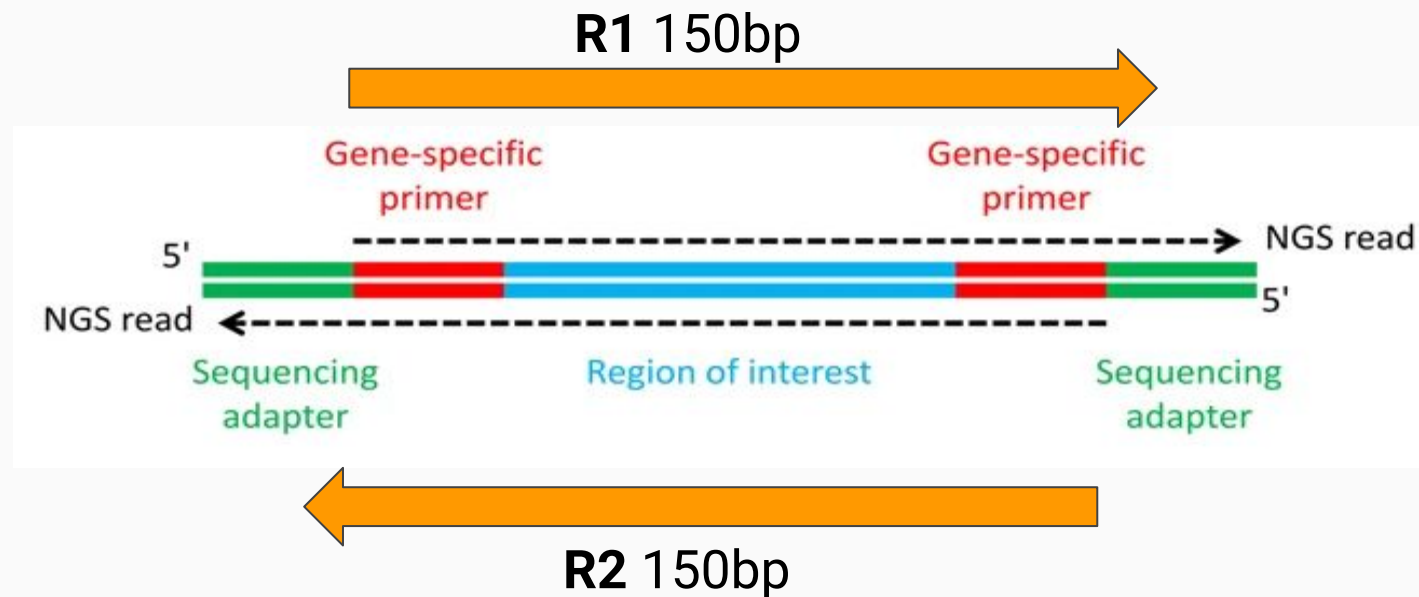
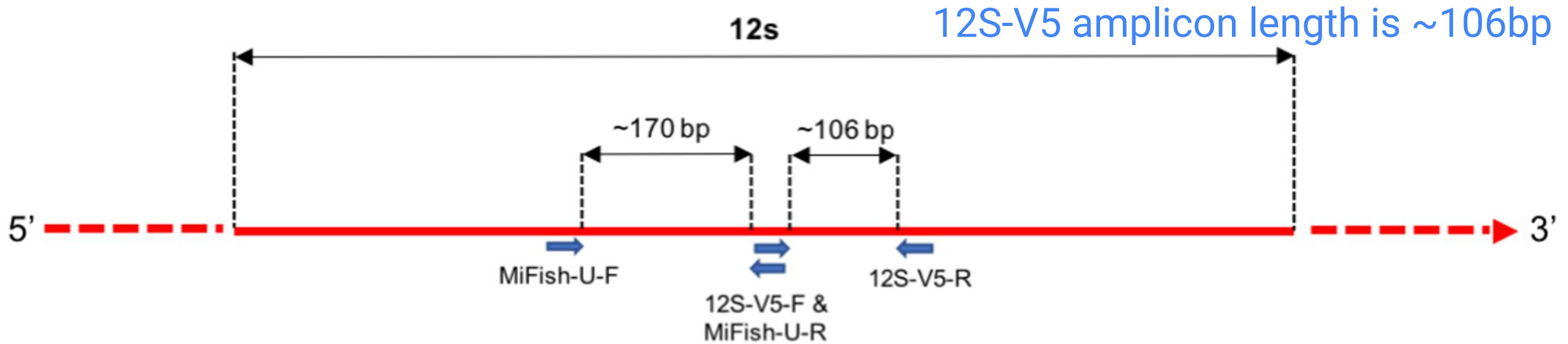


```
# Trim FWD and the reverse-complement of REV off of R1 (forward reads)
R1.flags <- paste("-g", FWD, "-a", REV.RC)

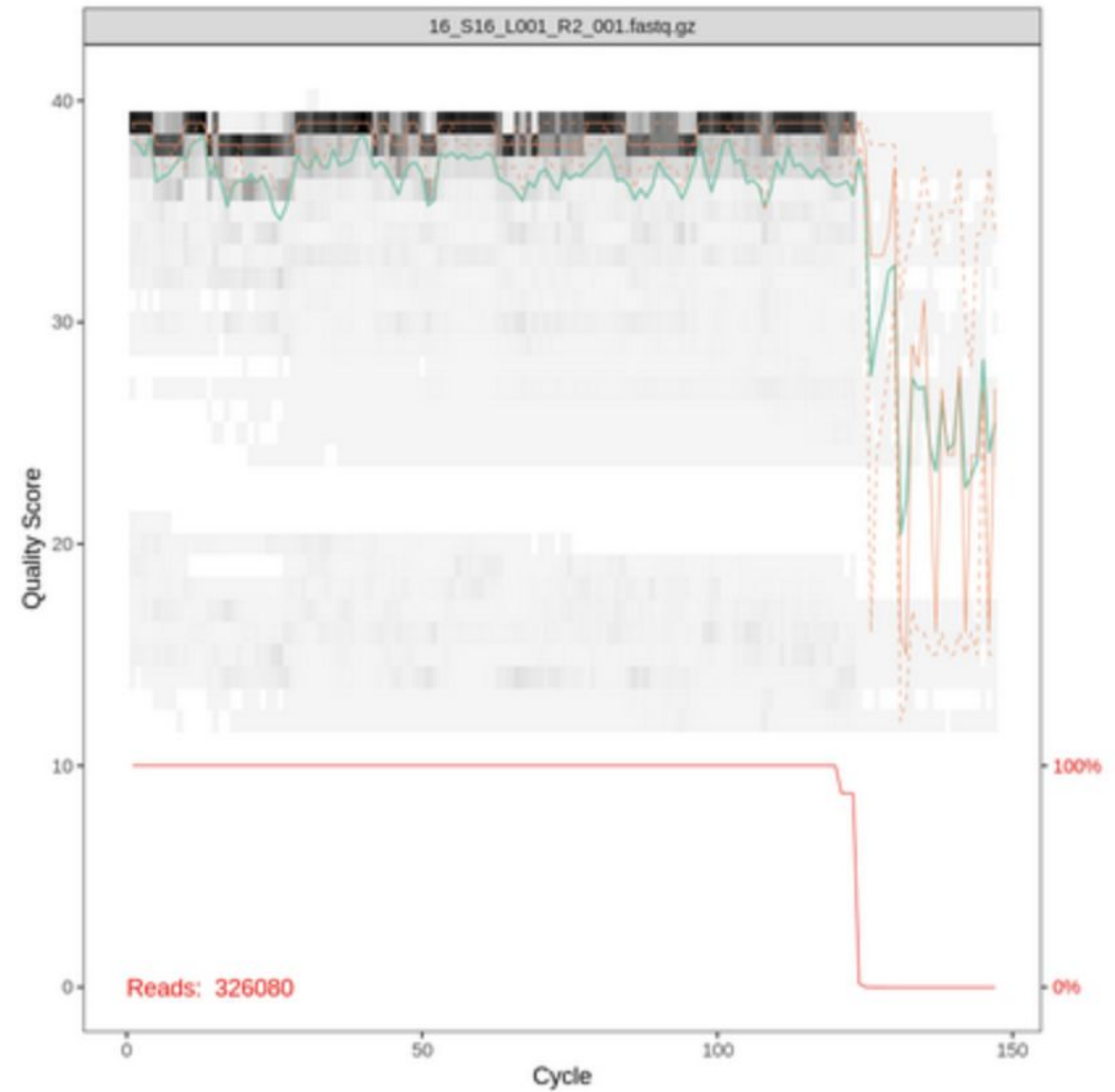
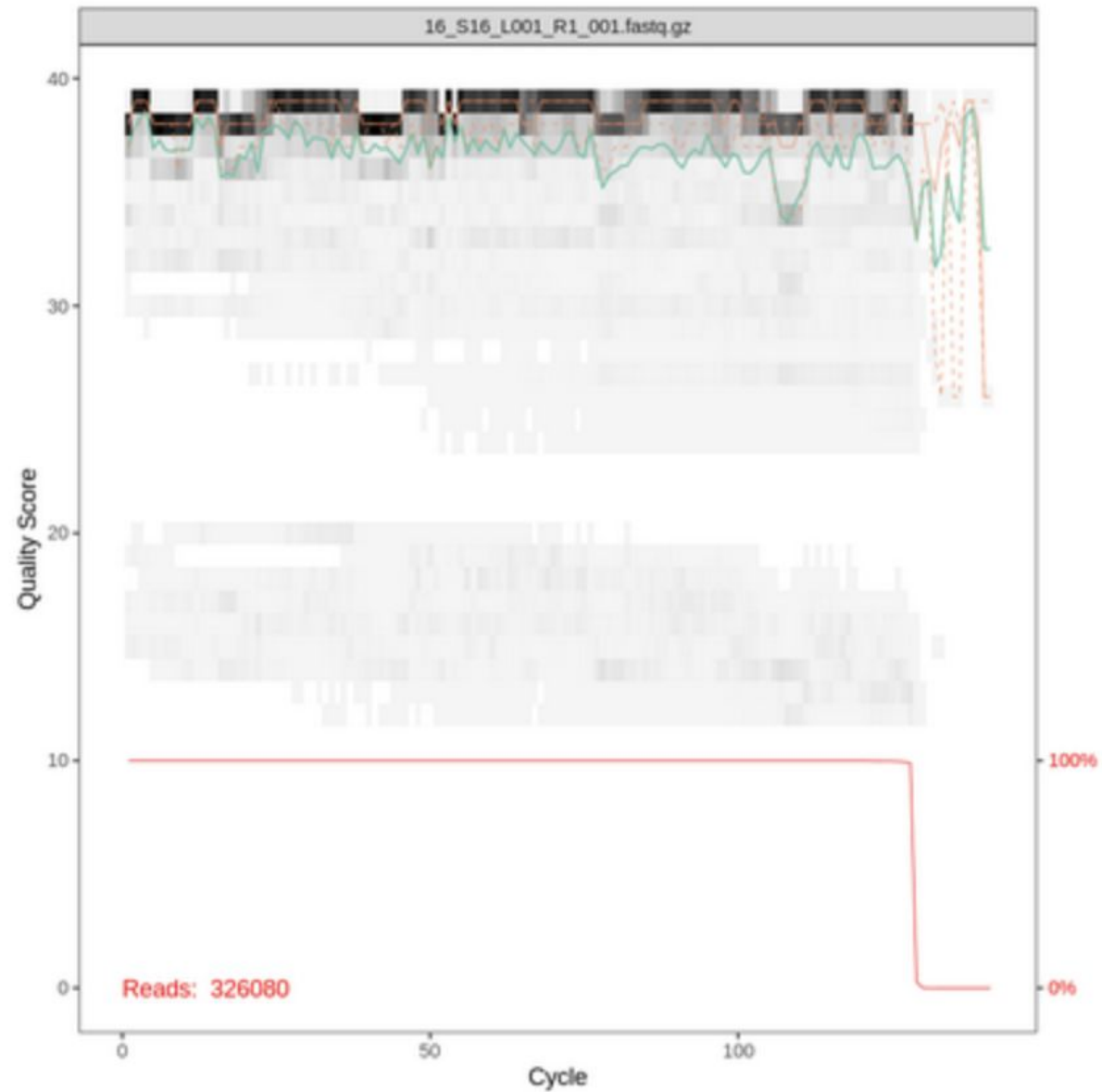
# Trim REV and the reverse-complement of FWD off of R2 (reverse reads)
R2.flags <- paste("-G", REV, "-A", FWD.RC)
```



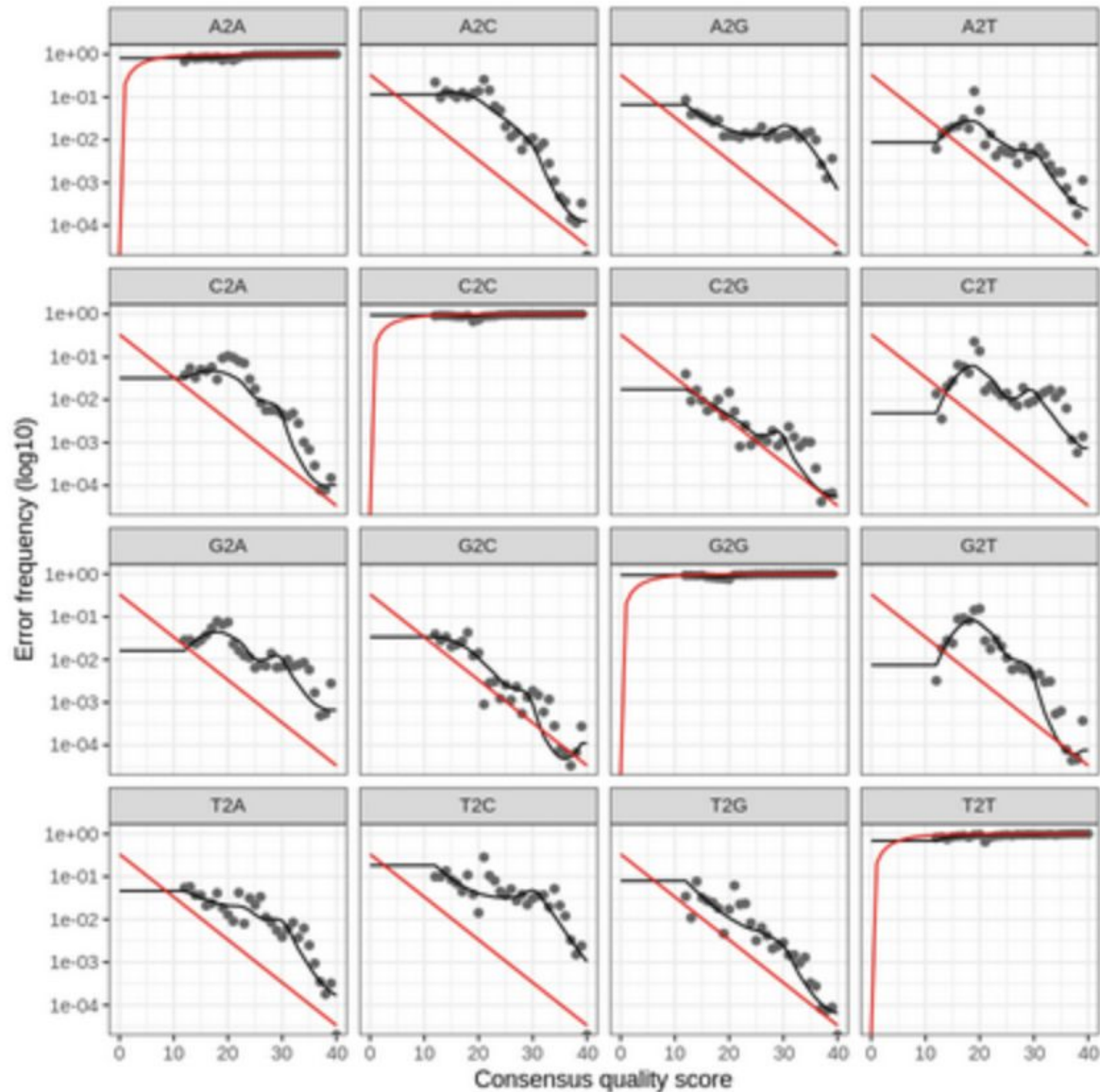
# Primer removal - cutadapt



# dada2 quality plots



# dada2 error model plots



- Red line - expected based on the quality score
- Black line - estimate
- Black dots - observed



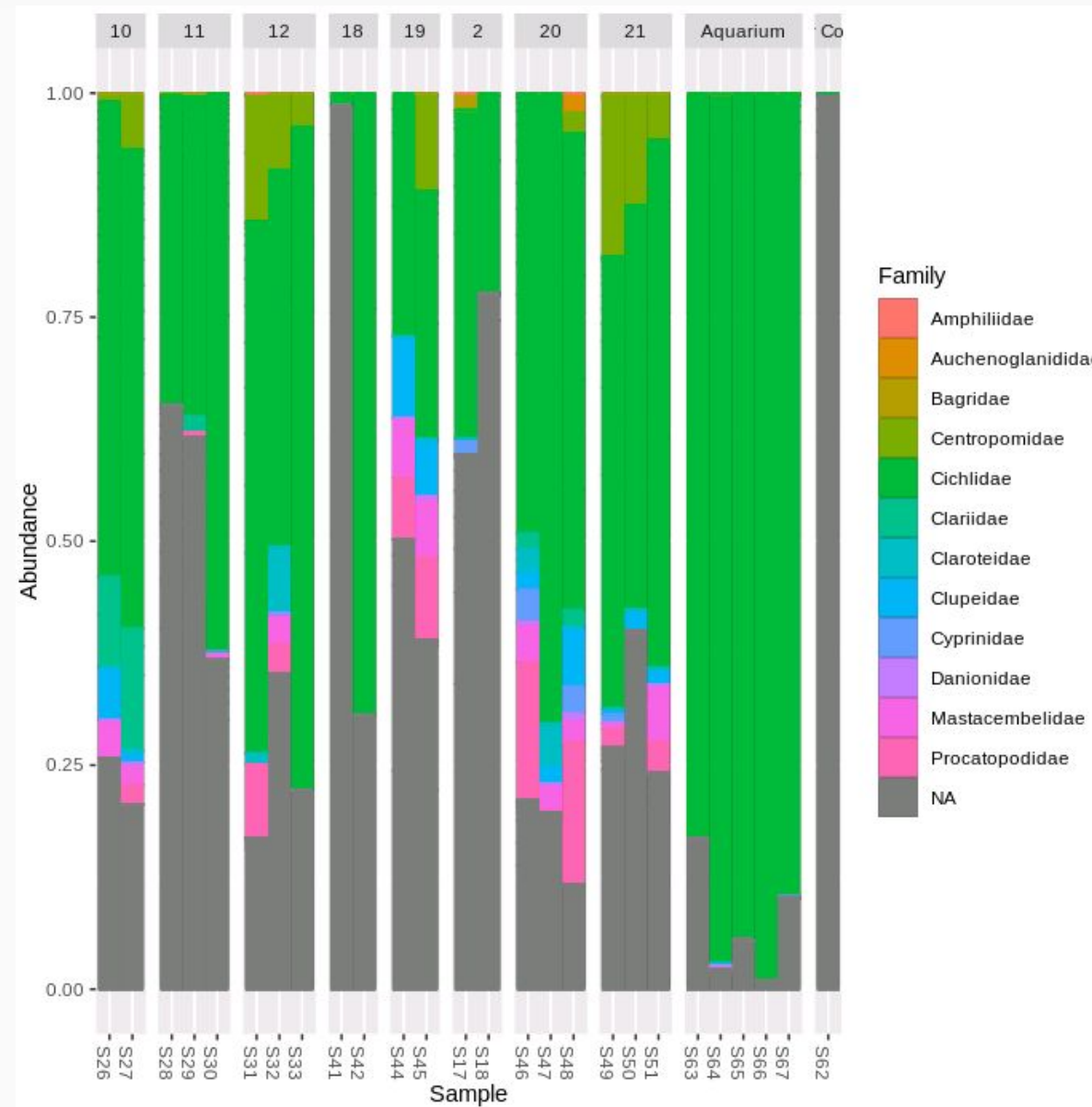


# Track reads table

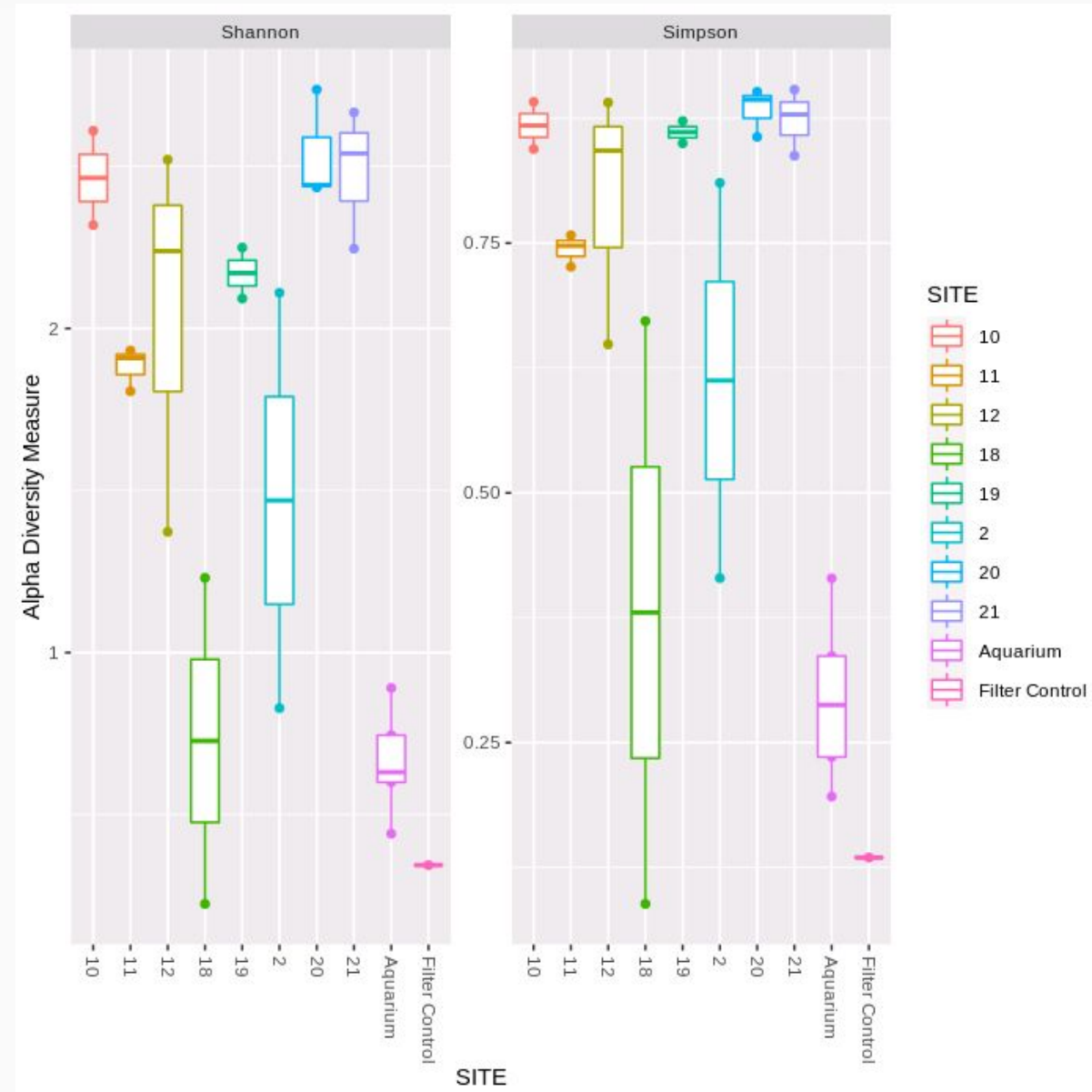


	input	filtered	denoisedF	denoisedR	merged	nonchim
<b>S16</b>	326080	323224	322464	322303	298970	258956
<b>S17</b>	47643	47305	47212	47086	45724	36541
<b>S18</b>	292556	289929	289795	289745	271083	249236
<b>S25</b>	80091	79456	79136	79249	64996	49766
<b>S26</b>	72054	71610	71240	71470	61042	48982
<b>S27</b>	88832	88026	87632	87633	70331	52770
<b>S28</b>	95054	94380	94056	94145	89173	73909
<b>S29</b>	118894	118040	117868	117854	112784	98042
<b>S30</b>	77321	76867	76356	76562	70981	60402
<b>S31</b>	159473	158262	158097	157649	152496	129603
<b>S32</b>	319652	317812	317195	317244	301715	243508
<b>S33</b>	268470	266299	265943	265813	252415	177440
<b>S40</b>	276096	265923	265485	265453	71	71
<b>S41</b>	95	89	81	80	74	74
<b>S42</b>	81764	81044	80938	80933	73528	68204

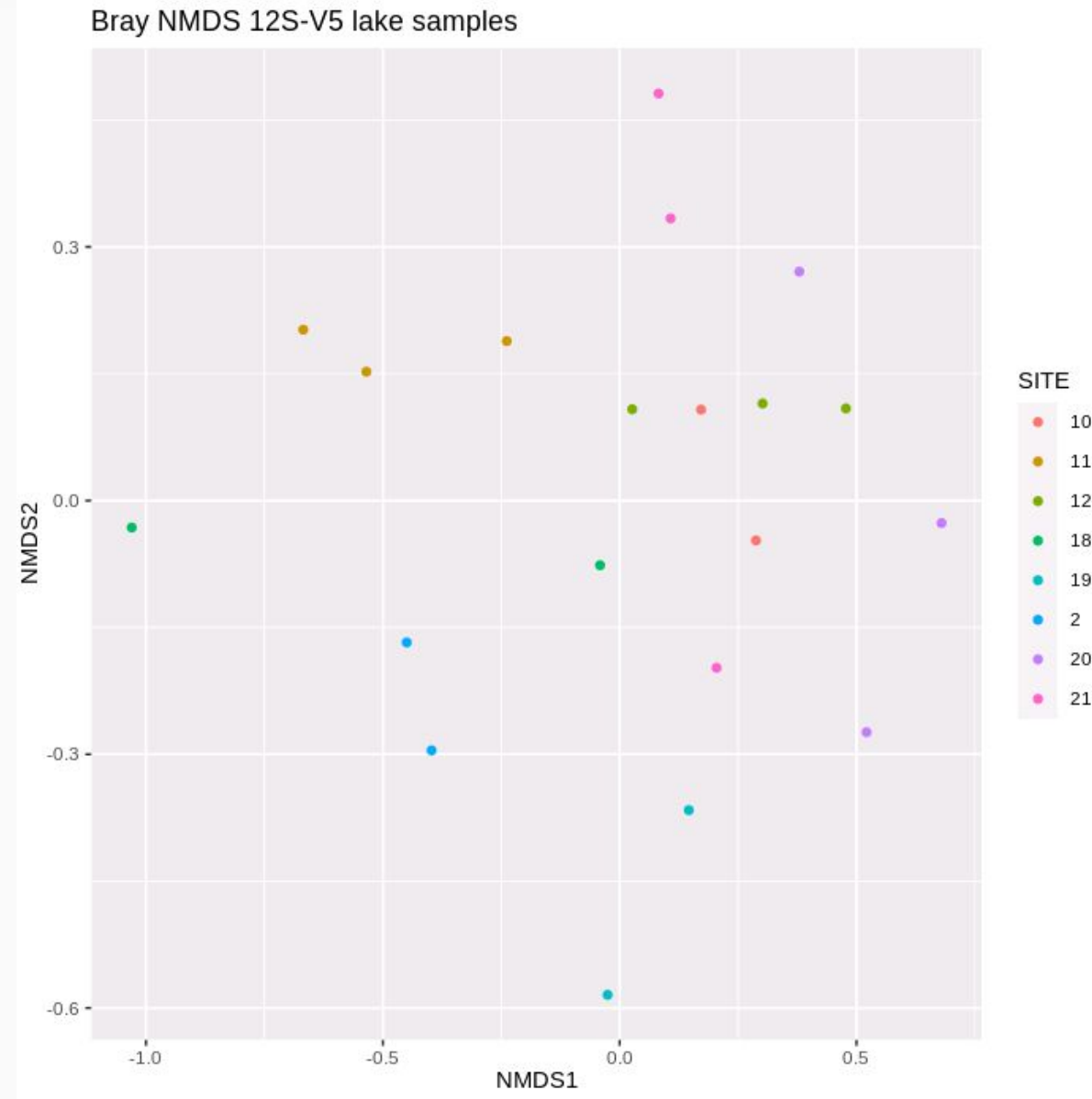
# Taxa plots: 12S-V5



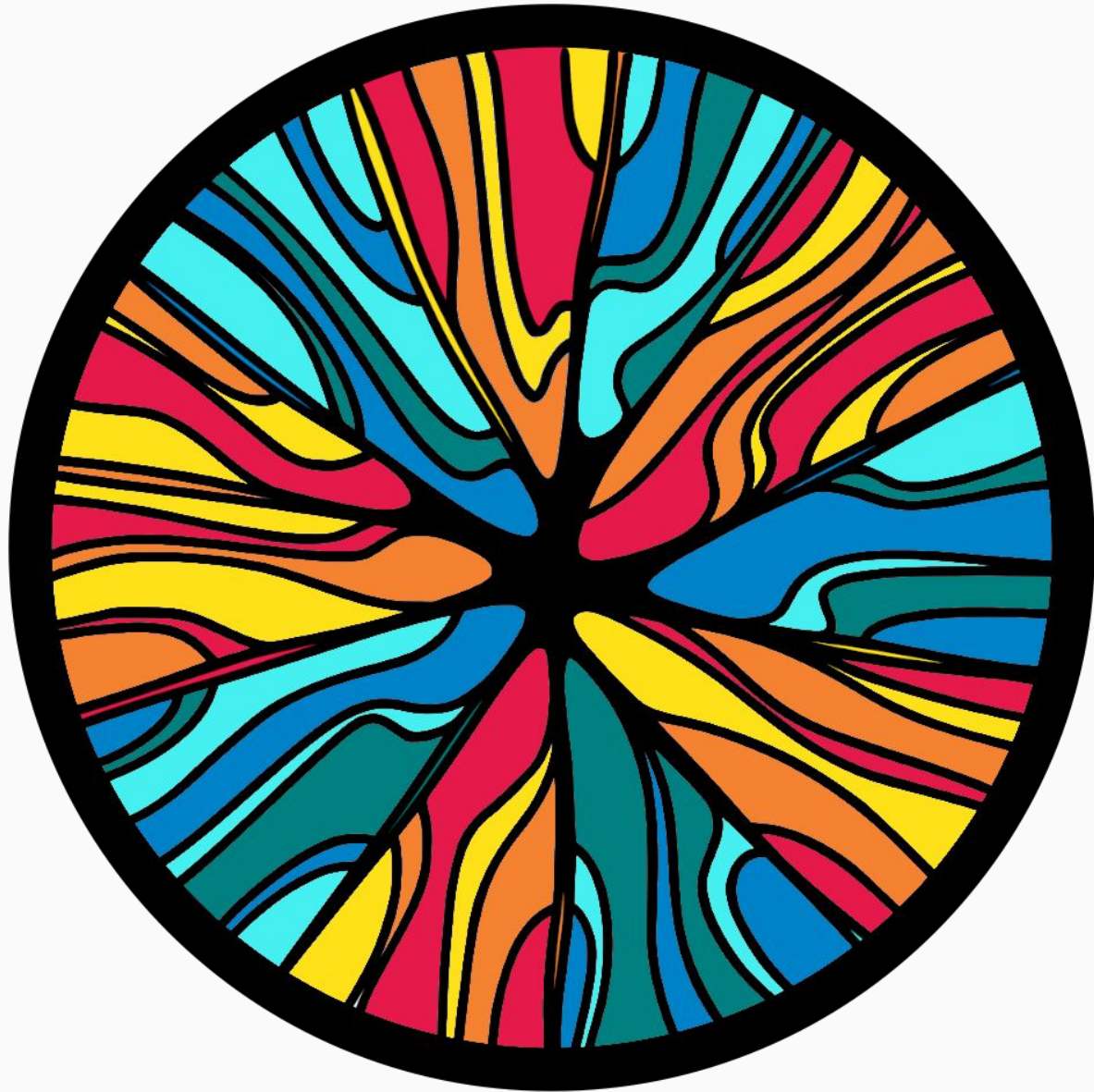
# Alpha diversity: 12S-V5



# Beta diversity: 12S-V5



# Further community analysis



- Further analysis and visualisations on community data
- Using phyloseq objects in R
- Online NEOF course
- Dates for 2024 to be announced soon!
- <https://neof.org.uk/training/>







# Thank you!

Questions?

